Changing the Paradigm: Simulation, Now a Method of First Resort

Thomas W. Lucas,[1] W. David Kelton,[1,2] Paul J. Sánchez,[1] Susan M. Sanchez,[1] Ben L. Anderson[3]

[1] *Department of Operations Research, Naval Postgraduate School, Monterey, California 93943-5219*

[2] *Department of Operations, Business Analytics, and Information Systems, University of Cincinnati, Cincinnati, Ohio 45221-0130*

[3] *College of Naval Command and Staff, United States Naval War College, Newport, Rhode Island 02841*

Received 19 December 2014; revised 24 March 2015; accepted 30 March 2015
DOI 10.1002/nav.21628
Published online 15 May 2015 in Wiley Online Library (wileyonlinelibrary.com).

Abstract: Decades ago, simulation was famously characterized as a "method of last resort," to which analysts should turn only "when all else fails." In those intervening decades, the technologies supporting simulation—computing hardware, simulation-modeling paradigms, simulation software, design-and-analysis methods—have all advanced dramatically. We offer an updated view that simulation is now a very appealing option for modeling and analysis. When applied properly, simulation can provide fully as much insight, with as much precision as desired, as can exact analytical methods that are based on more restrictive assumptions. The fundamental advantage of simulation is that it can tolerate far less restrictive modeling assumptions, leading to an underlying model that is more reflective of reality and thus more valid, leading to better decisions. Published 2015 Wiley Periodicals, Inc. Naval Research Logistics 62: 293–303, 2015

Keywords: simulation; last resort; analytical models; history of computing; simulation-modeling paradigms; simulation software; design of experiments; simulation analysis; future of simulation

1. INTRODUCTION

The belief that simulation should be a method of last resort has been entrenched since at least the 1950s. Harling [12], in an article entitled "Simulation Techniques in Operations Research—A Review," stated that "It has been often said that a simulation is a last resort." Variations of this phrase have seen continued use within the operations-research community over the years. Perhaps the best known example appears in Harvey Wagner's seminal textbook *Principles of Operations Research with Applications to Managerial Decisions* [60]. In his Chapter 21, "Computer Simulation of Management Systems"—the introductory section of which is entitled "When all else fails..."—he wrote that "Most operations research analysts look upon digital computer simulation as a 'method of last resort."' Wagner gave two primary reasons for his "gloomy attitude" toward simulation. The first was that when the simulation "includes uncertain events, the answers... must be viewed only as estimates subject to statistical error." The second reason was that if a system was so complicated that it was "beyond the reach" of traditional operations-research techniques, then "the required model-building effort and the subsequent analysis of the simulated results are likely to be difficult." As a consequence of these challenges, Wagner concludes that "computer simulation is often an expensive way to study a complex problem."

Unfortunately, the notion persists that simulation should be a last resort. Many researchers maintain the a priori attitude that analytical models are inherently superior. A relatively recent example is the paper by Parlar and Sharafali [37], which we evaluate in some detail in Section 4.

Times have changed. In the face of astounding advances in affordable processing power, modeling paradigms and tools, and supporting analysis capabilities, it should be clear that for many complex, real-world problems, simulation—done properly—should be the method of choice. After more than

Correspondence to: W. David Kelton (david.kelton@uc.edu)
Department of Defense Distribution Statement: Approved for public release; distribution is unlimited. The views expressed in this document are those of the authors and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

half a century of dramatic progress in simulation technology, it is time to retire the outdated notion that "simulation is a method of last resort." Our objective in this article is to change mindsets in support of those who are actually concerned about actually solving actual problems.

We state our fundamental position in Section 2, and in Section 3, we detail the current enabling technologies that now render our position practical and appealing. Section 4 exemplifies, in several ways, the risks associated with stylized models based on unrealistic simplifying assumptions required for analytical tractability. Section 5 looks ahead and Section 6 concludes.

2. SIMULATION: EMBRACING REALITY IN MODELING

The primary advantage of simulation is that it enables researchers to construct and study valid models of complex systems in relatively simple and straightforward ways. Put another way, simulation allows us to study problems without assumptions made principally to obtain an analytically tractable model, and that may jeopardize the veracity of the model.

We build and evaluate models (simplified representations) to gain insights into existing or prospective systems and phenomena. Ultimately, this should allow us to make better choices and improve outcomes—such as cutting costs or saving lives. Models can be abstracted as a mapping of inputs to outputs. In many cases, the mapping will be expressed as mathematical or logical representations of that transformation. If these relationships are sufficiently simple, they can be written as closed-form mathematical expressions. Such a model is called *analytically tractable* and yields exact results for output metrics of interest. Estimated results from analytically intractable models can usually be obtained via simulation [2, 23]—and increased computation power usually allows us to make the estimate extremely precise. If the models are equally valid, an exact analytical solution is, of course, preferred to an approximate simulation one. However, if the analytical tractability comes at the expense of the model's verisimilitude, then faulty insight and poor decisions may result. The imprecision of a simulation-generated estimate can be quantified and, if need be, reduced. Conversely, it is often difficult to assess the effect of unrealistic simplifying assumptions underpinning many analytical models without using simulation.

To facilitate analytical tractability, modelers commonly posit convenient probability distributions, such as the additive Gaussian or memoryless exponential. They also invoke assumptions such as linearity, deterministic relationships, homoscedasticity, identical distributions, Markovian behavior, and independence or stationarity. If these assumptions are made because they reflect the nature of the problem being examined, and they facilitate analyzing the model, then no one should quibble with them. However, it is our belief that such assumptions are often made solely to render a model analytically tractable. The consequences of such simplifying assumptions can be severe. As an all-too-common example, ignoring variance can easily lead to extremely poor decision making [28, 49].

Model validation is a central pillar within the simulation community, as evidenced by its ubiquity in the leading texts over the years [2, 5, 19, 20, 23, 26, 39]. We believe validation should be applied to *all* modeling, including analytical. A model is not *intrinsically* valid just because it is analytical. The allure of stylized analytical models, with little regard for their validity, may be partly due to the perception that analytical results are easier to get published than are "equivalent" conclusions derived via simulation [36].

