

Reliable Results from
**STOCHASTIC
SIMULATION MODELS**



**by Donald L. Gochenour, Jr.
Leonard R. Johnson**

USDA FOREST SERVICE RESEARCH PAPER NE-277
1973

FOREST SERVICE, U.S. DEPARTMENT OF AGRICULTURE
NORTHEASTERN FOREST EXPERIMENT STATION
6816 MARKET STREET, UPPER DARBY, PA. 19082
WARREN T. DOOLITTLE, DIRECTOR

The Authors

DONALD L. GOCHENOUR, JR., an assistant professor in industrial engineering at West Virginia University, received his Ph.D. from that institution, in 1972. He also received his B.S. and M.S. degrees in industrial engineering from West Virginia University and worked for the engineering work unit in Morgantown, NEFES, as a research industrial engineer from 1967 to 1968.

LEONARD R. JOHNSON works as a research industrial engineer with the Northeastern Forest Experiment Station, engineering work unit, Morgantown, W. Va. He graduated in 1968 from Montana State University with a B.S. degree in industrial engineering and received an M.S. degree from Montana State University in 1970.

MANUSCRIPT RECEIVED FOR PUBLICATION 30 MARCH 1973

Reliable Results from **STOCHASTIC SIMULATION MODELS**

ABSTRACT

Development of a computer simulation model is usually done without fully considering how long the model should run (e.g. computer time) before the results are reliable.

However construction of confidence intervals (CI) about critical output parameters from the simulation model makes it possible to determine the point where model results are reliable. If the results are not reliable, due to the variability within the model, the model user at least knows the degree of confidence that may be placed in the results.

Such factors as: 1) model variability, 2) prescribed width of the CI, and 3) the level of confidence of the simulation model results, have a profound effect on the amount of computer time needed to obtain reliable results. Since proper choice of these parameters can significantly affect the time and cost needed to obtain the results from the model, it is imperative that the user of the model consider these factors.

COMPUTER SIMULATION models of existing or hypothesized systems are often developed to obtain information about one or more system parameters. The simulation model provides the analyst with a tool for generating values of the system variables under study. However, the results obtained from stochastic simulation models are subject to the same questions of statistical reliability — precision of parameter estimates — as results of any other experiment. The analyst needs to know the statistical reliability of the results (*Reese 1968*) since accepting an unreliable estimate, because the simulation experiment was terminated too soon, could be costly when results are applied. Conversely, further computations after the results are reliable are unnecessary and wasteful.

The precision of estimates obtained from a simulation model depends on the length of the stochastic simulation run in the same way that the precision of a real-life experiment depends on the sample size. Failure to consider the precision of estimate may cause premature termination of the simulation experiment, resulting in unreliable parameter estimates or estimates no more reliable than those found by intuition or past experience in less time and at lower cost.

