Operations Research Letters 10 (1991) 437–443
North-Holland

Analysis of initial transient deletion for replicated steady-state simulations

Peter W. Glynn *
Department of Operations Research, Stanford University, Stanford, CA 94305-4022, USA

Philip Heidelberger
IBM Thomas J. Watson Research Center, P.O. Box 704, Yorktown Heights, NY 10598, USA

Received December 1989
Revised January 1991

Consider the method of independent replications with initial transient deletion for generating confidence intervals for 'steady-state' quantities. To produce intervals with good convergence characteristics, the relative growth rates of the number of replications, the length of each replication, and the deletion period must be controlled. Critical rates for these parameters are determined. The applicability of these results to simultaneously running multiple replications on a highly parallel computer is discussed.

simulation; steady-state; initial transient

1. Introduction

One of the major practical problems in steady-state simulation output analysis concerns the question of how to deal with 'initial transient' effects. The term 'initial transient' refers to that initial segment of the simulation that is biased because of initialization of the simulation with an initial condition that is atypical of the steady-state. For example, simulations of open queueing networks are often initialized by starting the simulation with all the stations empty. By permitting the simulation to run for a sufficiently long period of time, the queue-lengths in the network then build up to their respective steady-state values. Clearly, the observations collected as the system moves towards its steady-state are biased for the purposes of steady-state estimation. This biasing property of the 'initial transient' segment of the simulation can lead to significant degradation in the quality of steady-state estimators and confidence intervals. As a consequence, considerable research effort has been devoted to development of procedures for elimination of the initial transient segment; see, for example, Gafarian et al. (1978), Heidelberger and Welch (1983), Kelton (1980), Kelton and Law (1980), Schruben (1981, 1982), Schruben, Singh and Tierney (1983), and Wilson and Pritsker (1978).

In this paper, we provide some theoretical insight into the problem of initial transient deletion in the multiple replicate steady-state simulation setting. In particular, we show that without any initial transient deletion, significant convergence problems can arise if the length of each replication is not considerably larger than the total number of replications. On the other hand, if we delete an (asymptotically) negligible initial segment of each replication, these convergence difficulties disappear. One can then apply multiple replication procedures in which the total number of replicates is significantly larger than the length of each replication. Since variance estimation is advantageously affected through the presence of a large

* Research supported by the IBM Corporation under SUR-SST Contract 12480042 and by the U.S. Army Research Office under Contract DAAL-03-88-K-0063, and by a grant of the Natural Sciences and Engineering Research Council of Canada.

number of independent replications, this has important theoretical ramifications for steady-state simulation output analysis. (Although as Schmeiser (1982) has shown in the context of batch means, the practical advantage of increasing the number of batches/replications beyond a certain point is marginal). Our approach also permits us to quantify (in a rough sense) the amount of initial transient deletion that is appropriate for a given simulation. We view our analysis as offering theoretical support for the principle of initial transient deletion and providing some theoretical insight into how the deletion should be done.

The primary motivation for these results is that they have important ramifications for steady-state estimation in a parallel computing environment. Probably the simplest and most powerful approach to implementing a simulation on a parallel computer is to run independent replicates of the simulation, in parallel, on each of the available processors. The resulting estimator then corresponds to a multiple replicate estimator. Furthermore, since the cost per processor is widely expected to decline significantly in the future, one can expect to see parallel machines with (very) large numbers of processors available. Thus, it is appropriate to develop estimation procedures in which the number of replications may be large relative to the length of each replication. Such procedures would allow the simulation analyst to simultaneously run many simulation replications on a highly parallel computer and receive a valid, low variance estimator in a relatively short period of time. This results in a large turnaround time reduction over running the simulation on a single processor for a comparable amount of simulation time (divided into either one long replication, or multiple shorter replications).

The results in this paper are appropriate to the analysis of such estimators, provided that the length of each replication and the deletion period are determined in units of simulated time. In particular, our results suggest that initial transient deletion becomes increasingly advantageous as the number of processors increases. In Glynn and Heidelberger (1989 and 1990), we discuss the (more complicated) situation in which the replication length and deletion interval are specified in units of computer time.