Of course, the debate over analytical versus simulation approaches is not new. Discussions are in all standard simulation books such as those referenced in the preceding paragraph, in general operations-research overview texts (e.g., [14]), and in introductory simulation tutorials (e.g., [42]). There has been general agreement that simulation's rightful place is for addressing models that are too complicated to succumb to an analytical solution, and we agree. But our argument here is that, with modern hardware, modeling paradigms, software, and methods, simulation can produce results that are just as insightful and just as precise as can analytical methods. Moreover, simulation models have the undeniable advantage of not requiring overly-restrictive, and thus less realistic, modeling assumptions. Thus, the historically strong preference for analytical over simulation approaches is no longer warranted. With a well-done simulation study there is (now) nothing really lost in terms of insight or precision, but (as always) much to be gained by the opportunity to work with a model that is more valid, even if more complicated.

3. ADVANCES IN SIMULATION TECHNOLOGIES

Since the first utterance that simulation is a "method of last resort," there have been quantum advances in computing capabilities, modeling paradigms, and software environments for building simulation models. Likewise, both the methods and the software for statistical design and analysis have dramatically improved, allowing us to evaluate far more scenarios than ever before, and in statistically convincing ways. These parallel movements, taken together, imply that simulation is now an option that should be, in many ways, regarded as the method of choice for analyzing complex systems.

In Section 3.1 we explore the implications, for simulation purposes, of the dual facts that today's analyst has billions of times more processing power at her or his fingertips than all of the world's simulation pioneers combined had sixty years ago, and that these astronomical increases in capability are available at several orders of magnitude less expense. In Section 3.2 we explore modern foundational paradigms for simulation modeling, which have led to much-improved simulation software. In Section 3.3 we describe the salient elements of those resulting software environments for simulation modeling and execution that facilitate developing, testing, and designing/analyzing simulation experiments in ways that provide as much or more insight as do exact analytical models—and stem from more-valid modeling assumptions. In Section 3.4 we document the revolutionary advances in our ability to explore computational models efficiently using new and extremely efficient experimental designs developed specifically for this purpose. Once the experiments are run and the output data collected, modern data-mining and visualization software enables researchers to identify patterns interactively in complex, high-dimensional explorations that were inconceivable in simulation's early days. Finally, in Section 3.5 we sum up this confluence of events that we believe mandates a fundamental mindset change in the way analysts choose tools and, in particular, view the simulation tool as an option.

3.1. A Trillion-Fold Improvement in Computing

Simulation proved its value long before the invention of electronic computers. Perhaps the most famous early example is Buffon's 1777 "needle experiment," in which needles were manually tossed on a plane ruled with parallel lines to estimate π [7]. Another early use of manual simulation was performed by Gosset in the early 20th century, investigating the deviations from normality that constitute Student's-t distribution [53]. Shewhart [52] drew samples of marked chips from bowls to derive control charts—ushering in the field of statistical quality control, which has had tremendous impact on manufacturing around the world. Later, von Neumann used humans as computational units for performing Monte Carlo-based estimation during the Manhattan Project [27].

Prior to World War II, computation was done by humans, although increasingly with the assistance of electromechanical machines. For calculation-intensive efforts during the war, such as for the Manhattan Project and in developing the U.S. Army's Ballistic Research Laboratory's (BRL's) firing tables, computers were people—mostly highly educated women [27]. To increase the speed and accuracy of the BRL's calculations, the University of Pennsylvania developed the first modern computer, known as the Electronic Numerical Integrator And Computer (ENIAC) for a cost of \$500,000 (roughly \$6 million in today's dollars) [6]. The ENIAC was formally accepted by the U.S. Army in July 1946.

As was typical of early computing devices, ENIAC was large and cumbersome to use. The 30-ton behemoth had 18,000 vacuum tubes and occupied 1800 square feet of floor space. The ENIAC team consisted of "about 50 people" [41]. Furthermore, it required 150 kilowatts of power to operate. Researchers interacted with ENIAC via punch cards. If that was not enough, ENIAC had severe reliability issues, with a tube failure "about every two days" [41]. In 1954, the longest period of continuous operation was less than five days [15]. Despite all of those hardships, ENIAC provided scientists with a revolutionary new capability. It could add 5000 10-digit numbers per second and multiply 400 of them per second—roughly 1000 times faster than any previous device. Even with having a data-storage capacity of only 20 accumulators with 10 signed decimal digits each, ENIAC could solve previously unsolvable problems required in designing a hydrogen bomb, calculate "the path of a shell's trajectory faster than the shell could fly," and determine "2000 digits of π in only 70 hours" [6].

It is no wonder that early researchers such as Harling found simulation difficult, problematic, and prohibitively expensive. However, since that period, there have been amazing advances in processing speed, computer memory, reliability, and data storage—all at ever more affordable costs. This phenomenon is best exemplified by "Moore's Law." In 1965 Gordon Moore [33] predicted that the number of transistors per chip would continue to increase by "a factor of two per year." Later, Moore [34] revised the growth rate to "approximate a doubling every two years." Not only has Moore's law roughly held true [9], it now drives progress as industry regards it as a standard. As of this writing, the Tianhe-2, developed by China's National University of Defense Technology, is listed as the world's top supercomputer, with a performance of over 33 petaflops (quadrillions of floating point operations per second) on the LINPACK benchmark [25, 54]. Moreover, the amount of data storage per unit cost since the 1950s has doubled approximately every 14 months [13, 22]. All told, since Harling's time, there have been greater than trillionfold reductions in the cost of flops and data storage. The continuing increase in processing power is dramatically illustrated by the fact that in 2012 a \$399 Apple iPad 2 was faster than the \$16 million (in 1985 dollars) liquid cooled Cray-2 super computer, the world's most powerful computer in the late 1980s, as measured by the LINPACK benchmark [8].

3.2. New Paradigms for Modeling

Today's modelers and analysts have access to a far richer set of tools for building realistic and valid models than did Harling and his contemporaries.

Turing [56, 57] and von Neumann [59] built a foundation for modern computing that influences us to this day. The von Neumann architecture drew on Turing's insight that computing instructions were themselves data, and computation could be achieved by loading a mix of data and instructions from a memory device into registers to perform operations sequentially, negating the need for physical rewiring to achieve a different mix of computations. All programs must eventually be mapped—via compilers, interpreters, or virtual machines—to instructions corresponding to the specific architecture of the hardware on which the program will run, but this is not an easy or natural mode of thinking for human beings. Consequently, we have designed and created a variety of programming languages to facilitate the *human* expression of problem formulations and their solutions. The growth and evolution of domain-specific language and modeling paradigms over the last half century has greatly extended our ability to formulate, express, and solve complex problems in many fields.

