SCHOOL OF OPERATIONS RESEARCH
AND INDUSTRIAL ENGINEERING
COLLEGE OF ENGINEERING
CORNELL UNIVERSITY
ITHACA, NEW YORK

TECHNICAL REPORT NO. 594

October, 1983

# INITIALIZATION EFFECTS IN COMPUTER SIMULATION EXPERIMENTS

by

Lee Schruben

and

David Goldsman

This research was supported by the Office of Naval Research under contract N00014-81-k-0037.

Abstract

Issues relating to initializing a simulation program are discussed, and suggested procedures for handling this source of error are reviewed. We examine academic and practical aspects of simulation initialization. A bibliography is provided.

## 1. Introduction

When a simulation program is run on a computer, initial values for all variables must be specified. If the simulation is such that appropriate initial conditions are known or are fixed factors in an experiment, then this causes no problem. However, the experimenter typically does not know what values are appropriate for all of the variables in the program. As a matter of convenience, these initial conditions are often selected in a rather arbitrary manner. In this case, the initial conditions can potentially have a significant influence on the outcome of the experiment. If the setting of the initial conditions for a run has a major but unrecognized effect, then the results from the run can be misleading. Thus, initialization can be a serious source of error in simulation experiments. In section 2 of this paper, some simple examples of simulation initialization effects are presented. This is followed by a discussion of the mathematical aspects of initializing a simulation program in section 3. Suggested procedures for handling this source of error are reviewed in section 4 along with some of the academic and practical issues. A bibliography is included.

## 2. Some examples

By considering some simple examples, we illustrate some of the problems that can arise when initializing a simulation. If a simulation of a new factory is initiated with no work in progress, then the production of finished goods will be relatively low early in the run. Products simply have not had time to flow through the system. In addition, there will be little congestion in the system until it fills up with work in progress. A naive statistical analysis of the data from such a simulation might conclude that the production is positively correlated with congestion on the factory floor. This correlation is an artifact caused by the way in which the experiment was run; the correlation is not a natural characteristic of the system being studied. Nevertheless, the study might suggest that one way to increase production is to make the factory more congested. Low observed machine utilizations (perhaps again due to the simulated factory being initially empty) could lead the analyst to recommend that fewer machines than originally planned be installed. Unfortunately, the use of fewer machines could then result in severe system bottlenecks as production increases.

Another example illustrates the problem of visually detecting initialization effects in estimators of system performance. Figure 1 is the output (customer waiting times) of a simulated queueing system which was initially started with no customers in the system. The initial portion of the run does not appear to differ radically from the rest of the run. However, the average waiting time for the first 50 customers is 10 minutes. The average waiting time for the first 500 customers is 27 minutes - almost three times the value for the

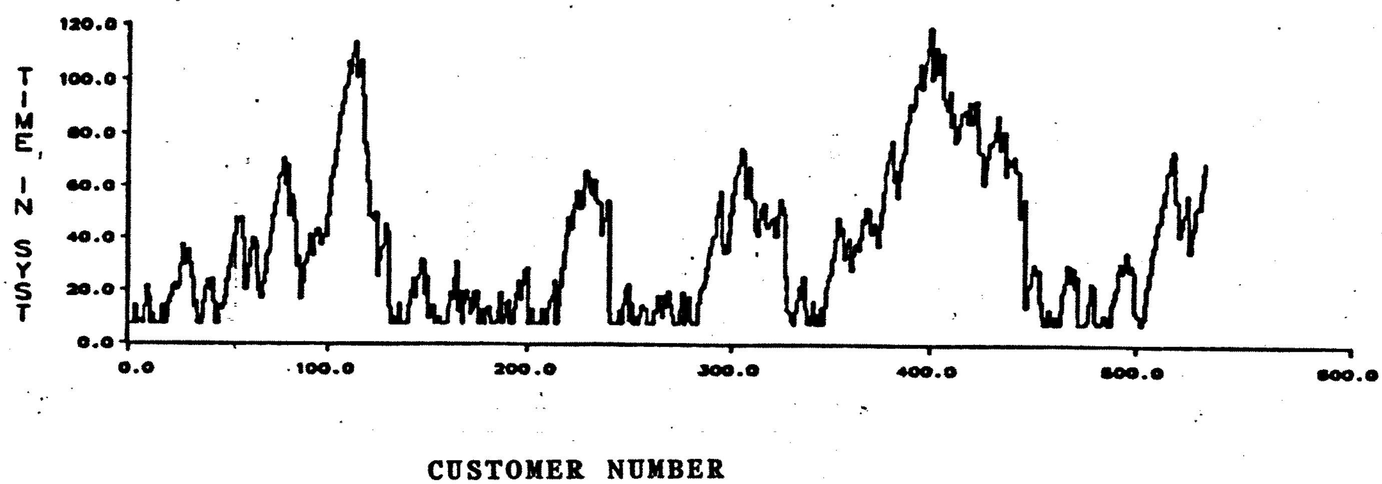


Figure 1 : Time in System for customers in a simple M/D/1 queue.

first 50 customers! Obviously, the relatively low observed waiting times for early customers are due in part to the empty and idle initialization of the system.

We continue with the above queueing example. It may be that the system closes each evening and re-opens the next morning with no customers waiting in line (in which case starting the system empty is appropriate). However, customers might return the next morning if they are not served by closing time. There would then be a backlogging of demand from day to day. In this case, each day is not independent; starting the simulation with the same initial condition of no demand is not appropriate. Such a simulation should have each day started with an initial demand that is dependent on the demand level from the end of the previous day.

An example from an actual simulation study [Schwarz (1976)] shows how secondary correlation of output measures with the initial conditions can lead to errors. The following discussion is not intended to be a criticism of this study. The purpose of the study was to investigate short-term behavior of different inventory policies. (Of course, if the purpose had been to estimate the long-term average performance of the different policies, then initialization effects could have produced serious errors.)