METHODS

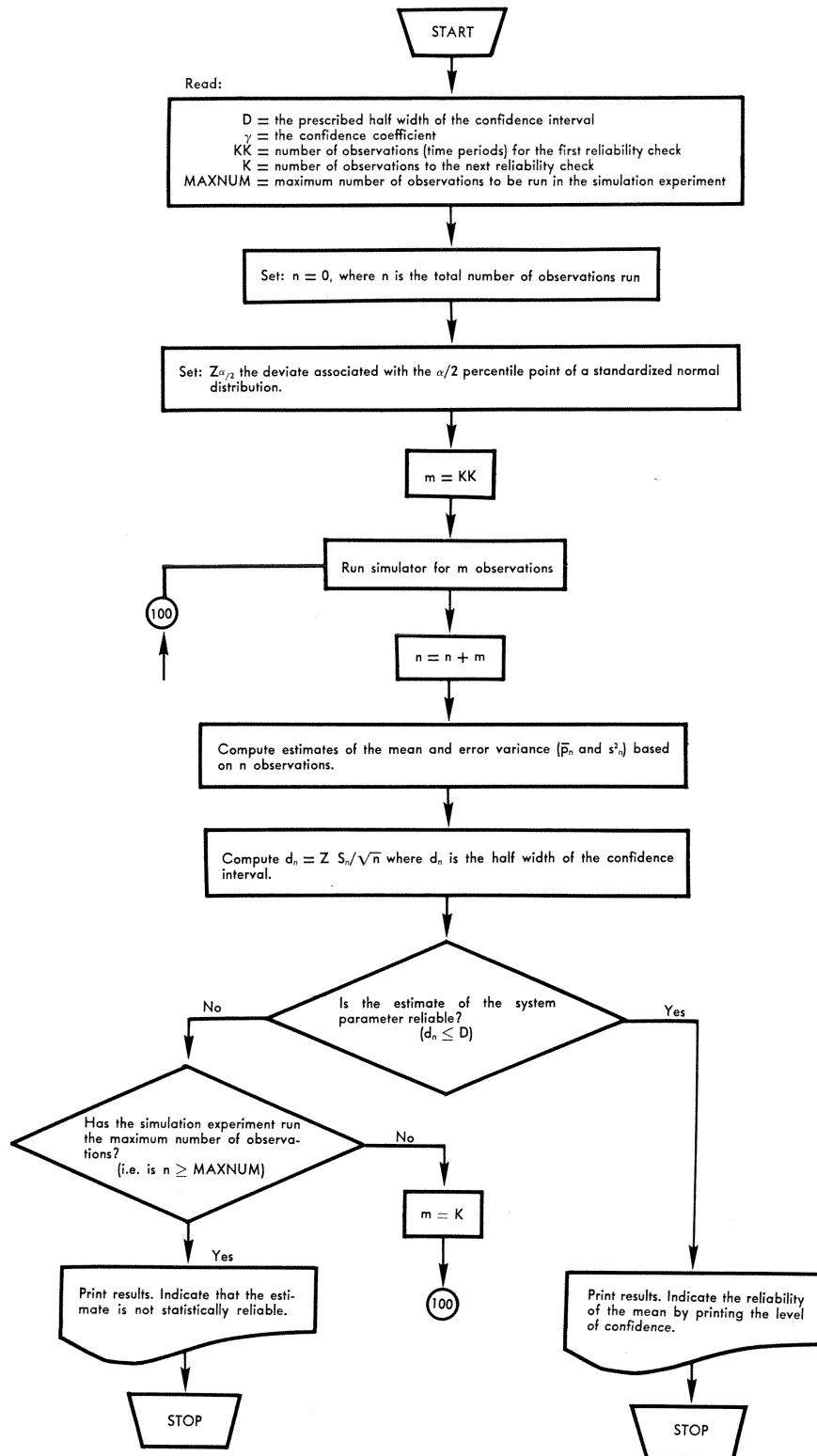
The sample size can be determined from interval estimates of the mean. This procedure involves setting a prescribed interval half width (D) and a probability that the population parameter falls within a distance D of the estimate and then calculating the sample size (n) needed to meet these requirements. This approach requires knowledge of the population variance (σ^2) or the assumption of a normally distributed variable. One advantage of simulation experiments is that obser-

vations can be obtained sequentially, and the results from previous observations can be retained. (One observation equals one time period in the simulator, where one time period could be 1 hour, 1 day, 1 week, 1 year, etc.) This allows one to make use of the past values in deciding whether the parameter estimates are reliable.

There are various procedures available that can be used to signal the termination of the simulation model. Reese used methods based on Chebychev's inequality, trend-corrected t tests, and sequential t tests to obtain parameter estimates after runs had reached steady state conditions. Stein's (1945) two-sample method, could also be used to stop the simulator run. However, in these works restrictions are imposed on the observations from the simulator. For example, normality of the variable under study has to be assumed or the simulator must run long enough to obtain a reliable estimate of σ^2 .

A procedure that can be used to obtain results reliable at a prescribed probability level, and that is less restrictive than the methods previously discussed, can be obtained using a theorem by Chow and Robbins. This procedure involves estimating a confidence interval (CI) with coefficient γ and termination of the simulator when the half width of the confidence interval (d_n) is less than a prescribed half width (D). If σ^2 is finite, the method will satisfy any (γ, D) reliability condition for n large enough. With σ^2 unknown and D sufficiently small, the population parameter u_p will lie within the estimated interval, centered about the sample mean, with probability γ (*Chow and Robbins 1965*).