Simulation has benefitted from the development of both general computing and simulation-specific modeling paradigms. The latter include activity scanning, event modeling, process modeling, and agent-based modeling. Each of these enhances our ability to build valid models. Problems that may seem difficult or impossible when viewed from one paradigm are often tractable or even trivial when viewed from another. For example, simulations of epidemics and tools to combat them (like contact tracing and vaccination programs) are straightforward to implement using agent-based modeling [17, 18], but less so with process modeling.

3.3. Better Environments for Simulating

Simulation-specific software did not even exist until three years after Harling made his "last resort" statement [50]. Since that time, simulation-modeling software has made significant leaps on multiple fronts. Together with the dramatic hardware advances described in Section 3.1, we now have far better platforms with the capability for rapidly and routinely building high-quality simulation models. This would seem to be a much better option than relying on heavily stylized models with strong, questionable assumptions made solely for the sake of pushing through an analytical solution, without regard for validity.

On the user-modeling front, drag-and-drop interfaces promote ease of modeling along with verification and validation. Laborious, error-prone, low-level coding of simulations in a procedural programming language is (mostly) gone. Current simulation software is not just easier, but also more flexible, using object-oriented design or hierarchical structures to facilitate creating models that can better mimic details of complicated systems. Low-level coding still has a place in the bottom layer, in those occasional cases when we wish to assert ultra-fine-grained control to capture particular nuances.

Another front is simulation software itself, which has improved markedly. Better programming and verification practices have increased software quality. Some software effectively and automatically exploits the multicore capabilities of today's processors, greatly increasing run speed.

The underlying random-number algorithms, many of which originated long ago—most notably linear-congruential generators, including the now-infamous RANDU [30]—can be replaced by newer algorithms with astronomical cycle lengths and much-improved statistical qualities that are essential in compute-intensive, high-precision applications [24]. We can enumerate the entire cycle of a legacy 31-bit linear congruential generator (2.1×10^9) in just a couple of minutes on a garden-variety personal computer [19]. By contrast, modern generators, such as the Mersenne Twister [31] with a cycle length of 10^{6001}, would keep that same machine busy for many multiples of the current age of the universe.

We now have improved lists of "standard" probability distributions and processes from which to generate simulation inputs, further enhancing validity. Empirical distributions are also available when fitting standard distributions proves unsatisfactory. Generating dynamic input processes with time-varying rates, like nonstationary arrivals, is standard, and often essential for validity; allowing these in exact analytical models is limited to only a very few simple cases. In simulation, it is no harder to specify a realistic (even if messy) input distribution and process than to specify one that is analytically tractable. Pursuing the analytically-tractable option, historically the more common choice, can do serious damage to model validity and output reliability; more frightening is that, even if we are aware of this risk, we cannot a priori quantify it.

Animation has been in simulation modeling environments for some time and continues to improve. Academics may sometimes belittle it, perhaps as undue effort is sometimes expended fine-tuning the cosmetics, but it can be useful, beyond just communicating and establishing credibility with decision makers who may not know (or care) about simulation per se. By watching an animation and noting anomalous behavior, modeling and coding errors can be revealed. We can also literally see dynamic behavior, like migrating bottlenecks, which could be masked by summary measures like means or maxima.

As Wagner [60] wrote in 1969, results from stochastic simulations are indeed uncertain statistical estimates. It is fair to criticize some simulation projects for failing to recognize this, and thus failing to deal with it appropriately. This has, perhaps deservedly, sometimes given simulation a bad name. But for decades, researchers have been developing robust statistical design-and-analysis protocols specifically for simulation output data. Until recently these protocols, some of which are not simple, have remained inaccessible to simulation practitioners, but that is changing. It is now at least almost automatic to get precision measures such as confidence-interval half widths in standard simulation output, making it easy to assess the precision of point estimates of means. Further, many software packages include effective ways to compare at least the means of multiple competing scenarios in statistically valid ways. Some software easily visualizes (e.g., histograms, box plots) post-run simulation output that goes beyond the traditional focus on means alone [20]; in an application like allocation of voting machines to minimize waiting-time inequity across precincts, it is not the *expected* waiting time that really matters, but rather the *maximum* or the *90th (say) percentile* of the underlying waiting-time distribution [61, 62]. Increasingly common are grafted-on third-party optimum-seeking packages to search for model configurations that are better, according to a critical performance metric, than anyone could find by just trying alternatives in an ad hoc manner. So we now have (and can easily execute) effective statistical procedures that practically eliminate any gnawing uncertainty about our results' precision; as real systems are, after all, mostly themselves stochastic, a simulation can capture that system variation in a realistic way, while still producing results that can be made as precise as desired.

With simulation, we are not afraid of complexity if we need it to build a valid model. An approximate estimate from a valid (simulation) model is preferable to an exact result from an invalid (analytical) model. In the former case we can estimate the imprecision and, if need be, reduce it via more computing. But in the latter case we generally have no way of knowing "how wrong" the results from an overly simplified or perhaps even stylized analytical model might be. Putting it another way, if we greatly simplify model assumptions and structure to arrive at an analytically-tractable model, we run a substantial risk of solving the (possibly-unrealistic) *model* as opposed to addressing the actual *problem* [48]. Mitroff and Featheringham [32] designated "the error ... [of] choosing the wrong problem representation..." to be a "*Type III error*." This catchy term is a nice shorthand that highlights the importance of modeling assumptions, and captures the spirit of Tukey's statement "Far better an approximate answer to the *right* question, which is often vague, than an *exact* answer to the wrong question, which can always be made precise" (pp. 13–14 of [55]).

3.4. Data Farming: Designing Experiments and Analyzing Output

Once your simulation has been built, verified, and validated, "it's time to have the model work for you" [44]. Kleijnen et al. [21] describe three types of goals: to (i) develop a basic understanding of the simulation model and the system it emulates; (ii) find robust policies and decisions, or (iii) compare the merits of various policies or decisions. Well-designed experiments can be efficient and effective ways to help you meet these goals. Even with the exponential increase in processing ability, efficient design of experiments is absolutely required for obtaining broad insights via large-scale simulation studies. A brute-force approach quickly falls victim to the curse of dimensionality. Attempting to explore all possible combinations of 100 input factors, each at just two levels, for a simulation that runs as fast as a single elementary operation, would require all the Tianhe-2's processing power for over 935 millenia. In contrast, even for a simulation with a run time measured in minutes and that has hundreds of factors and higher-order effects, analysts can use efficient designs like those in Vieira et al. [58] to complete the necessary runs over a weekend using a single multicore desktop computer, or in a few hours or minutes on a high-performance computing cluster [46].

