

AD-A078 535

WISCONSIN UNIV-MADISON DEPT OF INDUSTRIAL ENGINEERING

F/6 9/2

VALIDATION OF SIMULATION MODELS, I. AN OVERVIEW AND SURVEY OF R--ETC(U)

AUG 79 A M LAW

N00014-76-C-0403

UNCLASSIFIED

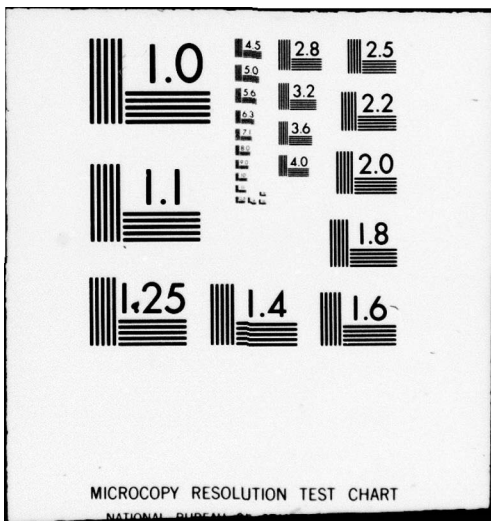
TR-78-14

NL

| OF |
ADA
078535



END
DATE
FILMED
-80
DDC



9/10/79

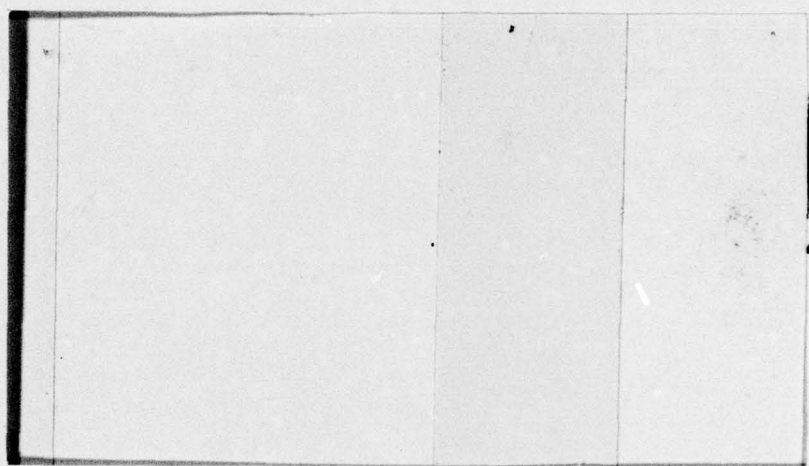
UH.

Handwritten initials and a large circled '1'.

DEPARTMENT OF INDUSTRIAL ENGINEERING

LEVEL

ADA 078535



DDC
RECEIVED
DEC 17 1979
E

DDC FILE COPY

This document has been approved for public release and sale; its distribution is unlimited.



UNIVERSITY OF WISCONSIN-MADISON

79 12 14 091

P

DDC
RECEIVED
DEC 17 1979
E

⑨ Technical Rept.

⑥

Validation of Simulation Models, I.

An Overview and Survey
of Real-World Practice

by

⑩ Averill M. Law

⑪

August 1979

⑭

TR-78-14

⑫

25

⑮ N00014-76-C-0403,
DAAG-29-75-C-0024

Accession For	
NTIS GRA&I	<input checked="" type="checkbox"/>
DDC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or special
A	

This document has been approved
for public release and sale; its
distribution is unlimited.

408848

4B

Validation of Simulation Models, I:
An Overview and Survey of Real-World Practice[†]

Averill M. Law
University of Wisconsin

ABSTRACT

We consider the problem of assessing the degree to which a model constructed for use in a simulation study faithfully represents the corresponding real system. ^{IS CONSIDERED} ~~Our~~ ^{THE} focus here is to survey general approaches and methods which have been used in practice, or could be implemented, including good programming techniques, verifying that the program itself is correct, overall goals of the validation of a model, and a general framework for the total validation effort. Several concrete examples are discussed to illustrate the methods as well as possible pitfalls.

