

INTRODUCTION TO SIMULATION

Jerry Banks

AutoSimulations, Inc.
Atlanta, GA 30067, U.S.A.

ABSTRACT

This introduction begins with an example of simulation done by hand. Modeling concepts in simulation are then introduced based on the example. Next, the advantages and disadvantages of simulation are discussed. The introduction ends with a discussion of the steps in a simulation study.

1 DEFINITION OF SIMULATION

Simulation is the imitation of the operation of a real-world process or system over time. Simulation involves the generation of an artificial history of the system, and the observation of that artificial history to draw inferences concerning the operating characteristics of the real system that is represented.

Simulation is an indispensable problem-solving methodology for the solution of many real-world problems. Simulation is used to describe and analyze the behavior of a system, ask "what if" questions about the real system, and aid in the design of real systems. Both existing and conceptual systems can be modeled with simulation.

2 SIMULATION EXAMPLE

Consider the operation of a one-teller bank where customers arrive for service between one and ten minutes apart in time, integer values only, each value equally likely. The customers are served in a time between one and six minutes, also integer valued, and equally likely. Restricting the times to integer values is an abstraction of reality, since time is continuous, but this aids in presenting the example. The objective is to simulate the bank operation, by hand, until twenty customers are served, and to compute measures of performance such as the percentage of idle time, the average waiting time per customer, and so on. Admittedly, twenty customers is far too few to draw conclusions about the operation of the system for the long run, but by following this example, the stage is set for further presentations in the introductory tutorials at WSC.

To simulate the process, random interarrival and service times need to be generated. Assume that the interarrival times are generated using a spinner that has possibilities for the values 1 through 10. Further assume that the service times are generated using a die that has possibilities for the values 1 through 6.

Table 1 is called an ad hoc simulation table. The setup of the simulation table is for the purpose of this problem, but does not pertain to all problems. Column 1, Customer, lists the 20 customers that arrive to the system. It is assumed that Customer 1 arrives at time zero, thus a dash is indicated in Row 1 of Column 2, Time between Arrivals. Rows 2 through 20 of Column 2 were generated using the spinner. Column 3, Arrival Time, shows the simulated arrival times. Since Customer 1 is assumed to arrive at time 0, and there is a 5 minute interarrival time, Customer 2 arrives at time 5. There is a 1 minute interarrival time for Customer 3, thus, the arrival occurs at time 6. This process of adding the interarrival time to the previous arrival time is called bootstrapping. By continuing this process, the arrival times of all 20 customers are determined. Column 4, Service Time, contains the simulated service times for all 20 customers. These were generated by rolling the die.

Now, the simulation of the service process begins. At time 0, Customer 1 arrived, and immediately began service. The service time was 2 minutes, so the service period ended at time 2. The total time in the system for Customer 1 was 2 minutes. The bank teller was not idle since the simulation began with the arrival of a customer. The customer did not have to wait for the teller.

At time 5, Customer 2 arrived, and immediately began service as shown in Column 6. The service time was 2 minutes so the service period ended at time 7 as shown in Column 6. The bank teller was idle from time 2 until time 5, so 3 minutes of idle time occurred. Customer 2 spent no time in the queue.

Customer 3 arrived at time 6, but service could not begin until time 7 as Customer 2 was being served until time 7. The service time was 6 minutes, so service was completed at time 13. Customer 3 was in the system from time 6 until time 13, or for 7 minutes as indicated in Column 7, Time in System. Although there was no idle

Table 1: Ad Hoc Simulation

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Mechanic	Time between Arrivals	Arrival Time	Service Time	Service Begins	Time Service Ends	Time in System	Idle Time	Time in Queue
1	-	0	2	0	2	2	0	0
2	5	5	2	5	7	2	3	0
3	1	6	6	7	13	7	0	1
.								
.								
18	4	84	5	84	89	5	0	0
19	7	91	3	91	94	3	2	0
20	7	98	1	98	99	$\frac{1}{72}$	$\frac{4}{34}$	$\frac{0}{7}$

time, Customer 3 had to wait in the queue for 1 minute for service to begin.