The term "data farming" has been used in the defense community over the past decade to capture the notion of purposeful data generation from simulation models. Large-scale designed experiments let us "grow" the simulation output efficiently and effectively. We can explore massive input spaces, uncover interesting features of complex simulation response surfaces, and explicitly identify cause-and-effect relationships. With this new mindset, we can achieve quantum leaps in the breadth, depth, and timeliness of the insights yielded by simulation models.

Sanchez [43] and Sanchez [45] contrast the data-mining and data-farming metaphors as follows. Miners seek valuable nuggets of ore buried in the earth, but have no control over what is out there or how hard it is to extract the nuggets from their surroundings. As they take samples from the earth they gather more information about the underlying geology. Similarly, data miners seek to uncover valuable nuggets of information buried within massive amounts of data. Data-mining techniques use statistical and graphical measures to try to identify *interesting correlations* or clusters in the dataset.

Farmers cultivate the land to maximize their yield. They manipulate the environment to their advantage using irrigation, pest control, crop rotation, fertilizer, and more. Small-scale designed experiments let them determine whether these treatments are effective. Similarly, data farmers manipulate simulation models to their advantage, using large-scale designed experimentation to grow data from their models in a manner that easily lets them extract useful information. This may result in datasets that are big, but still far smaller than what would be needed to gain insights if the results were obtained using ad hoc or randomly generated combinations of factor settings. They also contain better data, in the sense that the results can reveal *root cause-and-effect relationships* between the model input factors and the model responses, in addition to rich graphical and statistical views of these relationships.

Classic experimental designs are ill-suited for exploring simulations of large-scale, complex problems because they either limit the investigation to a handful of factors, or make assumptions about the responses that are unrealistic except over very narrow ranges. Many new designs have been developed in recent years, including single-stage and adaptive sequential methods. Sanchez and Wan [47] created a “consumer-report” chart that provides guidance to those interested in conducting large-scale simulation experiments. An updated version of this chart is kept on the Naval Postgraduate School’s SEED Center web pages [51]. It characterizes designs in terms of their factors, features, and flexibility; gives notes with additional guidance; provides citations for the source papers; and highlights designs that we have found to be good starting points.

Note that data farming need not be limited to stochastic simulation experiments, but can be applied to any model evaluated computationally. For example, many articles on stochastic modeling present formulas that characterize the models’ behaviors, and tabulate numeric results for a small number of test cases or excursions from a baseline. If, instead, the authors had chosen to use a modern design, this could reveal whether there are interactions that substantively alter the patterns in other regions of the design space, or whether the test-case behaviors can be generalized. We provide an example in Section 4.4.

The transformation in analysis capabilities means we can make use of the results from these modern designs. Fitting a single regression took 40 hours of computation near the end of World War II [11]. As late as 1970, central computer labs might have a one-day turnaround time on punch card submissions [40]. Now, fitting regressions involving orders of magnitude more data is a trivial task on a desktop computer. Computationally-intensive analysis techniques abound, including bootstrap methods, stepwise regression, clustering algorithms, partition trees, logistic regression, spline-fitting, and kriging metamodels. A plethora of visualization methods exploit the use of colors, small multiples, three-dimensional plots, contour plots, and other data-mining tools. Interactive plots make it easy to search through high-dimensional data for interesting features.

3.5. A Sea Change is Happening

We understand why, when viewed from the vantage point of the 1950s–1980s, the mere use of simulation tended to be regarded prima facie as an abject, disappointing failure. With the hardware, modeling paradigms, modeling software, and experimental-design and analysis methods of the day, the very idea of numerically acting out all the detailed operations of complex systems was extremely unappealing and unlikely to lead to useful, generalizable insights. One often heard the argument that simulation was just “too expensive,” in terms of both computing time—which cost (even at the margin) real money at the time—and in terms of analyst time to navigate the cumbersome programming languages that were the only option for building simulation models.

But with the multifront, dramatic developments in recent decades chronicled above, we believe that simulation now offers a highly appealing option. Simulation models, which can tolerate almost arbitrary complexity, can thus be far more general than a limited set of analytical models constrained by solvability concerns. With the technologies described here, simulation can, therefore, provide every bit as much insight as can exact analytical models that are based on sometimes-questionable underlying assumptions made expressly for the purpose of enabling an analytical solution in the first place.

The larger scientific community is embracing this new reality. Over the last 60 years, a variety of digital libraries show dramatic growth in the number of articles containing “simulation” as a keyword [38]. To see how much of this is due to a shift in modeling preferences, rather than an overall increase in the number of journal articles, the simulation trends were compared with those of “linear programming” and “optimization”—two other operations-research methods that also benefit from more powerful and affordable processing capabilities. In all cases, the growth of simulation-related articles is more dramatic than the growth of articles related to linear programming and optimization [38].

We are not claiming that simulation can solve all the world’s problems, nor do we argue that simulation is appropriate in all situations. For example, if an analyst has a linear-optimization problem in which all the coefficients and parameters are really known with exact certainty, then simulation is clearly inappropriate. We are merely pointing out that, increasingly, a well-done simulation study is an appealing option that should no longer be regarded as some kind of consolation prize, or even failure.

4. EXAMPLE: MODELING AIRLINE CHECK-IN

Despite the widespread use and acceptance of simulation in the broader scientific community, it is still discounted within some academic circles. As Nance and Sargent [36] noted, simulation has suffered from “scholarly disrespect” in some quarters due to the belief that “simulation was simply a programming exercise” and that “anyone could do it.” In this section, we illustrate the types of insights simulation can provide by comparing and contrasting simulation results to some of the analytical results from a published article that denigrates simulation. By relaxing some of the unrealistic assumptions that were originally made for reasons of analytical tractability, we show the potential consequences of Type III errors.

Our example is the recent paper by Parlar and Sharafali [37], mentioned in Section 1, in which they construct a complex queueing model to “optimize” the number of employees who should staff the check-in counters for an airline. In discussing prior literature, they say "... most of those that have looked into this problem have *resorted to simulation* to study the queue characteristics" and "The work still employed *only simulation* to determine the number of counters to open for each flight..." [emphases ours]. They then proceed to make a number of oversimplifying, unrealistic assumptions (e.g., exponential service times and no group arrivals) to obtain an analytically tractable model.