Thus, on a highly parallel machine there is strong motivation to use a larger number of replications. However, on a single processor (or a parallel processor with only a few processors), the motivation for many replications is less compelling since using either a single run procedure with, say, batch means and deletion to form confidence intervals, or using a modest number of replications and deletion is less sensitive to the effects of initialization bias. In addition, on a single processor, there is no turnaround time advantage to running multiple replications as there is on parallel processors.

Heidelberger (1988) and Glynn and Heidelberger (1991) study related questions in parallel simulation of terminating simulations. However, while those results do apply to steady-state analysis of regenerative simulations, our discussion here is valid much more generally. On the other hand, most of the results of this paper are more qualitative in nature and do not (immediately) suggest operational procedures for dealing with initial transient deletion in a parallel simulation setting. Thus, the results of this paper complement those of Heidelberger (1988) and Glynn and Heidelberger (1991), rather than subsume them.

This paper is organized as follows. Section 2 describes and discusses the main results of this paper, while Section 3 is devoted to proofs of the results.

2. Description of the main results

The basic setting for our analysis is the existence of a real-valued stochastic process $X = (X(t):\ t \geq 0)$ that represents the simulation output generated by a single replication of the stochastic system under study. We assume that there exist finite constants α and σ such that

$$t^{1/2}\left(\frac{1}{t}\int_0^t X(s)\,\mathrm{d}s - \alpha\right) \Rightarrow \sigma N(0,1) \tag{2.1}$$

as $t \to \infty$. We note that standard weak convergence arguments show that (2.1) implies that

$$\frac{1}{t}\int_0^t X(s)\,\mathrm{d}s \Rightarrow \alpha$$

as $t \to \infty$. As a consequence, (2.1) implies that X has a long-run steady-state average value α. The parameter α is called the *steady-state mean* of X; the constant σ^2 is known as the *time-average variance constant* of X. The goal of steady-state simulation is to estimate and produce confidence intervals for α.

We will further require that X be appropriately uniformly integrable (uniform integrability permits one to pass limits through expectations). Specifically, we shall assume that

$$tE\left(\frac{1}{t}\int_0^t X(s)\,\mathrm{d}s-\alpha\right)^2 \to E\sigma^2 N(0,1)^2=\sigma^2 \tag{2.2}$$

as $t\to\infty$. Assumptions (2.1) and (2.2) hold under a variety of different assumptions on the output process X. Regenerative processes (Smith (1955)), associated sequences (Newman and Wright (1981)), and mixing processes (Ethier and Kurtz (1986)) all satisfy the two assumptions under suitable regularity hypotheses. As a consequence, the central limit theorem assumptions (2.1) and (2.2) can be viewed as being typically satisfied by most discrete-event simulations.

We also need a further additional assumption that controls the extent to which the initial transient biases the observations of a simulation. To be precise, let $b(t)=EX(t)-\alpha$. We assume that

$$\int_0^\infty |b(t)|\,\mathrm{d}t<\infty. \tag{2.3}$$

We note that (2.3) holds whenever $EX(t)\to\alpha$ exponentially fast (i.e., the exist constants A, $\lambda>0$ such that $|EX(t)-\alpha|\le A\exp(-\lambda t)$ for $t\ge 0$). This exponential rate of convergence is typical of most reasonable stochastic processes. For example, autoregressive sequences and geometrically ergodic Markov chains (see Nummelin and Tuominen (1982)) have this type of behavior.

In particular, irreducible aperiodic finite state space discrete time Markov chains exhibit exponential convergence to steady-state (see Chapter 5 of Doob (1953)) as do irreducible finite state continuous time Markov chains (see, e.g., Chapter 4 of Karlin and Taylor (1975)).