This process continues for all 20 customers, and the totals shown in Columns 7, 8 (Idle Time), and 9 (Time in Queue) are entered. Some performance measures can now be calculated as follows:

$$\text{Average time in system} = 79/20 = 3.95 \text{ minutes}$$

$$\% \text{ idle time} = [30/99](100) = 30\%$$

$$\text{Average waiting time per customer} = 10/20 = 0.5 \text{ minutes}$$

$$\text{Fraction having to wait} = 5/20 = 0.25$$

$$\text{Average waiting time of those that waited} = 10/3 = 3.33 \text{ minutes.}$$

This very limited simulation indicates that the system is functioning well. Only 25% of the customers had to wait. About 30% of the time the teller is idle. Whether a slower teller should replace the current teller depends on the cost of having to wait versus any savings from having a slower server.

This small simulation can be accomplished by hand, but there is a limit to the complexity of problems that can be solved in this manner. Also, the number of customers that must be simulated could be much larger than 20 and the number of times that the simulation must be run for statistical purposes could be large. Hence, using the computer to solve real simulation problems is almost always appropriate.

This example raises some issues that will be addressed in the introductory tutorials at WSC. The issues include the following:

1. How is the form of the input data determined?

2. How are random variates generated if they follow statistical distributions other than the discrete uniform?
3. How does the user know that the simulation imitates reality?
4. What other kinds of problems can be solved by simulation?
5. How long does the simulation need to run?
6. How many different simulation runs should be conducted?
7. What statistical techniques should be used to analyze the output?

3 MODELING CONCEPTS

There are several concepts underlying simulation. These include system and model, events, system state variables, entities and attributes, list processing, activities and delays, and finally the definition of discrete-event simulation. Additional information on these topics is available from Banks, Carson and Nelson (1996) and Law and Kelton (1991). The discussion in this section follows that of Carson (1993).

System, Model and Events: A model is a representation of an actual system. Immediately, there is a concern about the limits or boundaries of the model that supposedly represent the system. The model should be complex enough to answer the questions raised, but not too complex.

Consider an event as an occurrence that changes the state of the system. In the example, events include the arrival of a customer for service at the bank, the beginning of service for a customer, and the completion of a service.

There are both internal and external events, also called endogenous and exogenous events, respectively. For example, an endogenous event in the example is the beginning of service of the customer since that is within the system being simulated. An exogenous event is the arrival of a customer for service since that occurrence is outside of the simulation. However, the arrival of a

customer for service impinges on the system, and must be taken into consideration.

WSC concerns discrete-event simulation models. These are contrasted with other types of models such as mathematical models, descriptive models, statistical models, and input-output models. A discrete-event model attempts to represent the components of a system and their interactions to such an extent that the objectives of the study are met. Most mathematical, statistical, and input-output models represent a system's inputs and outputs explicitly, but represent the internals of the model with mathematical or statistical relationships. An example is the mathematical model from physics,

$$\text{Force} = \text{Mass} \times \text{Acceleration}$$

based on theory. Discrete-event simulation models include a detailed representation of the actual internals.

Discrete-event models are dynamic, i.e., the passage of time plays a crucial role. Most mathematical and statistical models are static in that they represent a system at a fixed point in time. Consider the annual budget of a firm. This budget resides in a spreadsheet. Changes can be made in the budget and the spreadsheet can be recalculated, but the passage of time is usually not a critical issue. Further comments will be made about discrete-event models after several additional concepts are presented.

System State Variables: The system state variables are the collection of all information needed to define what is happening within the system to a sufficient level (i.e., to attain the desired output) at a given point in time. The determination of system state variables is a function of the purposes of the investigation, so what may be the system state variables in one case may not be the same in another case even though the physical system is the same. Determining the system state variables is as much an art as a science. However, during the modeling process, any omissions will readily come to light. (And, on the other hand, unnecessary state variables may be eliminated.)