4.1. Some of the Assumptions for an Analytically Tractable Airline Check-in Model

Parlar and Sharafali built and analyzed a model, to which we will refer as the P&S model, to study how check-in counters should be allocated at an airport. They stressed the importance of the fact that their model is "amenable to analytical treatment." Although they refer to their model as "analytical and more realistic," its applicability is quite narrow and reliant on many assumptions that seem unreasonable to us. Specifically, the situation modeled involved an airport with no "paperless tickets and self check-in kiosks." In addition, the P&S model is an exclusive-use counter system, that is, the counters are used only for passengers of a single airplane, rather than the more typical common-use counter system. Some of the other assumptions that seem risky and contrary to experience are no batch arrivals and exponential service times with an individual counter service rate proportional to the number of people in line. Indeed, we believe exponential service times are routinely assumed solely for mathematical convenience, as nobody would actually believe that the modal service time is zero. Even Parlar and Sharafali [37] question "the assumption of exponential service-time distribution" by asking "[i]sn't a truncated exponential... more suitable than a distribution with infinite support...?" We agree that this is another valid concern, but disagree with their stated reason for retaining the exponential distribution, which is: "such an assumption would also *destroy the Markovian nature* of the process and *make the model intractable*" [emphases ours]. We think their motivation is misguided, and that the more important question is: do such assumptions matter?

4.2. The Airline Check-in Model is Computationally Tractable

We define a model as *computationally tractable* if it can be computationally implemented to yield results that can be estimated to any required level of precision within the allotted time. For many models, using simulation, one can quickly obtain results equivalent to the analytical solution to within any desired degree of precision. This is true of the airport check-in model, which can be readily simulated. Moreover, with modern processing capabilities, the P&S model can be simulated many times in short order. Indeed, we simulated the P&S model using Java. Figure 1 shows a simulation-generated estimate of particular transient transition probabilities, given the same arrival and basic service rates ($\lambda = 1.5$, and $\mu = 5$, respectively) used in [37] for a 10-passenger flight, of going in t time units from a state of four passengers having arrived and two served, to a state of seven having arrived and three served, as a function of elapsed time t, for each of $c = 1$, 2, and 3 counters (i.e., parallel servers). We chose this particular output as it was the one used in [37]. This graph is identical to the analytically derived Fig. 2 in [37], at least to the human eye, and thus we glean identical insights.

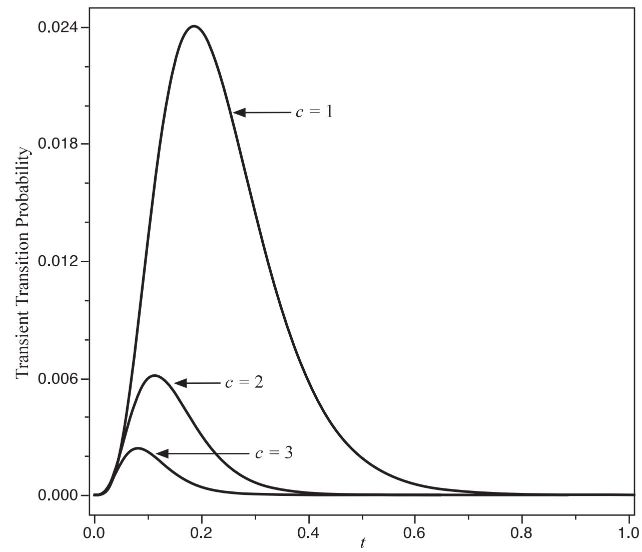

Figure 1. This simulation-estimated transient transition probability graph is visually identical to the analytically derived one in Fig. 2 of Parlar and Sharafali (2008).

Figure 1 is based on 100 million replications of the simulation, which were run in less than 15 minutes on a modern laptop. While the curves are "only" estimates, they are quite precise. The standard error of any point on this graph is much less than the thickness of the lines. The sampling error due to the inherent randomness of a stochastic simulation is easy to measure (and control), but what is the magnitude of the modeling error caused by the unrealistic assumptions required by the analytical approach?

4.3. Using Simulation We Can Solve the Airline Check-in Model Using More Reasonable Assumptions—And It Makes a Difference

The primary value of simulation is not that it can reproduce an analytical solution to an analytically tractable model, but that it facilitates modeling and analyzing many of the attributes of real-world situations that are analytically intractable. To illustrate this, we simulated the airport check-in process

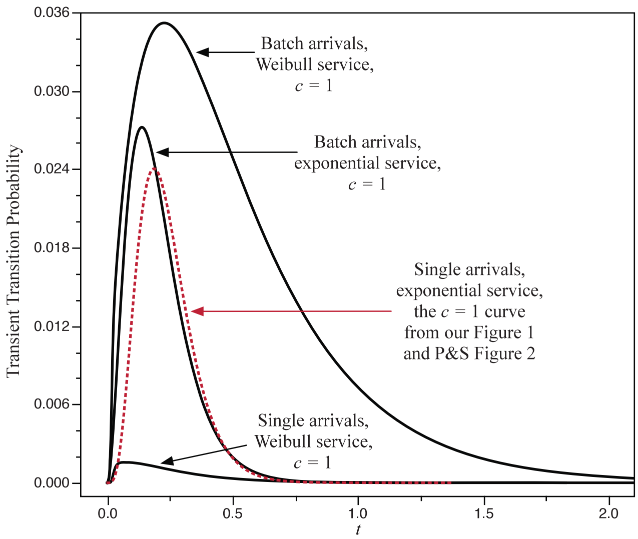

Figure 2. The transient transition probabilities for $c=1$ significantly change when allowing batch arrivals (while preserving the same overall arrival rate) or Weibull service times (while preserving the mean but using shape parameter 5). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

with batch arrivals, and also with a Weibull service-time distribution.

To generate batch arrivals, we used a shifted and truncated Poisson distribution, with a minimum batch size of one rather than zero and a maximum batch size of no more than the remaining number of unprocessed passengers. The rate at which batches arrive was scaled by the expected batch size to yield the same overall expected number of arrivals per time period. Figure 2 shows that, despite having the same mean arrival rate as the P&S model, the transition probabilities dramatically change when we allow batch arrivals.

Another assumption often made for analytical tractability, but which we consider to be unrealistic, is that service times are exponentially distributed. The mode of an exponential distribution is zero, but no reasonable analyst believes that zero is the most likely service time customers will experience. Instead, we chose to model the service time using a Weibull distribution with shape parameter 5 (which yields a "hump"-shaped probability density function with positive mode and support $[0, \infty)$, more typical of service-time data), and with the scale parameter adjusted to yield the same mean service time as the P&S model. Figure 2 shows that the results are highly sensitive to the form of the service-time distribution. Note that in practice, we could use any empirically "appropriate" distribution or even take bootstrap samples from observational data. Figure 2 also illustrates how changing more than one model aspect at a time can yield behavior that is substantially different from that associated with individual changes. The transition-probability curve associated with incorporating both batch arrivals and a Weibull service-time distribution is strikingly different from the original curve and the curves associated with either one of these enhancements. Additional changes would likely ensue if service times were modeled as serially dependent or nonstationary.