A consequence of (2.3) is that the bias of the estimator $\overline{X}(t)\triangleq t^{-1}\int_0^t X(s)\,\mathrm{d}s$ goes to zero at rate $1/t$. To see this, note that we can write

$$\begin{aligned}E\overline{X}(t)-\alpha&=\frac{1}{t}\int_0^t b(s)\,\mathrm{d}s\\&=\frac{1}{t}\int_0^\infty b(s)\,\mathrm{d}s-\frac{1}{t}\int_t^\infty b(s)\,\mathrm{d}s\\&=b/t+\mathrm{o}(1/t)\end{aligned}$$

as $t\to\infty$ where $b\triangleq\int_0^\infty b(s)\,\mathrm{d}s$.

We now turn to defining our replication estimator for the steady-state mean α. Let $X_1, X_2,\ldots,$ be a sequence of i.i.d. replicates of the process X. Suppose T is a computational budget constraint which defines the total amount of simulation to be performed over all replications. Let $m(T)$ be the number of replications associated with computational budget T. We assume that $m(T)$ is a deterministic quantity. Then, since the total simulation time equals T, it follows that each replicate is simulated for $l(T)=T/m(T)$ time units. Finally, we let $\beta(T)$ be the (deterministic) length of the initial segment that is to be deleted from each replication. As discussed earlier, this deletion is performed so as to reduce the bias of the replicated estimator. Given the above description, our estimator $\alpha(T)$, associated with computational budget T, then takes the form

$$\alpha(T)=\frac{1}{m(T)}\sum_{i=1}^{m(T)}\frac{1}{l(T)-\beta(T)}\int_{\beta(T)}^{l(T)}X_i(s)\,\mathrm{d}s.$$

Our first theorem considers the case where no initial bias deletion is performed, so that $\beta(T)\equiv 0$. The theorem asserts that in order for $\alpha(T)$ to have $T^{-1/2}$ rate of convergence, it is necessary to choose $m(T)$ so that $m(T)/T^{1/2}\to 0$ as $T\to\infty$ (or equivalently, $m(T)/l(T)\to 0$ as $T\to\infty$). Hence, if no initial bias deletion is performed, it is necessary to use a relatively small number of replicates (small relative to the length of each replication).

Theorem 1. *Assume* (2.1)–(2.3) *are in force and that* $\beta(T)\equiv 0$. *Then*:

(i) *if* $m(T)/T^{1/2}\to\infty$ *and* $b\ne 0$, *then* $T^{1/2}|\alpha(T)-\alpha|\Rightarrow\infty$ *as* $T\to\infty$,

(ii) *if* $m(T)/T^{1/2}\to m$ $(0<m<\infty)$, *then* $T^{1/2}(\alpha(T)-\alpha)\Rightarrow\sigma N(0,1)+bm$ *as* $T\to\infty$,

(iii) *if* $m(T)/T^{1/2}\to 0$, *then* $T^{1/2}(\alpha(T)-\alpha)\Rightarrow\sigma N(0,1)$ *as* $T\to\infty$.

This result appears as Theorem 3 of Glynn (1987).

Our next result shows that when initial transient deletion is applied to a stochastic process X in which the bias of $X(t)$ decreases polynomially fast (i.e., $|b(t)|=\mathrm{O}(t^{-r})$, for some $r>0$, as $t\to\infty$), the constraint on how many replications can be used relaxes considerably. Specifically, Theorem 2 shows that if $|b(t)|=\mathrm{O}(t^{-r})$ for $r>1$ (note that (2.3) is typically violated if $r\le 1$), then initial transient deletion can be implemented so that the number of replications can grow as fast as T^p ($p < 1 - 1/2r$). Note that when the constant r is large, the constraint on p essentially disappears.

Theorem 2. *Assume* (2.1)–(2.2) *are in force and that* $|b(t)| = O(t^{-r})$ *for some* $r > 1$. *Suppose that we choose* $\beta(T) = T^{\beta}$ *and* $m(T) = \lfloor T^p \rfloor$ *where* $1 > \beta \geq 1/2r$ *and* $p < 1 - \beta$. *Then*

$$T^{1/2}(\alpha(T) - \alpha) \Rightarrow \sigma N(0, 1)$$

as $T \to \infty$.