For example, suppose we want to obtain a reliable estimate of the average daily profit u_p by operating a given system for an unknown period of time. The sample size must be large enough to provide reasonable assur-



ance that $\bar{p}$, the estimated average daily profit, is close to u_p . Misleading results could occur if the sample size is too small. Using a sequential stopping procedure adapted from Chow and Robbin's theorem, it is necessary to make D small enough and perform the series of steps presented in figure 1.

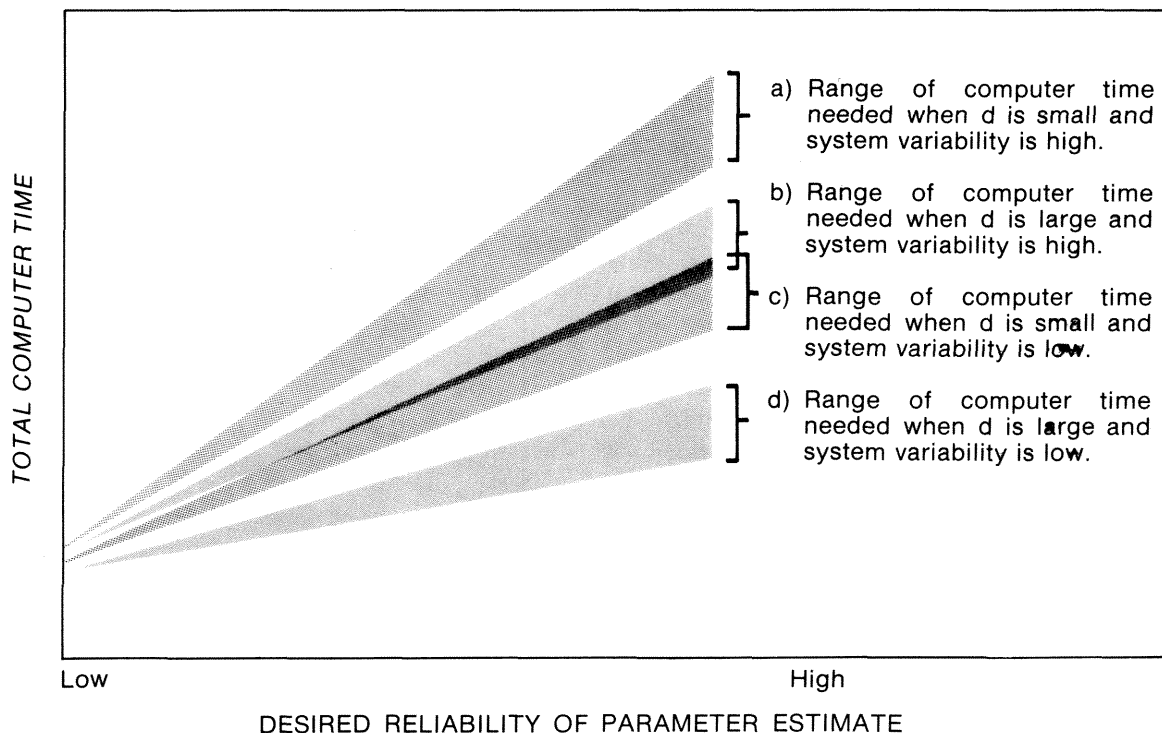
This procedure, resulting in a fixed width confidence statement that holds asymptotically makes the sample size (n) equal to the smallest integer multiple of K greater than or equal to $Z^2 S_n^2 / d_n^2$ and KK . Rearranging n , from above, to appear as $d_n = Z S_n / n^{1/2}$ a reasonable procedure for ending the simulator is to stop when the simulator-estimated half width is shorter than the prescribed half width D . Although the actual confidence coefficient is no longer exactly γ due to sequential sampling, this procedure can be used to signal an end to the simulation experiment. If the parameter estimate is assured reliability $\pm D$ at approximately the confidence coefficient γ , the results can be confidently used. If the variability is so large that reliable

estimates are too costly to obtain, the simulator will stop when the maximum number of observations MAXNUM is reached.