Having defined system state variables, a contrast can be made between discrete-event models and continuous models based on the variables needed to track the system state. The system state variables in a discrete-event model remain constant over intervals of time and change value only at certain well-defined points called event times. Continuous models have system state variables defined by differential or difference equations giving rise to variables that may change continuously over time.

Some models are mixed discrete-event and continuous. There are also continuous models that are treated as discrete-event models after some reinterpretation of system state variables, and vice versa.

Entities and Attributes. An entity represents an object that requires explicit definition. An entity can be dynamic in that it "moves" through the system, or it can be static in that it serves other entities. In the example, the

customer is a dynamic entity, whereas the bank teller is a static entity.

An entity may have attributes that pertain to that entity alone. Thus, attributes should be considered as local values. In the example, an attribute of the entity could be the time of arrival. Attributes of interest in one investigation may not be of interest in another investigation. Thus, if red parts and blue parts are being manufactured, the color could be an attribute. However, if the time in the system for all parts is of concern, the attribute of color may not be of importance. From this example, it can be seen that many entities can have the same attribute or attributes (i.e., more than one part may have the attribute "red").

Resources: A resource is an entity that provides service to dynamic entities. The resource can serve one or more than one dynamic entity at the same time, i.e., operate as a parallel server. A dynamic entity can request one or more units of a resource. If denied, the requesting entity joins a queue, or takes some other action (i.e., diverted to another resource, ejected from the system). (Other terms for queues include files, chains, buffers, and waiting lines.) If permitted to capture the resource, the entity remains for a time, then releases the resource.

There are many possible states of the resource. Minimally, these states are idle and busy. But other possibilities exist including failed, blocked, or starved.

List Processing: Entities are managed by allocating them to resources that provide service, by attaching them to event notices thereby suspending their activity into the future, or by placing them into an ordered list. Lists are used to represent queues.

Lists are often processed according to FIFO (first-in-first-out), but there are many other possibilities. For example, the list could be processed by LIFO (last-in-first-out), according to the value of an attribute, or randomly, to mention a few. An example where the value of an attribute may be important is in SPT (shortest process time) scheduling. In this case, the processing time may be stored as an attribute of each entity. The entities are ordered according to the value of that attribute with the lowest value at the head or front of the queue.

Activities and Delays. An activity is a duration of time whose duration is known prior to commencement of the activity. Thus, when the duration begins, its end can be scheduled. The duration can be a constant, a random value from a statistical distribution, the result of an equation, input from a file, or computed based on the event state. For example, a service time may be a constant 10 minutes for each entity; it may be a random value from an exponential distribution with a mean of 10 minutes; it could be 0.9 times a constant value from clock time 0 to clock time 4 hours, and 1.1 times the standard value after clock time 4 hours; or it could be 10 minutes when the preceding queue contains at most four entities and 8

minutes when there are five or more in the preceding queue.

A delay is an indefinite duration that is caused by some combination of system conditions. When an entity joins a queue for a resource, the time that it will remain in the queue may be unknown initially since that time may depend on other events that may occur. An example of another event would be the arrival of a rush order that preempts the resource. When the preempt occurs, the entity using the resource relinquishes its control instantaneously. Another example is a failure necessitating repair of the resource.

Discrete-event simulations contain activities that cause time to advance. Most discrete-event simulations also contain delays as entities wait. The beginning and ending of an activity or delay is an event.

Discrete-Event Simulation Model: Sufficient modeling concepts have been defined so that a discrete-event simulation model can be defined as one in which the state variables change only at those discrete points in time at which events occur. Events occur as a consequence of activity times and delays. Entities may compete for system resources, possibly joining queues while waiting for an available resource. Activity and delay times may "hold" entities for durations of time.

A discrete-event simulation model is conducted over time ("run") by a mechanism that moves simulated time forward. The system state is updated at each event along with capturing and freeing of resources that may occur at that time.