To further interpret Theorem 2, consider what happens if the deletion period $\beta(T)$ grows at its minimum rate $\beta(T) = T^{1/2r}$ and the number of replications grows at close to its maximum rate, i.e., $m(T) \sim T^{1-1/2r-\epsilon}$ for a (small) $\epsilon > 0$. In this case, the total fraction of the simulation that is deleted is $m(T)\beta(T)/T \sim T^{-\epsilon}$, which goes to zero as $T \to \infty$. Thus, we can delete an asymptotically negligible fraction of the simulation; this reflects itself in the asymptotic variance σ^2/T. If no deletion is performed, then by Theorem 1, we must have $m(T)/l(T)$ bounded in order to obtain a convergence rate of $T^{-1/2}$. However, in this case of asymptotically negligible deletion, we have $m(t)/l(T) \sim T^{1-2\epsilon-1/r} \to \infty$ as $T \to \infty$. As the convergence rate of the bias term $b(t)$ increases (σ increases), one can delete less and replicate more.

We note that application of Theorem 2 requires only that the simulation analyst have a lower bound on r, say r'. Then, choosing $\beta(T) = T^{1/2r'}$ and $m(T) = T^p$, where $p = 1 - (1/2r') - \epsilon (\epsilon > 0)$, automatically satisfies the conditions of Theorem 2. The use of coupling techniques in the Markov process setting (see, for example, Pitman (1974)) has proven successful in obtaining estimates of the form $|b(t)| = O(t^{-r})$. In particular, if the r'th moment of the expected return time to a fixed state is finite, such estimates ensue. Hence, a lower bound r' on r is available whenever it is known that the return time to the initial state for the simulation has a finite r''th moment.

As noted earlier in this section, many discrete-event simulations in fact have the property that the bias function $b(\cdot)$ goes to zero exponentially fast. For such simulations, the results of Theorem 2 can be improved. Specifically, Theorem 3 shows that the number of replications may grow at the rate T^p for any $p < 1$. Exponential decay of the bias function typically holds in simulations for which it is known that the return time to the initial state has a moment generating function which converges in a neighborhood of zero (see Nummelin and Tuominen (1982)).

We next consider asymptotically negligible deletion rules ($m(T)\beta(T)/T \to 0$ as $T \to \infty$), in the case of exponentially decreasing bias.

Theorem 3. *Assume* (2.1)–(2.2) *are in force and that* $b(t) = O(e^{-\lambda t})$ *for some* $\lambda > 0$. *Suppose* $m(T)\beta(T)/T \to 0$ *as* $T \to \infty$. *If either*

(i) $m(T)/T^{1/2} \to \infty$ *and* $\beta(T)/\log T \to \infty$, *or*

(ii) $m(T)/T^{1/2} \to m$ $(0 < m < \infty)$ *and* $\beta(T) \to \infty$, *or*

(iii) $m(T)/T^{1/2} \to 0$,

then

$$T^{1/2}(\alpha(T) - \alpha) \Rightarrow \sigma N(0, 1)$$

as $T \to \infty$.

Comparing Theorem 3 to Theorem 1, we see that, with no deletion, only sampling plan (iii) with $m(T)/T^{1/2} \to 0$ produces a valid (or usable) central limit theorem for α. However, with deletion, sampling plans (i) and (ii) also produce central limit theorems for α with $T^{-1/2}$ rate of convergence, provided $\beta(T) \to \infty$ appropriately. From Theorem 3, the deletion period $\beta(T)$ can grow quite slowly. We note that one particular permisible choice is $\beta(T) = (\log T)^{1+\delta}$ and $m(T) = T^p$ for $\delta > 0$ and $p < 1$.