DISCUSSION

Examination of the flow chart (fig. 1) indicates several decisions that affect the number of observations required. A decision must be made on the prescribed half width D and on the level of confidence γ . These two values, interacting with inherent variability of the simulator, have a direct effect on the length of the simulator run. Theoretically, as shown by Emshoff and Sisson (1970) for any simulation experiment, the length of the computer run will be proportional to the variance (see fig. 2 a and c). This would be true whether D was small or large (see fig. 2 a and b). When the variability is high, the observations vary so much that the half width will exceed the prescribed width until n gets very large. If the model variability is small, the half width becomes less than D sooner. Figure 2, which

Figure 2. — Total computer time of simulation experiment as a function of the reliability of the estimate.



makes use of Emshoff and Sisson's work, indicates general results from any arbitrary simulation model while showing the relationship between the computer time needed for the simulation experiment, the magnitude of D , and model variability. It shows that the user of the simulation model must be very careful when choosing values for D and γ .

Even though the value of the prescribed half width D is set by the analyst, the choice is not independent of the parameters of the simulation model. D must be small relative to σ^2 and relative to the mean. A rule that is often used in calculating the sample size is to set D equal to ± 10 percent of the mean. In many cases, the analyst will have preliminary runs of the simulator available to use for estimating the mean and variance. It may also be possible to derive approximations of the mean and variance by analytical methods. It is necessary to obtain reliable results from the model, but only if the cost and time required to obtain these results do not outweigh the potential benefits of using the model.

The following is an example of the stopping rule outlined in figure 1 and the relationships shown in figure 2:

An individual has contracted to do the log hauling from a logging site 12 miles to the sawmill. The logs are purchased at \$29 per thousand board feet (Mbf) from the logger and sold to the mill for \$34 per Mbf. The contract requires six loads per day. The contractor is presently hauling the six loads with two trucks, but is interested in the effect on total profit if one additional truck is used. The costs (owning, fuel, maintenance, etc.) involved in running two trucks are known and provide data for estimating the costs for a third truck.

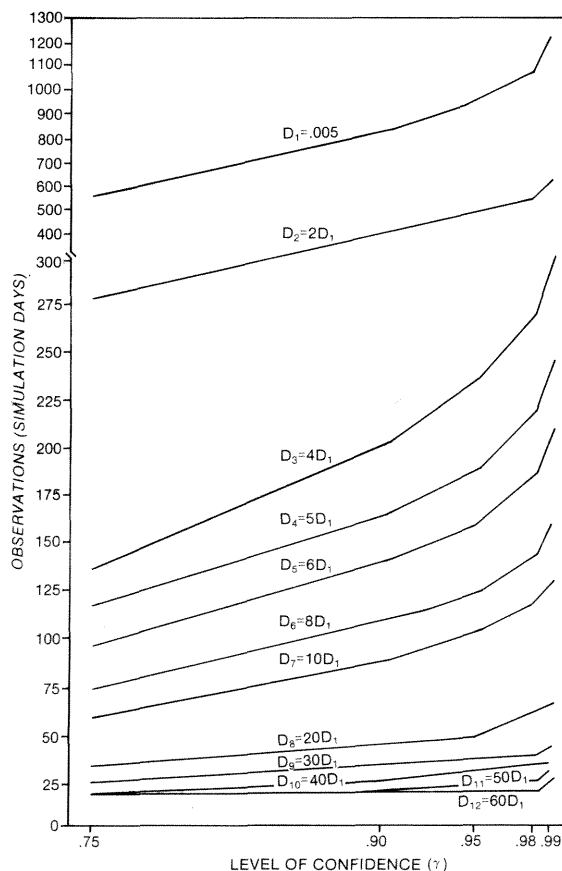
The profit for a three-truck system is determined by using a model which simulates the hauling operation through distributions of loading time, hauling time, and operator efficiency. The system is treated as a single-channel queue with the loader as the queuing point where trucks may wait for loading while another truck is loaded.

The parameter of primary interest in the simulation model is average profit per day. Since a decision whether to purchase the

third truck will be based upon this parameter, it must be reliable. The mean profit should be within D units of the true value ($100 \times \gamma$) percent of the time. However, the cost of obtaining profit estimates must not be too large.