In this simulation, three alternative inventory stocking policies were considered. Each alternative was simulated several times. In each run, the stock levels were the same; there were initially no backorders or backlogged demand. The measure of performance of the different policies was total unsatisfied demand (measured in 'backorder days') over a period of time. The specific measure of system performance is not actually important to our discussion. When the simulation was run for three years (simulated time) there were no marked differences in the performances of the policies. This was in spite of the fact that the systems were allowed to 'warm-up' for two years before data were collected. After five years of operation, there were significant differences in policy performances. The early similarity in performance of the policies was largely due to the fact that all of the simulation runs started with the same initial conditions.

Since we are trying to detect differences in policies it is good experimental practice to control other sources of difference (here, the initial conditions) by holding these sources at the same value. Indeed, in this problem, having the same initial conditions was *apparently* the right thing to do from an experimental design viewpoint. This reduces the variance of the estimator for the difference in policy performance. However, the difference estimator is typically biased by the initial conditions. In the problem at hand, differences in policies were underestimated since the policies behaved in a similar manner over the initial portion of each run. A common result of underestimation of the differences in the performances of systems is to retain the current system. This favors the status quo over potential new systems and policies that may in fact be superior.

Unfortunately, actual simulation initialization errors are not as transparent as in the above examples. Problems of bias (e.g. underestimated machine utilizations) and spurious secondary correlations in multiple output measures due to correlations with the initial conditions (e.g. congestion correlated with production) can be difficult to recognize.

Of course, there are situations where initialization effects are not of major concern. Examples of such cases are models where the initial conditions are known completely (e.g. certain elementary combat models) or in which there is no serial dependency in the output (and hence no dependency on the initial conditions of the model). The assumptions that the initial conditions of a model are known a priori or that the output has no serial dependency must be recognized as just that - model assumptions. They should be subjected to the same scrutiny as any other modeling assumption: Is the mathematical benefit gained in simplifying the analysis worth any loss in model accuracy due to the assumption?

In many models, careful examination will indicate that the assumptions necessary to ignore initialization effects are not easily supported. For instance, a simulation of a walk-in medical clinic may start each simulated day with no patients waiting for examinations. This may be equivalent to assuming that no appointments are taken and no patients are told to return for follow-up exams. Clearly, the first day's operation (with no patient backlog) of a real clinic is different than future days' operations. Starting each simulated day with no patients in the system is an assumption that may or may not be justifiable.

Initialization error is often thought not to be a problem in military combat simulation. Certainly, the strengths of the opposing forces are factors in the experiment. However, troop deployments, operational readiness, etc. should also be considered. That is, there are initial conditions that may have greater effects on the simulated combat than force strengths.

## 3. Mathematical effects of initialization

The output from a simulation can be modeled as a sample from a distribution *conditioned* on the manner in which the simulation is initialized. Denote this distribution function as $F_{Y|I}(y|i)$, where the multivariate random variables I and Y are the initial conditions and output, respectively. If somehow the initial conditions could be chosen by 'nature,' they would have the distribution function $F_I(i)$. If estimates of the characteristics of the unconditional distribution are desired, then we wish to sample from

$$F_Y(y) = \int F_{Y|I}(y|i)dF_I(i).$$

The problem is that we do not know $F_I(i)$. Estimates for moments (of Y) will be based on data from $F_{Y|I}(y|i)$ and not from $F_Y(y)$. In general, the expected value of an estimator that uses the conditional data will not be equal to the expected value of the corresponding unconditional estimator. This is commonly mislabelled as 'initialization bias.' It is the output that is 'biased,' not the estimator. Also, the above expression for $F_Y(y)$ suggests that I be included in the design of a simulation experiment, possibly randomized or sampled from a known $F_I(i)$. The inclusion of I as an experimental factor is discussed later.

As an example, consider the familiar M/M/1 queue with traffic intensity less than 1.0. Let $Y_i$ denote the waiting time of the i-th customer, i=1,2,... Suppose that the initial waiting time is $Y_0 = 0$. Denote $X_j$ as the average of the first j customer waiting times, j=1,2,...

$$X_j = (1/j) \sum_{i=1}^{j} Y_i .$$

Law (1983) plots the theoretical values for $E[X_j]$ of an M/M/1 queue as a function of j, the number of customers served. Further, he plots an actual realization of the $X_j$ process. The two graphs appear to be quite different (aside from the fact that both start at the (0,0) coordinate). See Figure 2 of this paper for a similar example. Some interesting work concerning the M/M/1 queue is also contained in Duket (1974).

The initial conditions also have an effect on the variance of estimators. Again, consider the output sequence $Y_0=0, Y_1, Y_2, \ldots$ of an M/M/1 queue. It is noteworthy that the variance of early output observations is actually *smaller* than the variance for later observations. I.e., $\mathrm{Var}(Y_s | Y_0=0) < \mathrm{Var}(Y_t | Y_0=0)$, where s is 'small' and t is 'large.' This phenomenon is illustrated in Figure 3, where we run a number of replications of an M/M/1 queue, each of which is started empty (typically, only a pseudo-random number seed is changed.) The simulated system simply has not had time to change from the initial state early in the run. This variance effect is perhaps more properly viewed as a secondary correlation of the output with the initial conditions.

A popular measure that accounts for both bias and variance of an estimator is the mean squared error (squared bias plus variance). The effect of initial conditions on the mean squared error (m.s.e) of the sample mean for a first-order autoregressive process,