The deletion plans given in Theorems 2 and 3 are such that the total deletion length of the simulation $m(T)\beta(T)$ is large, but that the fraction of the total simulation effort that is deleted is asymptotically negligible. This results in an asymptotic variance of $\sigma^2(T)$, the same as in the case where no data is deleted. However, in practice, a reasonable (and probably not uncommon) trancation rule is to delete a fixed fraction β $(0 < \beta < 1)$ of each replication, i.e., $\beta(T) = \beta l(T)$. In this case, the amount of simulation data that is retained is $T(1 - \beta)$; the asymptotic variance then becomes $\sigma^2/(1 - \beta)T$. For $\beta \approx 0$, the increase in variance over the asymptotically negligible scheme is small. Even for a rather sizable $\beta = 0.1$, the variance increases by only about 11%. Such a fixed fractional truncation rule provides robust protection against the initial transient since, for example, in the case of exponentially decreasing bias, $E[\alpha(T)-\alpha]$ decreases to zero exponentially fast in $l(T)$, the length of the replications.

Theorem 4 gives a precise analysis of fixed fractional truncation rules. It turns out that this result requires that we strengthen (2.1) to a functional central limit theorem (FCLT) for X: There exist finite constants α and σ such that

$$Y_\epsilon \Rightarrow \sigma W \qquad (2.4)$$

as $\epsilon \downarrow 0$ in the Skorohod space $D[0, \infty)$ where

$$Y_\epsilon(t)=\epsilon\int_0^{t/\epsilon^2}(X(s)-\alpha)\,\mathrm{d}s,$$

and W is a standard Wiener process (also known as a Brownian motion). From a practical standpoint (2.4) can be viewed as being essentially identical to (2.1), in the sense that (2.4) is typically valid whenever (2.1) is, and vice versa. In particular, FCLT's are known to hold for regenerative processes (Glynn and Whitt (1987)), associated sequences (Newmann and Wright (1981)), martingale processes, and mixing sequences (Ethier and Kurtz (1986)).

Theorem 4. *Assume* (2.2) *and* (2.4) *are in force. Let* $m(T)=\lfloor T^p \rfloor$ *and assume* $\beta(T)=\beta l(T)$ *for some* $\beta(0<\beta<1)$. *If either*

(i) $b(t)=\mathrm{O}(t^{-r})$ *for some* $r>1$ *and* $p \le 1-1/2r$, *or*

(ii) $b(t)=\mathrm{O}(\mathrm{e}^{-\lambda t})$ *for some* $\lambda>0$ *and* $p<1$,

then

$$T^{1/2}(\alpha(T)-\alpha) \Rightarrow (1-\beta)^{-1/2}\sigma N(0, 1)$$

as $T\to\infty$.

We conclude this paper with a discussion of the implications of the above results for periodic processes. In particular, suppose $X(t)=X_{\lfloor t \rfloor}$, where $(X_n\colon n \ge 0)$ is a periodic irreducible finite state Markov chain. Then, it is no longer true that $b(t)\to 0$ as $t\to\infty$, where $b(t)=EX(t)-\alpha$. The key to the theory developed in this paper is that a significant reduction in the bias of the replicated estimator $\alpha(T)$ can be obtained by deleting an initial segment of length $\beta(T)$. For a periodic Markov chain, such a bias reduction can still be accomplished, provided that $\beta(T)$ is constrained so that $l(T)-\beta(T)$ is an integer multiple of the period of the Markov chain. Hence, Theorems 2–4 continue to hold in the periodic case, provided that $\beta(T)$, in addition to satisfying the stated growth conditions, is selected so that $l(T)-\beta(T)$ is a multiple of the period of the chain. This observation may have important operational implications for how initial deletion should be implemented for simulations possessing periodicities.

3. Proofs

Proof of Theorem 2. Let $q=1-p$ and write $\alpha(T)-\alpha$ as

$$\alpha(T)-\alpha=\frac{T^{-q/2}}{m(T)}\sum_{i=1}^{m(T)} W_i(T)+B(T)$$

where

$$W_i(T)=\frac{T^{q/2}}{l(T)-\beta(T)} \times \int_{\beta(T)}^{l(T)}[X_i(s)-EX_i(s)]\,\mathrm{d}s$$