[†]This research was supported by the Office of Naval Research under contract N00014-76-C-0403 (NR 047-145) and the Army Research Office under contract DAAG29-75-C-0024.

1. Introduction

One of the most important problems facing a real-world simulator is that of trying to determine whether a simulation model is an accurate representation of the actual system which is being studied. Yet a review of the validation literature indicates that relatively little has been written on this subject. Furthermore, what has been written is often philosophical in nature, rather than being in the form of practical recommendations. (Important works on validation include [6], [13], [14], [15], and [17].) Somewhat surprised by this state of affairs, we decided to engage in a two-phase study to develop definitive qualitative and statistical procedures which actually can be used by a simulator in his validation efforts. In this paper, we present an overview of the entire field of validation and survey techniques that can be used for verifying and validating a simulation model. (See below for the distinction between these terms.) Information for this survey came not only from existing papers and books on validation, but also from conversations with notable members of the academic and industrial communities who have had firsthand experience with validation. It was hoped that we might uncover some validation techniques which have not been previously documented. The second phase of our study will seek to develop statistical procedures which can be used for comparing the output data from a simulation model and a corresponding real-world system (if the system exists), and will be reported by Law [10].

Since there appears to be some confusion in the simulation literature as to the meaning of verification, validation, and output analysis, we begin by giving simple definitions of these terms.

Verification is determining whether a simulation model performs as intended, i.e., "debugging" the computer program. Although verification is simple in concept, the debugging of a large-scale simulation model can be quite an arduous task. *Validation* is determining whether a simulation model (as opposed to the computer program) is an accurate representation of the real-world system under study. This is to be contrasted with *output analysis* which is concerned with determining a simulation model's (not necessarily the system's) true population parameters or characteristics. (For surveys of output analysis, see Law [9] and Law and Kelton [11,12].)

The remainder of this paper is organized as follows. In Section 2 we describe practical techniques for debugging the computer program of a simulation model. Some general thoughts on the entire validation effort are offered in Section 3, followed in Section 4 by a framework in which to carry out validation in practice. Several additional considerations in validation, such as calibration, are discussed in Section 5.

2. Verification of Simulation Models

In this section we discuss five techniques which can be used for debugging the computer program of a simulation model. Some of these techniques might be used for debugging any computer program, while others we believe to be unique to simulation modeling.

(1) In developing a simulation model, write and debug the computer program in modules or subprograms. For example, if one were to develop a 10000 statement simulation model, we feel that it would be poor programming practice to write the entire program before attempting any debugging. When this large, untested program is finally run, it almost certainly will not execute and, furthermore, determining

the location in the program of the errors will be an extremely difficult task. Instead, the simulation model's main program and a few of the key subprograms should first be written and debugged, perhaps representing the other required subprograms as "dummies" or "stubs." Then, additional subprograms or levels of detail are successively added and debugged until a model is developed which satisfactorily represents the system under study. In general, we believe that it is always better to start with a simple model which is gradually made as complex as needed, than to develop "immediately" a complex model which may turn out to be more detailed than necessary and excessively expensive to run (see Subsection 4.B for further discussion).

(2) We believe that it is advisable when developing large simulation models to have more than one person read the computer program, since the person who writes a particular subprogram may get into a "mental rut" and thus not be a good evaluator of its correctness. In some organizations, this idea is implemented in a formal manner and is called a structured walk-through. For example, all members of the modeling team (e.g., systems analysts, programmers, etc.) are assembled in a room each having a copy of a particular set of subprograms which are to be debugged. Then the subprograms' developer goes through the computer code but does not proceed from one statement to another until everyone is convinced that a statement is correct.