4.4. Data Farming the Airline Check-in Model

As we discuss in Section 3.4, it is possible—and can be quite useful—to use designed experiments to explore analytical models. Equation (2) of [37] provides an exact value (according to their analytical model) for the expected time to process all customers, $E[\tau]$, in the special case of $N=3$ potential airline passengers as a function of three parameters: λ and μ, related to the state-dependent arrival and service rates, and the number of counters c:

$$E[\tau] = \frac{22\lambda^4 + 101c\mu\lambda^3 + 143c^2\mu^2\lambda^2 + 89c^3\mu^3\lambda + 22c^4\mu^4}{6c\lambda\mu(2\lambda + c\mu)(\lambda + 2c\mu)(\lambda + c\mu)}.$$

Starting from a base case of $(\lambda, \mu, c) = (1, 5, 1)$, associated with $E[\tau] = 2.091$ hours, they do one-factor-at-a-time *ceteris paribus* excursions to show that $E[\tau]$ would decrease by 1.171 if λ were changed from 1 to 3, increase by 0.493 if μ were changed from 5 to 2, and decrease by 0.203 if c were changed from 1 to 4. A decision maker might anticipate that changing both μ and c simultaneously (while holding λ at 1) would result in $E[\tau] = 2.091 + 0.493 - 0.203 = 2.381$ hours, or a 13.9% *increase* in expected waiting time. However, because of interactions among the factor effects, the analytical result is $E[\tau] = 1.983$ hours, which is a 5.1% *reduction* in expected waiting time! This illustrates the danger of performing one-at-a-time variation from a baseline, instead of using a well-designed experiment, for analytical as well as simulation models. As an interesting side note, $E[\tau]$ provably converges monotonically downward to $11/(6\lambda)$ as c increases, which is $1.83\bar{3}$ in the present case of $\lambda = 1$. (We were able to find $\partial E[\tau]/\partial c$ *symbolically* and show that it is always negative for all admissible values of λ, μ, and c.) This is another instance where we believe that these analytical modeling assumptions are unrealistic—and lead to unrealistic results—because we see no reason why the service time should continue to decrease once the number of counters exceeds the maximum number of customers.

Even more information can be obtained if we use the newer space-filling designs and analysis approaches on a simulation model: we can explore a few orders of magnitude more factors if this is warranted by the system's complexity, yet do so in a computationally efficient way. See [1] for additional examples of these types of insights for an airline check-in simulation.

Another advantage of simulation is that it is straightforward to build a model that includes more than one airplane, more than one flight per day, and other more-realistic conditions—not so for the analytical model.

5. LOOKING AHEAD

Niels Bohr was fond of the Danish proverb, "it's hard to make predictions, especially about the future" [35]. We agree, but provide some general thoughts and hopes for the coming decades. Model-supported decision making will continue to expand, as decision makers seek solutions to problems of increasing complexity. Climate change, economics, transportation, warfare, epidemiology, health care, manufacturing, and social dynamics are just a few of the areas where closed-form analytic models will not suffice—simulation is better at capturing the nuanced behavior of the underlying systems. Those who are studying complex systems are not likely to be interested in answers to simple questions.

Since the emergence of the digital computer almost 70 years ago, there has been a steady trend of increasing capability and decreasing cost. We fully expect that this will continue for some time—whether by better exploitation of current technologies such as parallelization, or emerging technologies such as quantum computing, remains to be seen.

It is difficult to predict what will drive the creation of future simulation-modeling paradigms, but we have no doubt that resourceful people will develop them as new problems arise. Simulation researchers can support emerging application areas by expanding the portfolio of adaptive simulation exploration methods, particularly those that embrace multiple performance measures. Methods and software that facilitate building, running, verifying, updating, and analyzing simulations should and will continue to advance.

Singling out a few from among the many candidate technologies [10], we would like to see more widespread automation of massive, high-dimensional experimentation. Similarly, creating links that readily pull in real-time data from external sources will improve input modeling and enhance the potential of simulation for systems control. Both of these capabilities move a simulation study toward being an ongoing knowledge-acquisition process rather than yielding a fixed end product.

As simulation use expands, improving the practice requires more than just advances in technology—education is critical. Users of simulation are often unaware of, or do not fully use, available simulation technologies. Indeed, many simulation developers and users are experts in their application domain, but not in the discipline of simulation. As with any powerful technology, in the hands of unskilled or uneducated users, there is potential for misuse—avoiding this requires knowledge. Whatever the future holds, the simulation community should be poised to identify, respond to, and ideally blaze a trail that leverages emerging technologies.

6. FINAL THOUGHTS

Our motivations for writing this article go beyond any particular example. Those of us in the simulation community have had countless discussions with numerous colleagues over many years about the philosophy of modeling and analysis. Everybody makes assumptions, whether seeking an analytical or simulation-based solution. What matters is the utility of the resulting model and solution. We believe in parsimonious models [29, 42]. However, we also believe in the adage credited to Einstein that one should "keep your model as simple as possible, *but no simpler*" [emphasis ours]. George Box famously noted that "all models are wrong, but some are useful" [3]. He and Norman Draper further stated that "the practical question is how wrong do they have to be to not be useful" [4]. Because, it commonly requires fewer or weaker assumptions, simulation helps us avoid Type III errors and attain truly useful solutions to complex problems.

Harling had legitimate cause, nearly 60 years ago, to be skeptical of simulation. The complexity of implementation, the cost along many dimensions (hardware costs, compute time, reliability, staffing, energy, etc.), and the dearth of tools for both modeling and analysis, made it a daunting task. These constraints no longer apply. Simulation's ability to address the complexity of real problems has grown by leaps and bounds.

In 1986, John Tukey [16] spoke almost wistfully about the limitations of mathematics relative to computing:

> If anyone, here or later, can tell us how the approach of certainty—traditional mathematics—is going to answer the questions that practical data analysts are going to have to have answered, I will rejoice. Such a route will surely be easier and cheaper, and there will be many more ready to follow it up at once with effective work.
>
> But until I am reliably informed of such a utopian prospect, I shall expect the critical practical answers of the next decade or so to come from the approach of simulation—from a statistician's form of mathematics, in which ever more powerful computing systems will be an essential partner and effective, mathematically sound "swindles" will be of the essence.
>
> To take this view does nothing to discount the paraphrase made by the late great John von Neumann: "The only good Monte Carlo is a dead Monte Carlo!" This aphorism was coined to express the view that out of a well-conducted Monte Carlo should come enough insight to allow us to use newly-developed or newly-chosen approximations to solve other cases of that particular complexity, thus needing to use Monte Carlo

again only when we want to go still further or still deeper. I approve this goal; I only wish I could reach it more often.