and

$$B(T)=\frac{1}{l(T)-\beta(T)}\int_{\beta(T)}^{l(T)} b(s)\,\mathrm{d}s.$$

We note that

$$\begin{aligned} T^{1/2}B(T) &\sim T^{1/2-q}\int_{T^\beta}^{T^q} b(s)\,\mathrm{d}s \\ &= T^{1/2-q}\mathrm{O}(T^{\beta(1-r)}) \\ &= \mathrm{O}(T^{\beta-q}) \to 0 \end{aligned}$$

as $T\to\infty$. To conclude the proof, we therefore need to show that

$$\frac{T^{(1-q)/2}}{m(T)}\sum_{i=1}^{m(T)} W_i(T) \Rightarrow \sigma N(0, 1)$$

as $T\to\infty$. This follows from the Lindeberg-Feller theorem, provided that we show that Lindeberg's condition is satisfied (see p. 205 of Chung (1974)). However, Lindeberg's condition reduces here to showing that

$$E\{W_i(T)^2;\ W_i(T)^2 > T^{1-q}\} \to 0 \qquad (3.1)$$

as $T\to\infty$. But (2.1)–(2.2) together imply that

$$W_i(T) \Rightarrow \sigma N(0, 1),$$

$$EW_i^2(T) \to \sigma^2$$

as $T\to\infty$. As a consequence, $\{W_i(T)\}$ is appropriately uniformly integrable (see Theorem

4.5.4 of Chung (1974)); the uniform integrability then yields (3.1).

Proof of Theorem 3. We need only show that $T^{1/2}B(T) \to 0$ as $T \to \infty$, since the uniform integrability argument is the same as in Theorem 2. Observe that $m(T)\beta(T)/T \to 0$ is equivalent to $\beta(T)/l(T) \to 0$ as $T \to \infty$. Hence, $l(T) - \beta(T) = l(T)(1 + o(1))$. As a consequence,

$$T^{1/2}B(T) = \frac{T^{1/2}}{l(T)(1+o(1))} \int_{\beta(T)}^{l(T)} b(s)\, ds = \frac{T^{1/2}}{l(T)(1+o(1))} O(e^{-\lambda\beta(T)});$$

this immediately yields parts (ii) and (iii). For part (i), note that $l(T) \geq \beta(T) \to \infty$, so $T^{1/2}/l(T) = O(T^{1/2})$, and thus

$$T^{1/2}B(T) = O(T^{1/2}\, e^{-\lambda\beta(T)}),$$

which converges to zero, since $\beta(T)\log T \to \infty$.

Proof of Theorem 4. Note that

$$T^{1/2}(\alpha(T) - \alpha) = \frac{1}{\sqrt{m(T)}} \sum_{i=1}^{m(T)} \chi_i(T)/(1-\beta),$$

where

$$\chi_i(T) = l(T)^{-1/2} \int_{l(T)\beta}^{l(T)} (X_i(s) - \alpha)\, ds.$$

But

$$\chi_i(T) \overset{\mathcal{D}}{=} Y_{l(T)^{-1/2}}(1) - Y_{l(T)^{-1/2}}(\beta) \Rightarrow \sigma[W(1) - W(\beta)]$$

as $T \to \infty$, where $\overset{\mathcal{D}}{=}$ denotes equality in distribution. Furthermore, (2.2) implies that $\{Y^2_{l(T)^{-1/2}}(t): T \geq 0\}$ is uniformly integrable for any fixed t. Hence $\{\chi_i^2(T): T \geq 0\}$ is uniformly integrable since $\chi_i^2(T) \leq 2Y_{l(T)^{-1/2}}(1)^2 + 2Y_{l(T)^{-1/2}}(\beta)^2$. Also $T^{1/2}(E\alpha(T) - \alpha)^{l(T)^{-1/2}} \to 0$ (the proof is identical to that used in Theorems 2 and 3). As a consequence, $\{\hat{\chi}_i(T): T \geq 0\}$ is uniformly integrable, where $\hat{\chi}_i(T) = \chi_i(T) - E\chi_i(T)$. The Lindeberg-Feller theorem may then be applied (as in Theorem 2) to conclude that

$$\frac{1}{\sqrt{m(T)}} \sum_{i=1}^{m(T)} \hat{\chi}_i(T)/(1-\beta) \Rightarrow (1-\beta)^{-1/2}\sigma N(0, 1)$$

as $T \to \infty$. Using once again the fact that $T^{1/2}(E\alpha(T) - \alpha) \to 0$, we obtain the desired conclusion.