The simulation model was run using different values for D , γ , and σ^2 and graphs were plotted to show the resulting relationships. (Note: the graphs serve to point out the results from this particular example and should not be interpreted for any other use.) Figure 3 shows the length of computer runs for various γ and D values. As the confidence coefficient increases for a fixed D value the computer time increases. Conversely, as D is decreased for a fixed γ the time increases. The analyst may want to choose γ and D values between the extremes of a small γ and a

Figure 3. — Number of simulated days versus level of confidence (γ) for various levels of D .



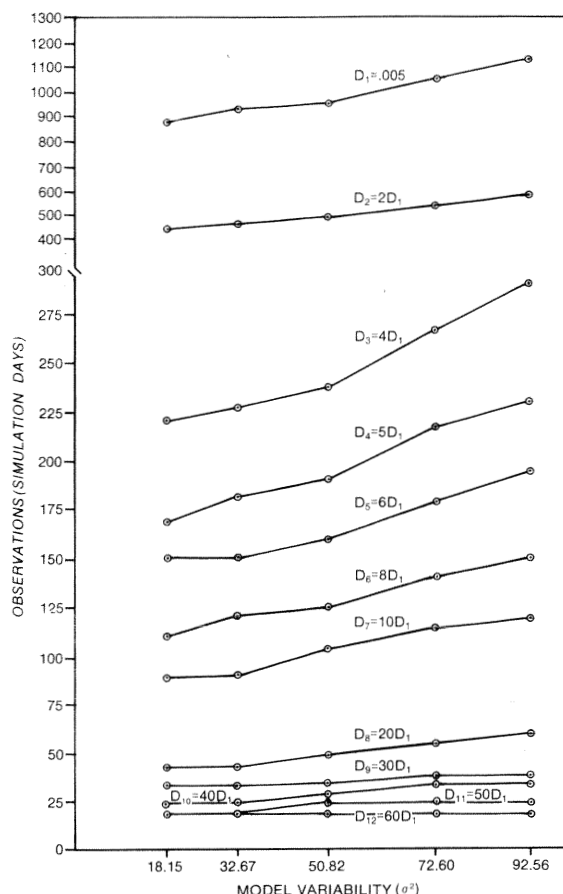
large D value resulting in a small amount of computer time, (20 observations for our example: one observation corresponds to one simulated day) and a large γ and a small D value resulting in a great amount of computer time (1,255 observations for our example).

Figure 4 shows the length of computer runs for 5 levels of (σ^2) and 12 D values. In our example the random variables were generated from normal distributions and variability was increased by increasing the variability of two of the components. This example is just one of many ways to change the model variability. For a given D level, as σ^2 of the model increases, the time needed to obtain a reliable parameter estimate increases. When the variance is held constant and D decreased, the computer time increases. For this example, when σ^2 was small and D was large, the number of observations needed to obtain a reliable parameter estimate was small (20 simulation days); however, when σ^2 was large and D was small, the required number of observations was large (1,125 simulation days).

Thus the three factors γ , D , and σ^2 must be carefully considered when analyzing the simulation experiment results.

The analyst has the responsibility of insuring reliable results from simulation experiments. The potential savings from use of the model should not be diminished by excessive costs in obtaining the results.

Figure 4. — Number of simulated days versus model variability (σ^2) for various levels of D .



REFERENCES

Chow, Y. S., and H. Robbins.

1965. ON THE ASYMPTOTIC THEORY OF FIXED-WIDTH SEQUENTIAL CONFIDENCE INTERVALS FOR THE MEAN. *Ann. Math. Statis.* 36(2):457-462.

Emshoff, J. R., and R. L. Sisson.

1970. DESIGN AND USE OF COMPUTER SIMULATION MODELS. The Macmillan Co., New York. p. 189-208.

Reese, P. A.

1968. STEADY STATE PARAMETER ESTIMATION IN COMPUTER SIMULATED SYSTEMS. M. S. thesis, State Univ. N. Y., Buffalo. 127 p.

Stein, C.

1945. A TWO SAMPLE TEST FOR A LINEAR HYPOTHESIS WHOSE POWER IS INDEPENDENT OF THE VARIANCE. *Ann. Math. Statis.* 16:243-250.