(3) One of the most powerful techniques that can be used to debug a discrete event simulation model is a "trace." In a trace, the state of the simulated system (i.e., the contents of the event list, the state variables, certain statistical counters, etc.) is

printed out just after each event occurs in order to see if the program is performing as intended. In performing a trace, it is desirable to evaluate each possible program path and also the program's ability to deal with "extreme" conditions. Sometimes in order to effect such a thorough evaluation, it may be necessary to prepare special (perhaps deterministic) input data for the model. It should be mentioned that all three of the major simulation languages in the United States (GASP, GPSS, and SIMSCRIPT) explicitly provide the capability to perform traces.

(4) In order to determine whether a simulation model is performing as intended, the model should, when possible, be run under simplifying assumptions for which the model's true characteristics are known or can be easily computed. To illustrate this idea, let us consider a detailed example. A manufacturing shop consists of five groups of machines with groups 1,2,...,5 consisting of 3,2,4, 3,1 identical machines, respectively. (However, machines in different groups are not the same.) Assume that jobs arrive to the shop with interarrival times that are independent identically distributed (i.i.d.) exponential random variables (r.v.'s). There are three types of jobs and each job type occurs with a specified probability. Each job type requires a certain number of tasks to be done and each task must be done at a specified machine group and in a prescribed order. For example, job type 1 requires four tasks to be done successively at machine groups 3, 1, 2, 5 (see Figure 1). If a job arrives at a particular machine group and finds all machines in that group already busy, then the job joins a single first-in, first-out queue at that machine group. The time to perform a task at a