References

Chung, K.L. (1974), *A Course in Probability Theory*, Academic Press, New York.

Doob, J.L. (1953), *Stochastic Processes*, John Wiley, New York.

Ethier, S.N., and Kurtz, T.G. (1986), *Markov Processes: Characterization and Convergence*, John Wiley, New York.

Gafarian, A.V., Ancker, C.J., Jr., and Morisaku, T. (1978), "Evaluation of commonly used rules for detecting 'steady-state' in computer simulation", *Naval Research Logistics Quarterly* **25**, 511–529.

Glynn, P.W. (1987), "Limit theorems for the method of replication", *Stochastic Models* **4**, 344–350.

Glynn, P.W., and Heidelberger, P. (1989), "Analysis of initial transient deletion for parallel, steady-state simulations", IBM Research Report RC 15260, Yorktown Heights, New York, to appear in *SIAM J. Sci. Statist. Comput.*

Glynn, P.W., and Heidelberger, P. (1990), "Experiments with initial transient deletion for parallel, replicated steady-state simulations", IBM Research Report RC 15770. Yorktown Heights, New York, to appear in *Management Sci.*

Glynn, P.W., and Heidelberger, P. (1991), "Analysis of parallel, replicated simulations under a completion time constraint", *ACM Transactions on Modeling and Computer Simulation* **1**/1, 3–23.

Glynn, P.W., and Whitt, W. (1987), "Sufficient conditions for functional-limit-theorem versions of $L = \lambda W$", *Queueing Systems, Theory and Applications* **1**, 279–287.

Heidelberger, P., and Welch, P.D. (1983), "Simulation run length control in the presence of an initial transient", *Operations Research* **31**, 1109–1114.

Heidelberger, P. (1988), "Discrete event simulations and parallel processing: statistical properties", *SIAM Journal on Scientific and Statistical Computing* **9**, 1114–1132.

Karlin, S. and H.M. Taylor (1975), *A First Course in Stochastic Processes*, Academic Press, New York.

Kelton, W.D. (1980), "The startup problem in discrete event simulation", Technical Report 80-1, Department of Industrial Engineering, University of Wisconsin, Madison, WI.

Kelton, W.D., and Law, A.M. (1980), "A new approach for dealing with the startup problem in discrete event simulation", *Naval Research Logistics Quarterly* **30**, 641–658.

Newman, C.M., and Wright, A.L., (1981), "An invariance principle for certain dependent sequences", *Annals of Probability* **9**, 671–675.

Nummelin, E., and Tuominen, P. (1982), "Geometric ergodicity of Harris recurrent Markov chains with applications to renewal theory", *Stochastic Processes and Their Applications* **12**, 187–202.

Pitman, J. (1974), "Uniform rates of convergence for Markov chain transition probabilities", *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* **29**, 193–227.

Schmeiser, B.W. (1982), "Batch size effects in the analysis of simulation output", *Operations Research* **30**, 556–568.

Schruben, L.W. (1981) "Control of initialization bias in multivariate simulation response", *Communications of the ACM* **24**, 246–252.

Schruben, L.W. (1982), "Detecting initialization bias in simulation output", *Operations Research* **30**, 569–590.

Schruben, L.W., Singh, H., and Tierney, L. (1983), "Optimal tests for initialization bias in simulation output", *Operations Research* **31**, 1167–1178.

Smith, W.L. (1955), Regenerative stochastic processes. *Proceedings of the Royal Society of London Series A* **232**, 6–31.

Wilson, J.R., and Pritsker, A.A.B. (1978), "A survey of research on the simulation start-up problem", *Simulation* **31**, 55–58.