The gap between mathematics and computing has only widened in the three decades since Tukey made these observations. The elements for going both "further" and "still deeper" are readily available for modern-day analysts, making simulation an appealing *first* choice.

ACKNOWLEDGMENTS

We wish to thank two anonymous referees and the Associate Editor for their helpful comments, as well as many colleagues for many engaging conversations over many years.

REFERENCES

[1] B.L. Anderson, Changing the paradigm: Simulation, a method of first resort. Master's thesis, Department of Operations Research, Naval Postgraduate School, Monterey, California, September 2011.
[2] J. Banks, J.S. Carson, II, B.L. Nelson, and D.M. Nicol, Discrete-Event System Simulation, 5th ed., Prentice-Hall, Upper Saddle River, NJ, 2010.
[3] G.E.P. Box, "Robustness in the strategy of scientific model building," in: R. Launer and G. Wilkinson (Editors), Robustness in Statistics, Academic Press, New York, 1979, pp. 201–236.
[4] G.E.P. Box and N.R. Draper, Empirical model-building and response surfaces, Wiley, New York, 1987.
[5] P. Bratley, B.L. Fox, and L.E. Schrage, A guide to simulation, Springer-Verlag, New York, 1983.
[6] A.W. Burks and A.R. Burks, The ENIAC: The first general-purpose electronic computer, IEEE Ann Hist Comput 3 (1981), 310–399, 389–399 (commentary).
[7] G.L.L. comte de Buffon, Essai d'arithmétique morale, Supplément á l'Histoire naturelle, générale et particulière, Tome Quatrième, De l'imprimerie royale, Paris, 1777.
[8] J. Dongarra and P. Luszczek, "Anatomy of a globally recursive embedded linpack benchmark," in: IEEE High Performance Extreme Computing (HPEC) Conference, IEEE Xplore Digital Library, 2012.
[9] J. Dongarra, H. Meuer, H.D. Simon, and E. Strohmaier, High performance computing today (lbnl-45666), 2000. http://escholarship.org/uc/item/81m1d8cx, accessed on 2 April 2015.
[10] B.E. Elmegreen, S.M. Sanchez, and A.S. Szalay, "The future of computerized decision making," in: A. Tolk, S. Y. Diallo, O. Ryzhov, L. Yilmaz, S. Buckley, and J. A. Miller (Editors), Proceedings of the 2014 Winter Simulation Conference, IEEE, Piscataway, NJ, 2014, pp. 943–949.
[11] M. Friedman and A.J. Schwartz, Alternative approaches to analyzing economic data, Am Econ Rev 81 (1991), 39–49.
[12] J. Harling, Simulation techniques in operations research—A review, Oper Res 3 (1958), 307–319.
[13] J. Hennessy and D. Patterson, Computer architecture, A quantitative approach, 5th ed., Elsevier, Amsterdam, 2012. http://booksite.elsevier.com/9780123838728/figures.php, accessed on 2 April 2015.
[14] F.S. Hillier and G.J. Lieberman, Introduction to operations research, 10th ed., McGraw-Hill, New York, 2015.
[15] Y. Igarashi, T. Altman, M. Funada, and B. Kamiyama, Computing: A historical and technical perspective, CRC Press, Boca Raton, FL, 2014.
[16] L.V. Jones, The collected works of John W. Tukey, Volume IV: Philosophy and principles of data analysis, Wadsworth & Brooks/Cole, Murray Hill, NJ, 1986.
[17] P. Kasaie, J.R. Andrews, W.D. Kelton, and D.W. Dowdy, Timing of tuberculosis transmission and the population-level impact of household contact tracing: An agent-based simulation model, Am J Respir Crit Care Med 189 (2014), 845–852.
[18] P. Kasaie and W.D. Kelton, A simulation-optimization framework for epidemic resource allocation, IIE Trans Healthc Syst Eng 3 (2013), 78–93.
[19] W.D. Kelton, R.P. Sadowski, and N.B. Zupick, Simulation with Arena, 6th ed., McGraw-Hill, New York, 2015.
[20] W.D. Kelton, J.S. Smith, and D.T. Sturrock, Simio and simulation: Modeling, analysis, applications, 3rd ed., Simio LLC, Pittsburgh, PA, 2015.
[21] J.P.C. Kleijnen, S.M. Sanchez, T.W. Lucas, and T.M. Cioppa, A user's guide to the brave new world of designing simulation experiments, INFORMS J Comput 17 (2005), 263–289.
[22] M. Komorowski, A history of storage costs, http://www.mkomo.com/cost-per-gigabyte-update, accessed on 2 April 2015.
[23] A.M. Law, Simulation modeling & analysis, 5th ed., McGraw-Hill, New York, 2015.
[24] P. L'Ecuyer, R. Simard, E.J. Chen, and W.D. Kelton, An object-oriented random-number package with many long streams and substreams, Oper Res 50 (2002), 1073–1075. http://pubsonline.informs.org/doi/suppl/10.1287/opre.50.6.1073.358, accessed on 2 April 2015.
[25] S. Lewin, Beyond tianhe-2, IEEE Spectrum (2015). http://spectrum.ieee.org/tech-talk/computing/hardware/beyond-tianhe2, accessed on 2 April 2015, Posted 24 June 2014.
[26] P.A.W. Lewis and E.J. Orav, Simulation methodology for statisticians, operations analysts, and engineers, Vol. I, Wadsworth & Brooks/Cole, Belmont, CA, 1989.
[27] J.S. Light, When computers were women, Technol Cult 40 (1999), 455–483.
[28] T.W. Lucas, The stochastic versus the deterministic argument for combat simulations: Tales of when the average won't do! Mil Oper Res 5(3) (2000), 9–28.
[29] T.W. Lucas and J. McGunnigle, When is model complexity too much? Illustrating the benefits of simple models with Hughes' salvo equations, Nav Res Logistics 50 (2003), 197–217.
[30] G. Marsaglia, Random numbers fall mainly in the planes, Proc Natl Acad Sci 61 (1968), 25–28.
[31] M. Matsumoto and T. Nishimura, Mersenne twister: A 623-dimensionally equidistributed uniform pseudo-random number generator, ACM Trans Model Comput Simul 8 (1998), 3–30.
[32] I.I. Mitroff and T.R. Featheringham, On systemic problem solving and the error of the third kind, Behav Sci 19 (1974), 383–393.
[33] G.E. Moore, Cramming more components onto integrated circuits, Electronics 38 (1965), 114–117.
[34] G.E. Moore, Excerpts from a conversation with Gordon Moore: Moore's law, 1975. http://large.stanford.edu/courses/2012/ph250/lee1/docs/Excepts_A_Conversation_with_Gordon_Moore.pdf, accessed on 2 April 2015.