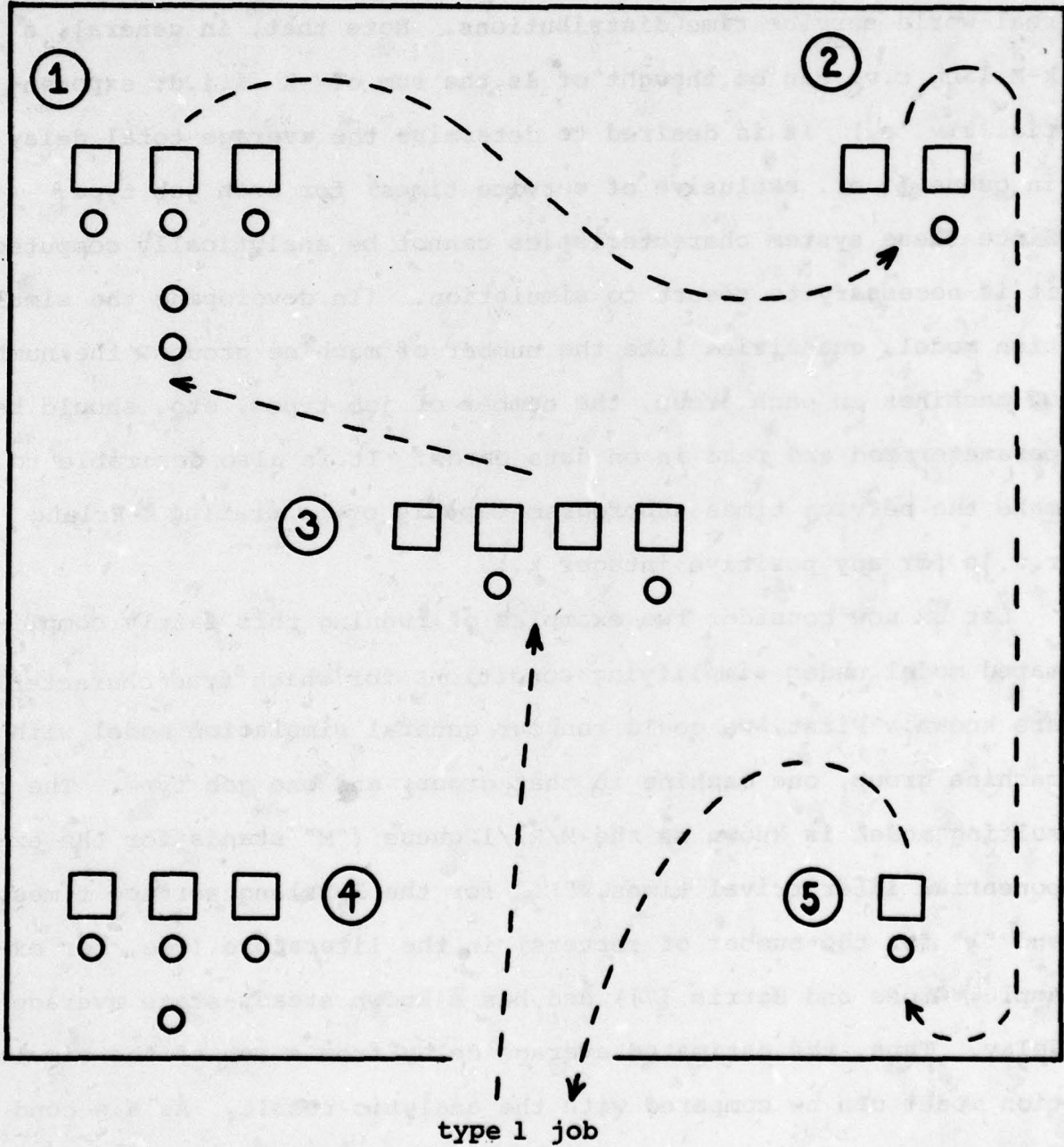


Figure 1. Manufacturing shop with five machine groups, showing the route of type 1 jobs.

particular machine is an independent 2- Erlang r.v. whose mean depends on the job type and on the machine group. (Our experience indicates that a 2-Erlang distribution is representative of many real-world service time distributions. Note that, in general, a k-Erlang r.v. can be thought of as the sum of k i.i.d. exponential r.v.'s.) It is desired to determine the average total delay in queue (i.e., exclusive of service times) for each job type. Since these system characteristics cannot be analytically computed, it is necessary to resort to simulation. (In developing the simulation model, quantities like the number of machine groups, the number of machines in each group, the number of job types, etc. should be parameterized and read in on data cards. It is also desirable to make the service times subprogram capable of generating k-Erlang r.v.'s for any positive integer k.)

Let us now consider two examples of running this fairly complicated model under simplifying conditions for which true characteristics are known. First, we could run our general simulation model with one machine group, one machine in that group, and one job type. The resulting model is known as the $M/E_2/1$ queue ("M" stands for the exponential interarrival times, " E_2 " for the 2-Erlang service times, and "1" for the number of servers) in the literature (see, for example, Gross and Harris [7]) and has a known steady-state average delay. Thus, the estimated average delay from a run of the simulation model can be compared with the analytic result. As a second example, one could run the model with the desired number of machine groups and number of machines in the groups, but with only type 1 jobs and with exponential service times (with the same mean as the

corresponding 2-Erlang service time) at each machine group. The resulting model is, in effect, four multi-server queues in series with the first queue's being an M/M/3, the second an M/M/2, etc. (The interdeparture times from an M/M/s queue (s the number of servers), which has been in operation for a long period of time, are i.i.d. exponential r.v.'s.) For this model, one can analytically compute the steady-state average total delay of a (type 1) job (see [7]) and compare this result to the simulation estimate.

(5) With some types of simulation models, it may be helpful to display the simulation output on a graphics terminal as the simulation actually progresses. Let us illustrate this idea by means of a real-life example. A simulation model of a network of automobile traffic intersections was developed, supposedly debugged, and used for some period of time to study such issues as the effect of various light sequencing policies. However, when the simulated flow of traffic was displayed on a graphics terminal, it was found that simulated cars were actually colliding in the intersections; subsequent inspection of the computer program revealed several errors which had not been previously detected. (The author would like to thank Professor Robert Sargent for this example.)