$$Y_t = \mu + \alpha(Y_{t-1} - \mu) + \epsilon_t$$

(where $\{\epsilon_t\}$ is a sequence of constant mean and variance

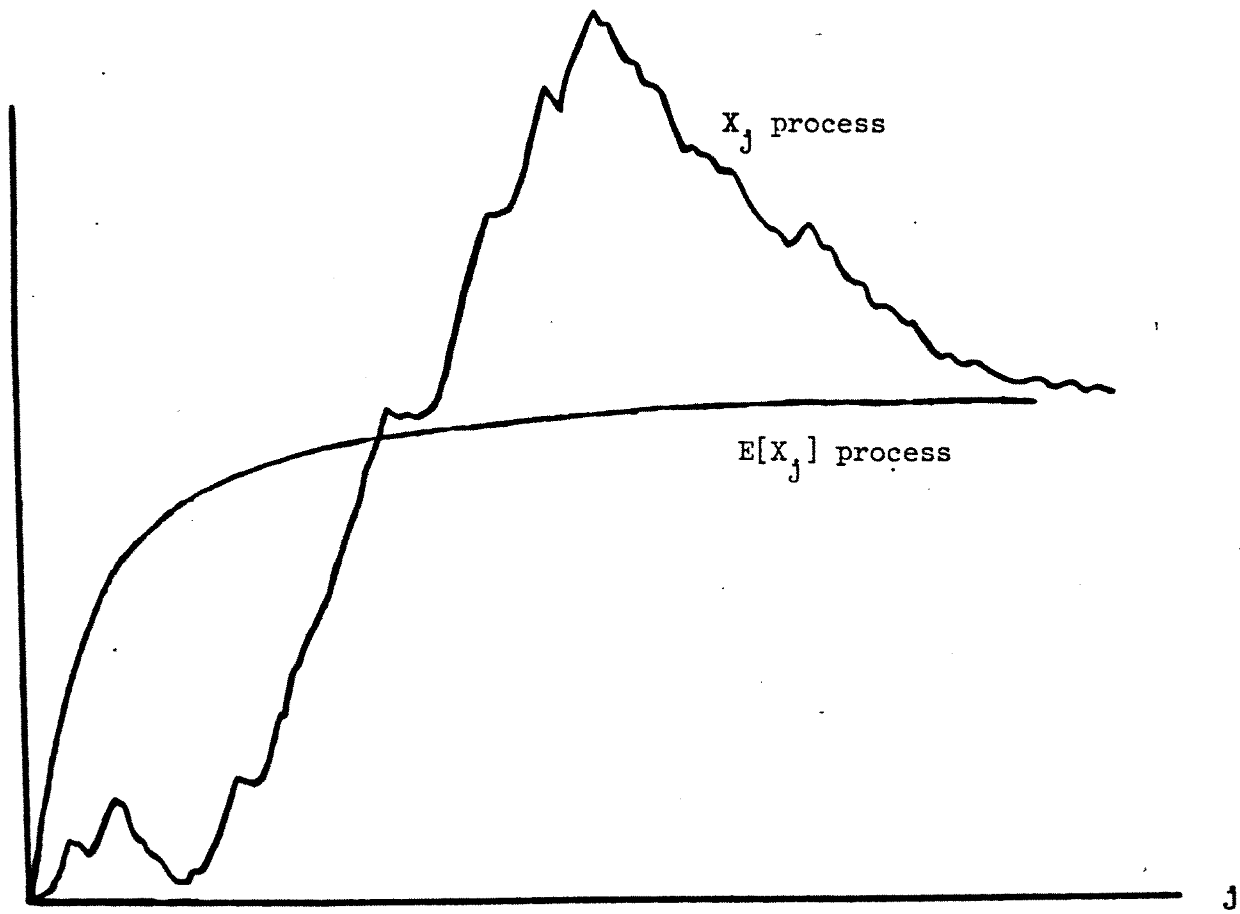


Figure 2: Cumulative average waiting times of first j customers in an M/M/1 queue.