4 ADVANTAGES AND DISADVANTAGES OF SIMULATION

Competition in the computer industry has led to technological breakthroughs that are allowing hardware companies to continually produce better products. It seems that every week another company announces its latest release, each with more options, memory, graphics capability, and power. (Much of this section is from Banks and Norman 1995.)

What is unique about new developments in the computer industry is that they often act as a springboard for other related industries to follow. One industry in particular is the simulation-software industry. As computer hardware becomes more powerful, more accurate, faster, and easier to use, simulation software does too.

The number of businesses using simulation is rapidly increasing. Many managers are realizing the benefits of utilizing simulation for more than just the one-time remodeling of a facility. Rather, due to advances in software, managers are incorporating simulation in their daily operations on an increasingly regular basis.

For most companies, the benefits of using simulation go beyond just providing a look into the future. These benefits are mentioned by many authors (Banks, Carson and Nelson 1996; Law and Kelton 1991; Pegden, Shannon and Sadowski 1995; and Schriber 1991) and are included in the following:

Choose Correctly. Simulation lets you test every aspect of a proposed change or addition without committing resources to their acquisition. This is critical, because once the hard decisions have been made, the bricks have been laid, or the material-handling systems have been installed, changes and corrections can be extremely expensive. Simulation allows you to test your designs without committing resources to acquisition.

Time Compression and Expansion. By compressing or expanding time simulation allows you to speed up or slow down phenomena so that you can thoroughly investigate them. You can examine an entire shift in a matter of minutes if you desire, or you can spend two hours examining all the events that occurred during one minute of simulated activity.

Understand "Why?" Managers often want to know why certain phenomena occur in a real system. With simulation, you determine the answer to the "why" questions by reconstructing the scene and taking a microscopic examination of the system to determine why the phenomenon occurs. You cannot accomplish this with a real system because you cannot see or control it in its entirety.

Explore Possibilities. One of the greatest advantages of using simulation software is that once you have developed a valid simulation model, you can explore new policies, operating procedures, or methods without the expense and disruption of experimenting with the real system. Modifications are incorporated in the model, and you observe the effects of those changes on the computer rather than the real system.

Diagnose Problems. The modern factory floor or service organization is very complex--so complex that it is impossible to consider all the interactions taking place in one given moment. Simulation allows you to better understand the interactions among the variables that make up such complex systems. Diagnosing problems and gaining insight into the importance of these variables increases your understanding of their important effects on the performance of the overall system.

The last three claims can be made for virtually all modeling activities, queueing, linear programming, etc. However, with simulation the models can become very complex and, thus, have a higher fidelity, i.e., they are valid representations of reality.

Identify Constraints. Production bottlenecks give manufacturers headaches. It is easy to forget that bottlenecks are an effect rather than a cause. However, by using simulation to perform bottleneck analysis, you can

discover the cause of the delays in work-in-process, information, materials, or other processes.

Develop Understanding. Many people operate with the philosophy that talking loudly, using computerized layouts, and writing complex reports convinces others that a manufacturing or service system design is valid. In many cases these designs are based on someone's thoughts about the way the system operates rather than on analysis. Simulation studies aid in providing understanding about how a system really operates rather than indicating an individual's predictions about how a system will operate.

Visualize the Plan. Taking your designs beyond CAD drawings by using the animation features offered by many simulation packages allows you to see your facility or organization actually running. Depending on the software used, you may be able to view your operations from various angles and levels of magnification, even 3-D. This allows you to detect design flaws that appear credible when seen just on paper on in a 2-D CAD drawing.

Build Consensus. Using simulation to present design changes creates an objective opinion. You avoid having inferences made when you approve or disapprove of designs because you simply select the designs and modifications that provided the most desirable results, whether it be increasing production or reducing the waiting time for service. In addition, it is much easier to accept reliable simulation results, which have been modeled, tested, validated, and visually represented, instead of one person's opinion of the results that will occur from a proposed design.