[35] W.J. Moore, Schrödinger, life and thought, Cambridge University Press, Cambridge, UK, 1989.
[36] R.E. Nance and R.G. Sargent, Perspectives on the evolution of simulation, Oper Res 50 (2002), 161–172.
[37] M. Parlar and M. Sharafali, Dynamic allocation of airline check-in counters: A queueing optimization approach, Manage Sci 54 (2008), 1410–1424.
[38] M.J. Powers, S.M. Sanchez, and T.W. Lucas, "The exponential expansion of simulation in research," in: C. Laroque, J. Himmelspach, R. Pasupathy, O. Rose, and A. Uhrmacher (Editors), Proceedings of the 2012 Winter Simulation Conference, IEEE, Piscataway, NJ, 2012, pp. 1552–1563.
[39] A.A.B. Pritsker, C.E. Sigal, and R.D.J. Hammesfahr, SLAM II: Network models for decision support, Prentice-Hall, Englewood Cliifs, NJ, 1989.
[40] R. Ramcharan, Regressions: Why are economists obssessed with them? Finance Dev 43 (2006).
[41] A. Randall, A lost interview with ENIAC co-inventor J. Presper Eckert. Computer World, 2006. http://www.computerworld.com/article/2561813/computer-hardware/q-a–a-lost-interview-with-eniac-co-inventor-j–presper-eckert.html, accessed on 2 April 2015.
[42] P.J. Sánchez, "As simple as possible, but no simpler: A gentle introduction to simulation modeling" in: L.F. Perrone, F.P. Wieland, J. Liu, B.G. Lawson, D.M. Nicol, and R.M. Fujimoto (Editors), Proceedings of the 2006 Winter Simulation Conference, IEEE, Piscataway, NJ, 2006, pp. 2–10.
[43] P.J. Sanchez, "Data farming," in: CENIC Conference, Corporation for Education Network Initiatives in California, Monterey, CA, 2010. http://harvest.nps.edu/papers/DataFarming.mp4, accessed on 2 April 2015.
[44] S.M. Sanchez, "Work smarter, not harder: Guidelines for designing simulation experiments," in: M.E. Kuhl, N.M. Steiger, F.B. Armstrong, and J.A. Joines (Editors), Proceedings of the 2005 Winter Simulation Conference, IEEE, Piscataway, New Jersey, 2005, pp. 69–82.
[45] S.M. Sanchez, "Simulation experiments: Better data, not just big data," in: A. Tolk, S.Y. Diallo, O. Ryzhov, L. Yilmaz, S. Buckley, and J.A. Miller (Editors), Proceedings of the 2014 Winter Simulation Conference, IEEE, Piscataway, New Jersey, 2014, pp. 805–816.
[46] S.M. Sanchez, T.W. Lucas, P.J. Sanchez, C.J. Nannini, and H. Wan, "Designs for large-scale simulation experiments, with applications to defense and homeland security," in: K. Hinkelmann (Editor), Design and Analysis of Experiments: Special Designs and Applications, Vol. 3, 1st ed., Chapter 12, Wiley, New York, 2012, pp. 413–441.
[47] S.M. Sanchez and H. Wan, "Work smarter, not harder: A tutorial on designing and conducting simulation experiments," in: C. Laroque, J. Himmelspach, R. Pasupathy, O. Rose, and A.M. Urmacher (Editors), Proceedings of the 2012 Winter Simulation Conference, IEEE, Piscataway, New Jersey, 2012, pp. 1929–1943.
[48] R.G. Sargent, Personal communication, 2010.
[49] S.L. Savage, The flaw of averages: Why we underestimate risk in the face of uncertainty, Wiley, Hoboken, NJ, 2009.
[50] T.J. Schriber, S. Cox, J.O. Henriksen, P. Lorenz, J. Reitman, and I. Ståhl, "GPSS turns 40: Selected perspectives," in: B.A. Peters, J.S. Smith, D.J. Medeiros, and M.W. Rohrer (Editors), Proceedings of the 2001 Winter Simulation Conference, IEEE, Piscataway, NJ, 2001, pp. 565–576.
[51] SEED Center for Data Farming, Home page. http://harvest.nps.edu, accessed on 2 April 2015.
[52] W.A. Shewhart, Economic control of manufactured product, Van Nostrand Reinhold Co., Princeton, New Jersey, 1931.
[53] Student, The probable error of a mean. Biometrika 6 (1908), 1–25.
[54] TOP500, 2014. http://www.top500.org/, accessed on 2 April 2015.
[55] J.W. Tukey, The future of data analysis, Ann Math Stat 33 (1962), 1–67.
[56] A.M. Turing, On computable numbers, with an application to the Entscheidungsproblem, Proc Lond Math Soc 42 (1936), 230–265.
[57] A.M. Turing, On computable numbers, with an application to the Entscheidungsproblem. A correction, Proc Lond Math Soc 43 (1938), 544–546.
[58] H. Vieira, Jr., S.M. Sanchez, K.H.K. Kienitz, and M.C.N. Belderrain, Efficient, nearly orthogonal-and-balanced, mixed designs: An effective way to conduct trade-off analyses via simulation, J Simul 7 (2013), 264–275.
[59] J. von Neumann, First draft of a report on the EDVAC, Technical Report, University of Pennsylvania, Moore School of Electrical Engineering, June 30, 1945.
[60] H. Wagner, Principles of operations research with applications to managerial decisions, Prentice-Hall, Englewood Cliffs, New Jersey, 1969.
[61] M. Yang, T.T. Allen, M.J. Fry, and W.D. Kelton, The call for equity: Simulation-optimization models to minimize the range of waiting times, IIE Trans 45 (2013), 781–795.
[62] M. Yang, M.J. Fry, W.D. Kelton, and T.T. Allen, Improving voting systems through service-operations management, Prod Oper Manage 23 (2014), 1083–1097.