$E[X_j]$ = theoretical expected cumulative average waiting times

$X_j$ = a realization of the cumulative average waiting times

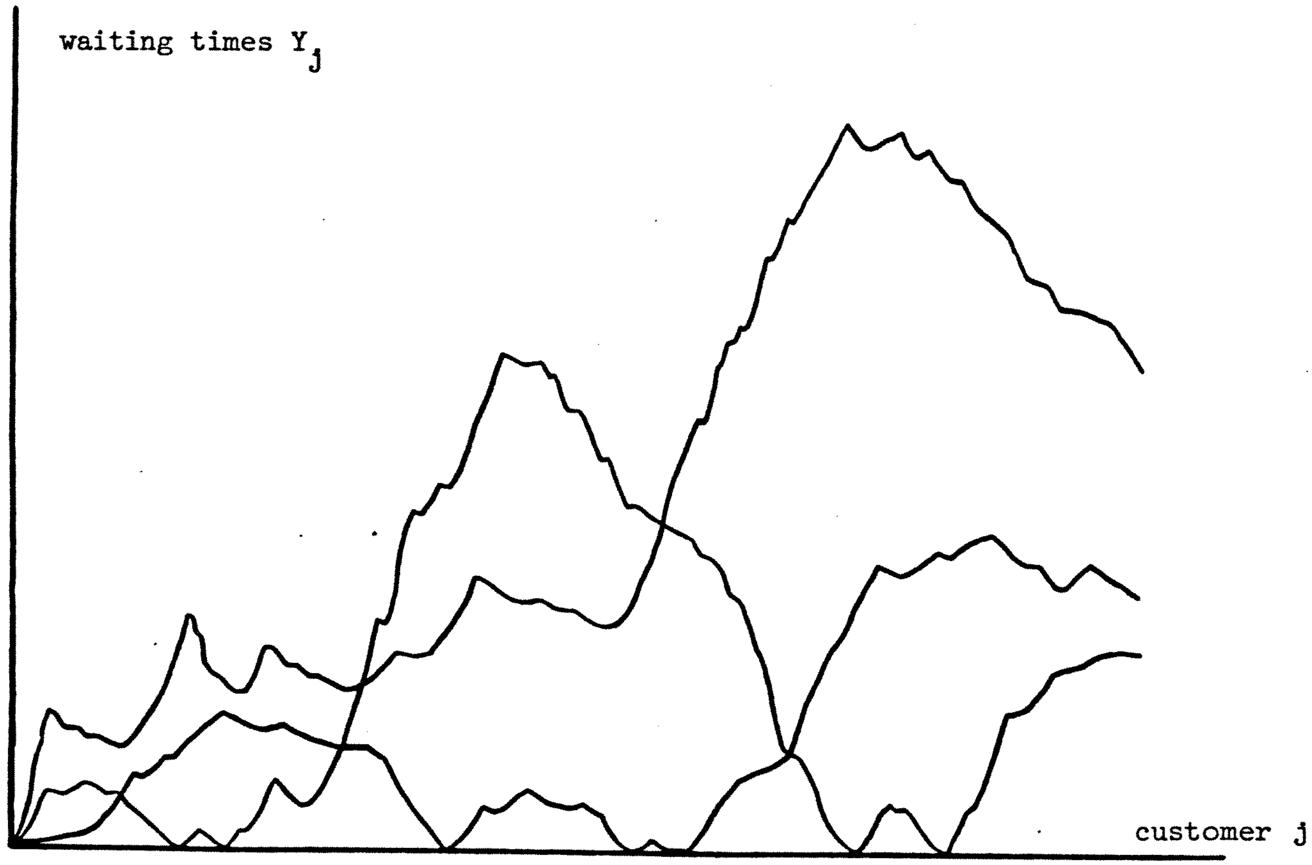


Figure 3: Some realizations of an M/M/1 queue waiting time process. Notice that $V(Y_s \mid Y_0=0) < V(Y_t \mid Y_0=0)$ when s is 'small' and t is 'large.'

uncorrelated random variables),

is examined in Fishman (1972). The lag j serial correlation for this process is,

$$\operatorname{corr}(Y_i, Y_{i+j}) = \alpha^{|j|}.$$

The first-order autoregression is a useful model for illustrating many of the issues in simulation initialization. In his paper, Fishman demonstrates that the common practice of letting the simulation warm up before beginning the collection of data (known as output truncation) is not necessarily advisable (from a mean squared error point of view).

Suppose we have two independent output series from a first-order autoregressive process. Suppose also that one series has been truncated and that one series has not been truncated. Snell and Schruben (1979) give the conditions for which the non-truncated series has a *lower* m.s.e than the truncated series. Several points in that paper are worth mentioning. First, consider the space formed by the possible values for $\alpha$, $\operatorname{Var}(\epsilon_t)$, $Y_0$ (the initial observation), and certain other factors. If the run is rather long, there will be a large region of this space where no truncation is called for. However, no matter how long the series is run, there are initial conditions that are so atypical that truncation will reduce the m.s.e. Of course, the benefit is reduced as the series becomes longer, but truncation is still of value in some cases. This contradicts the notion that truncation is not beneficial if a simulation is run a very long time [see Madanski (1976)]. It can be less costly to truncate some initial data than to overwhelm poor initial conditions with a long run.

The second point is that, from a m.s.e. viewpoint, positive and negative serial correlations have much the same effect. The region where no truncation is optimal is larger for positive $\alpha$ than for negative $\alpha$ (because for negative $\alpha$ the sign of the serial correlation alternates with lag). For the same magnitude of $\alpha$, the amount of positive correlation in the series is greater when this coefficient is positive than the amount of negative correlation when this coefficient is negative.

Data truncation can be viewed as a special case of weighting the output from a simulation. The weight on an observation is zero if it is truncated and one if it is retained for analysis. In their paper, Snell and Schruben also derive optimal weightings of mean estimators for the first-order autoregressive process. Least squares estimators (both ordinary and weighted) of the process mean are presented. Expressions are given for finding optimal truncation points for several different estimator performance criteria (in particular, for minimizing the m.s.e.)

The mathematical effect of secondary correlations among output variables can be illustrated by considering jointly normal estimators, $M_1, M_2, \ldots, M_p$, of parameters $\nu_1, \nu_2, \ldots, \nu_p$. For example, the $M_i$'s might be sample means of different output series that are long enough to justify applying the central limit theorem. Consider two situations: uncorrelated (here, implying independent) estimators and correlated estimators.

Suppose we form joint confidence ellipses for the unknown $\nu_i$'s. The volumes of these ellipses are inversely proportional to the determinant of the dispersion matrix (variance-covariance matrix) of the estimators. The effect of correlation on orthogonal estimators is to reduce the volume of the confidence ellipses [see the proof of

Lemma 1 in Schruben and Margolin (1978)]. The sign of the correlation is not material.

Assume for simplicity of argument that two normally distributed estimators are unbiased and have a common variance, $\sigma^2$. Then correlation between the estimators is beneficial if it is properly taken into account in the analysis [see Schruben (1979)]. When the correlation is not accounted for, then confidence ellipses may be too small. These confidence regions will be less likely to cover the unknown parameters, $\nu_1, \nu_2, \ldots, \nu_p$, than anticipated.

## 4. The problems of initialization

The simulation literature contains many papers on *'the initialization bias problem'* (see the bibliography at the end of this paper.) Probably as much has been written about this issue as any other single area of simulation output analysis. There does not seem to be any definitive solution to the problem. Indeed, there is not any universally accepted definition of the problem. No attempt at such a definition will be presented in this paper. We offer the suggestion that there is not just a single issue in initialization of simulations; rather, there are several.

At the highest level, there are issues that can be loosely referred to as academic issues and practical issues. The two sets are closely related, but they are not the same. Different approaches are needed.

The academic focus has been on obtaining accurate (low bias) and precise (low variability) estimators of a measure of simulated system performance. There is a basic trade-off between these two objectives. Meanwhile, the practical concern is that the effects of initialization do not lead to erroneous experimental conclusions. A 'significant' (in the statistical sense) initialization effect for the academic is somewhat arbitrary. A 'significant' initialization effect for the practitioner is one that indicates a wrong decision. A closer look at the academic question will be followed by a discussion of the practical question.

### Academic issues:

The academic question involves estimating what is often called the 'steady state' performance of a simulated system. The concept of a steady state for a stochastic process is inherited from queueing theory where it has great mathematical utility. Steady state analysis may or may not have particular interest in practical systems studies. Say that the output at simulated time t, $Y_t$, has a distribution function denoted by $F_t$. The problem is to estimate properties (e.g. moments, quantiles) of the limiting distribution,

$$F = \lim_{t \to \infty} F_t.$$

The difficulty is that in many simulations (e.g. the M/M/1 queue), the limiting distribution is *asymptotic*. None of the data in the output series will be sampled from the distribution function F. Thus, the problem has characteristics similar to forecasting problems [see Kimbler, et. al. (1979) and Snell (1980)]. Also, note that the problems of initialization are closely related to the decision of how long to run a simulation program. Truncation will not actually eliminate initialization bias in such 'asymptotic' systems no matter how long the run.