Prepare for Change. We all know that the future will bring change. Answering all of the "what-if" questions is useful for both designing new systems and redesigning existing systems. Interacting with all those involved in a project during the problem-formulation stage gives you an idea of the scenarios that are of interest. Then you construct the model so that it answers questions pertaining to those scenarios. What if an AGV is removed from service for an extended period of time? What if demand for service increases by 10 percent? What if...? The options are unlimited.

Wise Investment. The typical cost of a simulation study is substantially less than 1% of the total amount being expended for the implementation of a design or redesign. Since the cost of a change or modification to a system after installation is so great, simulation is a wise investment.

Train the Team. Simulation models can provide excellent training when designed for that purpose. Used in this manner, the team provides decision inputs to the simulation model as it progresses. The team, and individual members of the team, can learn by their mistakes, and learn to operate better. This is much less expensive and less disruptive than on-the-job learning.

Specify Requirements. Simulation can be used to specify requirements for a system design. For example, the specifications for a particular type of machine in a

complex system to achieve a desired goal may be unknown. By simulating different capabilities for the machine, the requirements can be established.

The disadvantages of simulation include the following:

Model Building Requires Special Training. It is an art that is learned over time and through experience. Furthermore, if two models of the same system are constructed by two competent individuals, they may have similarities, but it is highly unlikely that they will be the same.

Simulation Results May Be Difficult to Interpret. Since most simulation outputs are essentially random variables (they are usually based on random inputs), it may be hard to determine whether an observation is a result of system interrelationships or randomness.

Simulation Modeling and Analysis Can Be Time Consuming and Expensive. Skimping on resources for modeling and analysis may result in a simulation model and/or analysis that is not sufficient to the task.

Simulation May Be Used Inappropriately. Simulation is used in some cases when an analytical solution is possible, or even preferable. This is particularly true in the simulation of some waiting lines where closed-form queueing models are available, at least for long-run evaluation.

In defense of simulation, these four disadvantages, respectively, can be offset as follows:

Simulators. Vendors of simulation software have been actively developing packages that contain models that only need input data for their operation. Such models have the generic tag "simulators" or templates.

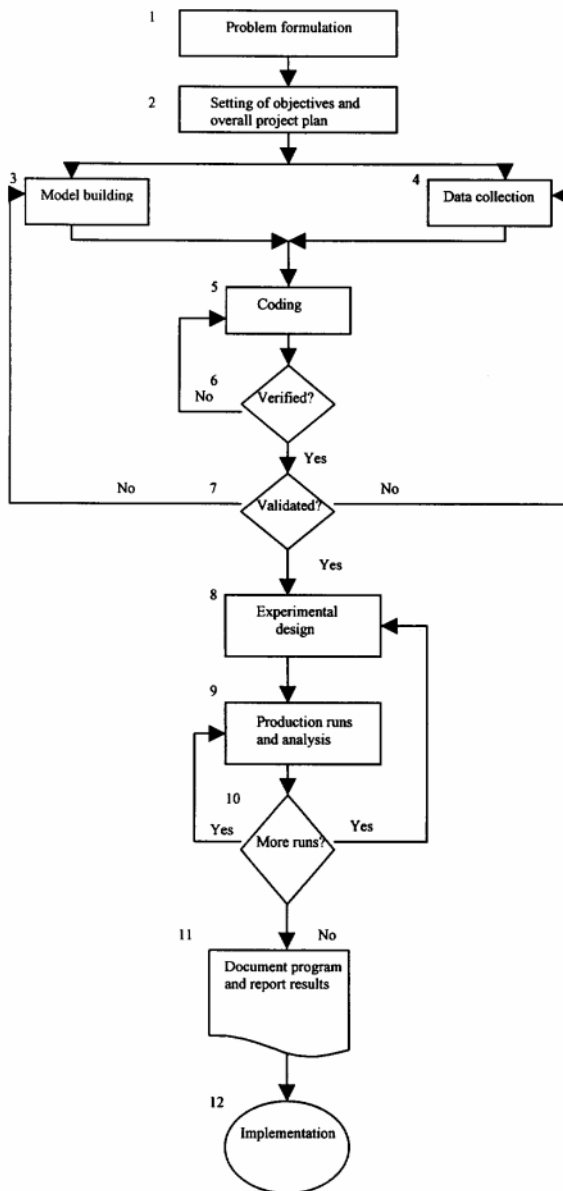
Output Analysis. Most simulation-software vendors have developed output-analysis capabilities within their packages for performing very extensive analysis. This reduces the computational requirements on the part of the user, although they still must understand the analysis procedure.