If the simulation run is very short, good estimation is not possible, regardless of the truncation procedure which we choose to use. The objectives of procedures for dealing with this problem have varied (and have often been only heuristic). In general, one tries to obtain as good an estimator as possible recognizing that without data from the population of interest, compromise will be necessary. Some of the criteria for judging the performance of various procedures for dealing with this problem are presented in Gafarian, et. al. (1977). There are two popular estimators of interest when considering the effectiveness of a simulation initialization bias control procedure.

One is the point estimator (for the parameter in question); the other is an interval estimator or confidence interval. We desire an initialization bias control procedure which yields 'good' point estimates and/or confidence intervals.

In point estimation, accuracy (measured by bias) and precision (measured by variance) have been considered important. There is usually a trade-off between estimator bias and variance; decreasing one often means that the other will increase. As mentioned earlier, mean squared error is a popular (but more or less arbitrary) criterion for combining the bias and variance of an estimator into a single quantity [see, for example, Blomqvist (1970), Fishman (1972), Law (1982), and Wilson and Pritsker (1978a)].

There have been two general approaches to this problem of point estimation. By far the most attention has been given to weighting of the data. As discussed previously, data truncation is a special case of weighting. In fact, truncation simply allows the simulation to warm up before data are retained for analysis. Many 'truncation rules' have been proposed and studied [see, for example, Gafarian, et. al. (1978), Kelton (1980), Morisaku (1976), Sargent (1979), and Wilson and Pritsker (1978a,1978b)]. The consensus is that simple truncation rules do not in general perform well in all situations.

The second approach is to directly model the transient mean function. In Snell (1980), economic growth models are used for this purpose. In Narasimhan, et. al. (1982) and Richards (1983), so-called 'intervention time series' models are used. More experience is needed with both of these transient models before their merits can be judged.

We now consider the problem of confidence interval estimation for some parameter of interest. In the recent simulation literature, confidence interval estimator performance has been measured using the following criteria [see, e.g., Schmeiser (1982)]:

1) observed interval coverage frequency (of the parameter in question) for a sample of confidence intervals that use the procedure,

2) interval width, typically measured by the sample average half-width of interval estimators using the procedure, and

3) interval stability, typically measured by the sample variance of the observed interval half-widths.

Some authors use the coefficient of variation (c.v.) of the interval half-width as another criterion. This can be misleading: the c.v. might be reduced (an apparent improvement) by merely increasing the interval width (not an improvement).

Clearly, initialization bias may result in poor confidence interval performance in terms of all of the above criteria. The effect of bias on interval coverage can be seen using the coverage function in Schruben (1980). The trade-off between interval half-width and coverage for various confidence interval estimators can be studied using a graphical technique developed in Kang and Schmeiser (1983). They suggest plotting the upper confidence interval bounds against the lower confidence interval bounds for a sample of several interval estimates. A $45^{o}$ rotation of this plot would be a graph of the observed confidence interval center points against the observed half-widths in a sample of confidence interval estimators. See Figure 4. Coverage frequency, estimator bias, half-width bias, as well as dependencies between the center point estimator and interval half-width estimator can also be seen from such a plot.

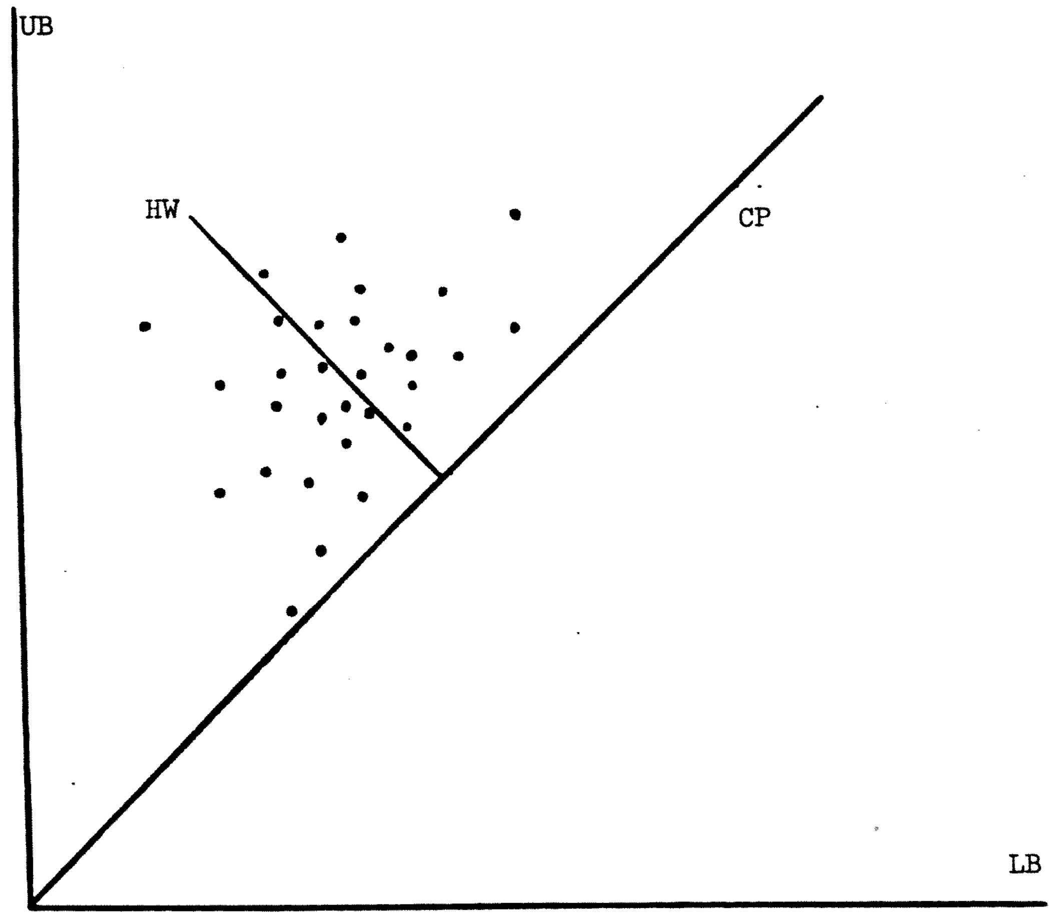


Figure 4: Kang and Schmeiser's scatter diagram technique for a sample of 30 confidence interval estimators.