Faster and Faster. Simulation can be performed faster today than yesterday, and even faster tomorrow. This is attributable to the advances in hardware that permit rapid running of scenarios. It is also attributable to the advances in many simulation packages. For example, many simulation software products contain constructs for modeling material handling using transporters such as conveyors, and automated guided vehicles.

Limitations of Closed-Form Models. Closed-form models are not able to analyze most of the complex systems that are encountered in practice. In nearly eight years of consulting practice, not one problem was encountered that could have been solved by a closed-form solution.

5 STEPS IN A SIMULATION STUDY

Figure 1 shows a set of steps to guide a model builder in a thorough and sound simulation study. Similar figures and



[Banks, et al., DISCRETE EVENT SYSTEM SIMULATION, 2/e, ©1996, p.14. Reprinted by permission of Prentice-Hall, Upper Saddle River, New Jersey.]

Figure 1: Steps in a Simulation Study

their interpretation can be found in other sources such as Pegden, Shannon, and Sadowski (1995) and Law and Kelton (1991). This presentation is built on that of Banks, Carson and Nelson (1996).

1. Problem formulation. Every simulation study begins with a statement of the problem. If the statement is provided by those that have the problem (client), the simulation analyst must take extreme care to insure that the problem is clearly understood. If a problem statement is

prepared by the simulation analyst, it is important that the client understand and agree with the formulation. It is suggested that a set of assumptions be prepared by the simulation analyst and agreed to by the client. Even with all of these precautions, it is possible that the problem will need to be reformulated as the simulation study progresses.

2. Setting of objectives and overall project plan. Another way to state this step is "prepare a proposal." This step should be accomplished regardless of location of the analyst and client, viz., as an external or internal consultant.

The objectives indicate the questions that are to be answered by the simulation study. The project plan should include a statement of the various scenarios that will be investigated. The plans for the study should be indicated in terms of time that will be required, personnel that will be used, hardware and software requirements if the client wants to run the model and conduct the analysis, stages in the investigation, output at each stage, cost of the study and billing procedures, if any.

3. Model conceptualization. The real-world system under investigation is abstracted by a conceptual model, a series of mathematical and logical relationships concerning the components and the structure of the system. It is recommended that modeling begin simply and that the model grow until a model of appropriate complexity has been developed. For example, consider the model of a manufacturing and material handling system. The basic model with the arrivals, queues and servers is constructed. Then, add the failures and shift schedules. Next, add the material-handling capabilities. Finally, add the special features. Constructing an unduly complex model will add to the cost of the study and the time for its completion without increasing the quality of the output. Maintaining client involvement will enhance the quality of the resulting model and increase the client's confidence in its use.

4. Data collection. Shortly after the proposal is "accepted" a schedule of data requirements should be submitted to the client. In the best of circumstances, the client has been collecting the kind of data needed in the format required, and can submit these data to the simulation analyst in electronic format. Oftentimes, the client indicates that the required data are indeed available. However, when the data are delivered they are found to be quite different than anticipated. For example, in the simulation of an airline-reservation system, the simulation analyst was told "we have every bit of data that you want over the last five years." When the study commenced, the data delivered were the average "talk time" of the reservationist for each of the years. Individual values were needed, not summary measures.