UB = observed upper confidence bound
LB = observed lower confidence bound
CP = observed confidence interval center point
HW = observed confidence interval half-width

It is not too surprising that truncation rules do not in general produce good confidence interval coverage for some of the systems tested in the literature. The most popular systems for testing simulation procedures have been simple queueing systems. These systems only asymptotically reach steady state and can have persistent serial dependence [see Schruben (1978) in his reply to Fox (1978)]. None of the data is from the (steady state) population that is actually being studied. See Fox (1978), Kleijnen (1979), and Schruben (1978) for lively but inconclusive discussions of some of the issues concerning initialization effects.

Practical issues:

The practical aspects of simulation initialization all relate to the purpose of the simulation study. If system performance estimation is an objective, then the same academic issues discussed earlier apply. However, many simulation models are used as aids in system design and evaluation. Here, simulation initialization effects must be considered in the context of a decision problem. In design studies, the concern is whether a better design could be possible if initialization errors were not present. In evaluation studies, the question is whether or not the presence of initialization errors favors other than the best of several alternatives that are being studied. In short, does the presence of initialization effects potentially result in an incorrect decision?

The authors are not aware of any definitive research that ties initialization error control directly with problems in decision and design. Perhaps the paper by Adlakha and Fishman (1979) is an exception. There it is suggested that overestimating system congestion is preferred to underestimating system congestion in queueing systems; simulations of such systems should be initialized accordingly. Adlakha and Fishman did not explicitly relate their work to decision making. Naturally, if system service level is the paramount consideration, then underestimating system congestion is again undesirable. Alternatively, if system cost is important, then overestimation of system congestion may result in an expensive, over-designed system.

A straightforward approach for controlling simulation initialization effects in practical studies is to sequentially truncate various amounts of the initial data and simply see if this changes the decisions indicated by the full output data set. This can be impractical in large simulation studies. A sequential procedure similar to that suggested in Kelton (1980) might be adapted effectively in a decision making context. Kelton tested the slope of the output regressed on simulated clock time for a zero slope.

The tests for initialization effects suggested in Schruben (1981,1982) and Schruben, Singh, and Tierney (1982) appear to be more powerful and robust than the regression slope test used by Kelton. In these tests, a quality control viewpoint is taken (to control for inconsistent 'quality' in the simulation output data). Goldsman and Schruben (1983) give generalizations of these tests.

Another approach is to treat simulation initialization effects as a nuisance factor in the design of simulation experiments. For example, the various initial values might be considered as a factor in an ANOVA analysis. Several possible initial states might be run for each set of experimental factors. One might block such experiments on the initial conditions or regress the output on the initial conditions to try to control this source of error. Indeed, consider the beneficial effects in terms of variance reduction that might result from blocking on the initial conditions: Estimators based on the outputs from runs with the same initial conditions could be expected to be positively correlated. Also, estimators from runs with radically different initial conditions might be expected to have negative correlations. See Schruben and Margolin (1978) and Schruben (1979) for a discussion of the relationships among blocking, correlation of the outputs, and variance reduction.

The experimental design approach probably deserves more attention than it has received. It may still be of limited practical value since there are usually a large number of variables to be initialized in a simulation model.

<u>State of practice</u>:

Current practice seems to be to simply look at plots of the output to try to visually detect any initialization effects. (In any event, one should look at the plots of the output, if feasible.) However, this can be very time consuming in large scale simulation experiments involving many runs. Also, visually scanning the output might not actually be of much benefit (recall Figure 1 of this paper). The smoothed output (a moving average, say) of several pilot runs as suggested in Sargent (1979) and Welch (1981) makes visual analysis easier. Averaging observations across several runs (i.e., the first observations from each run are averaged, the second observations are averaged, and so forth) also helps in the detection of initialization effects upon the mean of the output series. The 'CUSUM' type plots in Schruben (1979) are particularly sensitive to changes in the mean of a data series; these plots can be most useful in detecting initialization effects.

In Heidelberger and Welch (1982), an automatic test for initialization bias based on Schruben (1982) was included in a confidence interval procedure. This resulted in improved performance over an earlier version of the procedure that did not have any initialization error control [see Heidelberger and Welch (1981)]. However, automatic statistical procedures for initialization error control will probably not perform as well as an experienced analyst looking at a plot of the output. Such procedures and tests for initialization bias are useful when there is a large amount of data to analyze. The analyst simply does not have the time to plot and look at all of the data. Of course, the condensation of information contained in data (so that it can be understood more easily) is one of the purposes of statistics. The better the information is condensed, the closer the decisions based on sample statistics will be to those based on the entire data set.

Conclusions:

Initialization bias has been recognized as a difficult problem involving many issues. Much progress has been made in dealing with this question. Indeed, simple 'truncation rules' are being replaced by sequential procedures such as those given in Kelton (1980) and Heidelberger and Welch (1982). More attention should be given to the problem in the decision making and experimental design contexts. Also, we note that different issues may be involved in simulation experiments with a single system (or single performance measure) vs. experiments with multiple systems (or multiple performance measures). Finally, simulation initialization errors may have different effects depending on whether the simulations are used for design, optimization, evaluation, selection, or feasibility decisions. The problems of initializing a simulation program are important, and the issues are not at this time completely understood.