Model building and data collection are shown as contemporaneous in Figure 1. This is to indicate that the simulation analyst can readily construct the model while the data collection is progressing.

5. Model translation. The conceptual model constructed in Step 3 is coded into a computer recognizable form, an operational model.

6. Verified? Verification concerns the operational model. Is it performing properly? Even with small textbook sized models, it is quite possible that they have verification difficulties. These models are orders of magnitude smaller than real models (say 50 lines of computer code versus 2,000 lines of computer code). It is highly advisable that verification take place as a continuing process. It is ill advised for the simulation analyst to wait until the entire model is complete to begin the verification process. Also, use of an interactive run controller, or debugger, is highly encouraged as an aid to the verification process.

7. Validated? Validation is the determination that the conceptual model is an accurate representation of the real system. Can the model be substituted for the real system for the purposes of experimentation? If there is an existing system, call it the base system, then an ideal way to validate the model is to compare its output to that of the base system. Unfortunately, there is not always a base system. There are many methods for performing validation.

8. Experimental design. For each scenario that is to be simulated, decisions need to be made concerning the length of the simulation run, the number of runs (also called replications), and the manner of initialization, as required.

9. Production runs and analysis. Production runs, and their subsequent analysis, are used to estimate measures of performance for the scenarios that are being simulated.

10. More runs? Based on the analysis of runs that have been completed, the simulation analyst determines if additional runs are needed and if any additional scenarios need to be simulated.

11. Documentation and reporting. Documentation is necessary for numerous reasons. If the simulation model is going to be used again by the same or different analysts, it may be necessary to understand how the simulation model operates. This will enable confidence in the simulation model so that the client can make decisions based on the analysis. Also, if the model is to be modified, this can be greatly facilitated by adequate documentation.

The result of all the analysis should be reported clearly and concisely. This will enable the client to review the final formulation, the alternatives that were addressed, the criterion by which the alternative systems were compared, the results of the experiments, and analyst recommendations, if any.

12. Implementation. The simulation analyst acts as a reporter rather than an advocate. The report prepared in step 11 stands on its merits, and is just additional information that the client uses to make a decision. If the client has been involved throughout the study period, and the simulation analyst has followed all of the steps

rigorously, then the likelihood of a successful implementation is increased.

REFERENCES

- Banks, J., B. Burnette, H. Kozloski, J. D. Rose, 1995. *Introduction to SIMAN V and CINEMA V*, New York: John Wiley.
- Banks, J., J.S. Carson, B.L. Nelson, 1996. *Discrete Event System Simulation*, 2nd Ed., Upper Saddle River, New Jersey: Prentice-Hall.
- Banks, J. and V. Norman, November, 1995. "Justifying Simulation in Today's Manufacturing Environment," *IIE Solutions*.
- Carson, J.S., 1993. "Modeling and Simulation World Views," in *Proceedings of the 1993 Winter Simulation Conference*, eds., G.W. Evans, M. Mollaghasemi, E.C. Russell, and W.E. Biles, Institute of Electrical and Electronics Engineers, pp. 18-23. 1-4, Piscataway, NJ.
- Law, A.M. and W.D. Kelton, 1991. *Simulation Modeling and Analysis*, 2nd Ed., New York: McGraw-Hill.
- Pegden, C.D., R.E. Shannon and R.P. Sadowski, 1995. *Introduction to Simulation Using SIMAN*, 2nd Ed., New York: McGraw-Hill.
- Pidd, M., 1992. *Computer Modelling for Discrete Simulation*, Chichester, England: John Wiley.
- Schriber, T.J. 1991. *An Introduction to Simulation Using GPSS/H*, New York: John Wiley.

AUTHOR BIOGRAPHY

JERRY BANKS is Director of Simulation Technology, AutoSimulations, Inc. He was formerly Professor of Industrial and Systems Engineering at Georgia Tech.