## REFERENCES


Adlakha, V.G. and Fishman, G.S. (1979). A starting rule for data collection in queueing simulations. Technical Report 79-3, Curriculum in Operations Research and Systems Analysis, University of North Carolina, Chapel Hill.

Adlakha, V.G. and Fishman, G.S. (1982). Starting and stopping rules for simulations using a priori information. European Journal Operational Research 10, 379-394.

Anderson, O.D. (1980). On warming-up time series simulations generated by Box-Jenkins models. J. of the Operational Research Society 30, 587-589.

Blomqvist, N. (1970). On the transient behavior of the GI/G/1 waiting times. Skand. AktuarTidskr., 118-129.

Cheng, R.C.H. (1976). A note on the effect of initial conditions on a simulation run. Operational Research Quarterly, Vol. 27, No. 2, 467-470.

Conway, R.E. (1963). Some tactical problems in digital simulation. Management Science, Vol. 10, No. 1, 47-61.

Crane, M.A. and Iglehart, D.L. (1974a). Simulating stable stochastic systems, I: general multiserver queues. Journal of A.C.M., Vol. 21, 103-113.

Crane, M.A. and Iglehart, D.L. (1974b). Simulating stable stochastic systems, II: Markov chains. Journal of A.C.M., Vol. 21, 114-123.

Crane, M.A. and Iglehart, D.L. (1975). Simulating stable stochastic systems, III: regenerative processes and discrete event simulations. Operations Research, Vol. 23, No. 1, 33-45.

Deutsch, J. and Richards, J. (1981). The initialization problem in simulation. Technical Report J-80-17, Industrial Systems Engineering, Georgia Tech., Atlanta, Georgia.

Duket, S.D. (1974). Simulation Output Analysis. Pritsker and Assoc., West Lafayette, Indiana.

Eilon, S. and Chowdhury, I.G. (1974). A note on steady-state results in queueing and job-shop scheduling. Simulation, 85-87.

Fishman, G.S. (1967). Problems in the statistical analysis of simulation experiments: the comparison of means and the length of sample records. Communications of the A.C.M., Vol. 10, 94-99.

Fishman, G.S. (1971). Estimating sample size in computer simulation experiments. Management Science, Vol. 18, No. 1, 21-38.

Fishman, G.S. (1972). Bias considerations in simulation experiments. Operations Research, Vol. 20, No. 4, 785-790.

Fishman, G.S. (1978). Grouping observations in digital simulation. Management Science, Vol. 24, No. 5, 510-521.

Fishman, G.S. (1980). A perspective and proposal for the initial transient problem in simulation. Tech. Report 80-6, Curriculum in Operations Research and Systems Analysis, University of North Carolina, Chapel Hill.

Fishman, G.S. and Moore, L.R. (1978). Starting times for data collection in a queueing simulation I: experiments with a single server model. Tech. Report 78-1, Curriculum in Operations Research and Systems Analysis, University of North Carolina, Chapel Hill.

Fox, B. (1978). Estimation and simulation. Management Science, Vol. 24, 860-861.

Gafarian, A.V. and Ancker, C.J., Jr. (1966). Mean value estimation from digital computer simulation. Operations Research, Vol. 14, No. 1, 25-44.

Gafarian, A.V. and Ancker, C.J., Jr. (1976). Some additional results on the problem of the initial transient. ORSA/TIMS National Meeting, Miami, Florida.

Gafarian, A.V., Ancker, C.J., Jr. and Morisaku, T. (1977). The problem of the initial transient with respect to mean value in digital computer simulation and the evaluation of some proposed solutions. Technical Report 77-1, Department of Industrial and Systems Engineering, University of Southern California, Los Angeles, California.

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.

Goldsman, D. and Schruben, L.W. (1983). Extensions of some tests for initialization bias in simulation output. Tech. Report (in preparation), School of OR&IE, Cornell University, Ithaca, NY.

Grassmann, W.K. (1982). Initial bias and estimation error in discrete event simulation. Proceedings Winter Simulation Conference, San Diego, California.

Hanscomb, D. (1969). Monte Carlo techniques: theoretical. Design of Computer Simulation Experiments. (T. Naylor, ed.), Duke University Press.

Hazen, G. and Pritsker, A.A.B. (1980). Formulas for the variance of the sample mean in finite state Markov processes. J. Statist. Comput. Simul. 12, 25-40.

Heidelberger, P. (1978). Steady state estimation from finite length simulation. I.B.M. Research Report RC 7451 (#32153), Yorktown Heights, NY.

Heidelberger, P. and Welch, P.D. (1981). A spectral method for confidence interval generation and run length control in simulations. Comm. Assoc. Comput. Mach. 24, 233-245.

Heidelberger, P. and Welch, P.D. (1982). Simulation run length control in the presence of an inital transient. Report RC 9393, Computer Science, I.B.M. Watson Research Center, Yorktown Heights, NY 10598.

Kang, K. and Schmeiser, B. (1983). Evaluation and comparison of confidence interval procedures. Tech. Memo 83-3, School of IE, Purdue University, W. Lafayette, Ind.

Kelton, W.D. (1980). The startup problem in discrete-event simulation. Technical Report No. 80-1, Department of Industrial Engineering, University of Wisconsin.

Kelton, W.D. (1982). Initial transient techniques for stochastic simulation. Technical Report No. WS-8204, Department of Administrative Sciences, Kent State University. Also to appear in Encyclopedia of Systems and Control, Pergamon Press, 1983.

Kelton, W.D. and Law, A.M. (1981a). A new approach for dealing with the startup problem in discrete event simulation. Tech. Report No. 81-2, Dept. of Management Information Systems, University of Arizona, Tucsón, AZ. Also to appear in Naval Research Logistics Quarterly, 1983.

Kelton, W.D. and Law, A.M. (1981b). An analytical evaluation of alternative strategies in steady-state simulation. Technical Report No. 81-3, Department of Management Information Systems, University of Arizona. Also to appear in Operations Research 1983.

Kelton, W.D. and Law, A.M. (1982). The transient behavior of the M/M/s queue; with implications for steady-state simulation. Tech. Report 82-7, Dept. of Management Information Systems, University of Arizona, Tucson, AZ.

Kimbler, D.L., Jr. (1978). Applications of exponential smoothing to the initial transient problem in digital simulation. M.S. Thesis, Department of Industrial Engineering and Operations Research, Virginia Polytechnic Institute and State University.

Kimbler, D.L., Schmidt, J.W. and Davis, R.P. (1979). Bias reduction and initial conditions in digital simulation. Reprint Series No. 7905, Department of Industrial Engineering and Operations Research, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24016.

Kleijnen, J.P.C. (1979). The simulation startup problem. Simulation, Vol. 32, 29.

Kleijnen, J.P.C. (1983). Statistical analysis of steady-state simulations: survey of recent progress. Dept. of Economics, Tilburg University, Netherlands.

Lavenberg, S.S., Fishman, G.S., Kelton, W.D., Schruben, L.W. and Welch, P.D. (1981). The initial transient problem in steady state simulations. Proceedings of the 1981 Winter Simulation Conference.

Law, A.M. (1977). Confidence intervals in discrete event simulation: a comparison of replication and batch means. Naval Research Logistics Quarterly, Vol. 24, 667-678.

Law, A.M. (1980). Statistical analysis of the output data from terminating simulations. Naval Research Logistics Quarterly, Vol. 27, 131-143.

Law, A.M. (1982). Statistical analyses of simulation output data: the state of the art. Tech. Report 82-1, Dept. of Management Information Systems, University of Arizona, Tucson, AZ. Also to appear in Operations Research.

Law, A.M. and Kelton, W.D. (1978). Confidence intervals for steady-state simulations, II: a survey of sequential procedures. Technical Report 78-6, Department of Industrial Engineering, University of Wisconsin.

Law, A.M. and Kelton, W.D. (1979). Confidence intervals for steady-state simulations, I: a survey of fixed sample size procedures. Tech. Report 78-5, Dept. of IE, University of Wisconsin, Madison, Wis.

Madansky, A. (1976). Optimal initial conditions for a simulation problem. Operations Research, Vol. 24, No. 3, 572-577.

Meketon, M. (1979). Analysis of the regenerative method for finite run lengths. School of Operations Research and Industrial Engineering, Cornell University, Technical Report 400, Ithaca, NY.

Morisaku, T. (1976). Techniques for data-truncation in digital computer simulation. Ph.D. dissertation, Dept. of Industrial and Systems Engineering, University of Southern California.

Narasimhan, R., Rubin, P.A. and Melnyk, S.A. (1982). Analyzing simulation output data using time series intervention analysis. Technical Report, Graduate School of Business, Michigan State University, East Lansing, Michigan 48824.

Richards, J.E. (1983). Minimizing run length in simulations subject to initial transients. Ph.D. dissertation, School of Industrial and Systems Engineering, Georgia Institute of Technology, Altanta, Georgia.

Robinson, D.W. (1976). Determination of run length in simulations of stable stochastic systems. Technical Report 86-21, Control Analysis Corporation, Palo Alto, California.

Sargent, R.G. (1979). An introduction to statistical analysis of simulation output data. Proceedings of the AGARD Symposium, Paris, France.

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

Schriber, T.J. and Andrews, R.W. (1981). A conceptual framework for research in the analysis of simulation output. Comm. A.C.M. 24, 218-232.

Schruben, L.W. (1977). Approaches to simulation initialization. Talk presented at the 1977 Joint ORSA/TIMS National Meeting in Atlanta, Georgia.

Schruben, L.W. (1978). Reply to Fox. Management Science, Vol. 24, No. 8, 862.

Schruben, L.W. (1979). Designing correlation induction strategies for simulation experiments. Current Issues in Computer Simulation. (Adam, Dogramaci, Eds.) Academic Press, NY.

Schruben, L.W. (1980). A coverage function for interval estimators of simulation response. Management Science, Vol. 26, 18-27.

Schruben, L.W. (1981). Control of initialization bias in multivariate simulation response. Communications of the A.C.M., Vol. 24, 246-252.

Schruben, L.W. (1982). Detecting initialization bias in simulation output. Operations Research, Vol. 30, No. 3, 569-590.

Schruben, L.W. and Kulkarni, R. (1982). Some consequences of estimating parameters for the M/M/1 queue. Tech. Report 460, School of OR&IE, Cornell University, Ithaca, NY. Also to appear in OR Letters.

Schruben, L.W. and Margolin, B. (1978). Pseudorandom number assignment in statistically designed simulation and distribution sampling experiments. Journal of the American Statistical Association 73, 363, 504-25.

Schruben, L.W., Singh, H. and Tierney, L. (1982). Optimal tests for initialization bias in simulation output. Tech. Report 471, School of OR&IE, Cornell University, Ithaca, NY.

Schwarz, L.B. (1976). An examination of the United States Air Force (Q,R) policies for managing depot-base inventories. Technical Report from the University of Rochester Graduate School of Management.

Shannon, R.E. (1975). Simulation: a survey with research suggestions. AIIE Transactions, Vol. 7, No. 3, 289-301.

Snell, M.K. (1980). The application of regression methods to the initial transient problem in computer simulations. Unpublished Ph.D. thesis, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, NY 14853.

Snell, M.K. and Schruben, L.W. (1979). The use of weighting functions to correct for initialization bias in the AR(1) process. Technical Report No. 395, School of Operations Research and Industrial Enginering, Cornell University.

Turnquist, M.A. and Sussman, J.M. (1977). Toward guidelines for designing experiments in queueing simulation. Simulation, Vol. 28, No. 5, 137-144.

Welch, P.D. (1981). On the problem of the initial transient in steady state simulations. I.B.M. Watson Research Center, Yorktown Heights, NY.

Wilson, J.R. and Pritsker, A.A.B. (1978a). A survey of research on the simulation start-up problem. Simulation, Vol. 31, No. 2, 55-58.

Wilson, J.R. and Pritsker, A.A.B. (1978b). Evaluation of startup policies in simulation experiments. Simulation, Vol. 31, No. 3, 79-89.