3. General Perspectives on Validation

We now describe six general perspectives on validation. These should not be thought of as definitive recommendations on how to validate a simulation model, but rather as somewhat philosophical considerations which should be kept in mind when contemplating how to validate a model of a real-world system.

(1) Experimentation with a simulation model is a surrogate for actually being able to experiment with an existing or proposed system. Thus, a reasonable goal in validation is to ensure that a model is developed which can actually be used by a decision-maker to make the same decision that would be made if it were feasible and cost-effective to experiment with the actual system. Although this statement is hard to disagree with in theory, knowing how to effect it in practice is a different story. We hope to shed some light on this matter in the next section where we discuss a three-step approach to validation.

(2) A simulation model of a complex, real-world system is always only an approximation to the actual system regardless of how much effort is put into the model. Thus, one should not speak of the absolute validity or invalidity of a model, but rather of the degree to which the model agrees with the system. The more time (and hence money) that is spent on validation, the closer will be the agreement of the model with the system. However, the most "valid" model will not necessarily be the most cost-effective. One should always keep in mind the overall objective of the simulation study, which is often to save money by determining an efficient system design.

(3) A simulation model should always be developed for a particular purpose. Indeed, a model valid for one purpose may not be for another. (Since simulation models often evolve over time and are used for different purposes, we believe that every simulation study should include thorough documentation not only of the computer program but also of the assumptions underlying the model itself.) For example, consider a company which builds a simulation model of its computer system.

Since simulation models are generally better at comparing alternatives than at determining absolute answers, a model of the computer system which is sufficiently valid to compare, in a relative sense, three proposed job scheduling policies may not be valid enough to determine quite as precisely the average response time of the computer for a particular scheduling policy when the arrival rate of jobs is hypothesized to increase by fifty percent.

(4) A simulation model should be validated relative to a specified set of criteria, namely, the criteria that will actually be used for decision-making.

(5) Validation is not something to be attempted after the simulation model is already developed only if there is time and money still remaining. Instead, model development and validation should be done hand-in-hand throughout the course of the simulation study. (Our experience indicates that this recommendation is often not followed.)

(6) The use of formal statistical procedures is only part of the validation process; at the present time, most of the "validation" done in practice seems to be of the subjective variety as discussed in Subsection 4.A. One reason for this is that most classical statistical techniques cannot be directly applied in the context of simulation model validation. (See Subsection 4.C and [10] for further discussion.)

4. A Three-Step Approach to Validation

Probably the most important paper in the validation literature is that of Naylor and Finger [13], where a three-step approach is given for "validating" a simulation model. Here, we augment their approach

which was described in rather philosophical terms, by giving specific recommendations and examples of how to carry out each of the three steps.

A. Develop a Model with High Face Validity

The primary objective during the first step of validation is to develop a model with high "face" validity, i.e., a model which, on the surface, seems reasonable to people who are knowledgeable about the system under study. In order to develop such a model, the simulation modelers should make use of all existing information including:

- (i) Conversations with "experts" - A simulation model is not an abstraction developed by a modeler working in isolation, but rather the modeler should work closely with people that are intimately familiar with the system.
- (ii) Existing theory - For example, if one is modeling a service system such as a bank and the arrival rate of customers is constant over some period of time, then theory tells us that the interarrival times of customers are quite likely to be i.i.d. exponential r.v.'s or, in other words, customers arrive in accordance with a Poisson process. (See Cinlar [3] for a more complete discussion.)
- (iii) Observations of the system - If one is modeling a multi-teller bank with jockeying, then interarrival times are collected and used to fit a theoretical interarrival time distribution, service times are collected and used to fit a theoretical service time distribution, and the bank is observed to

construct a model of how people jockey from one line to another. (In collecting data on the system under study, however, care must be taken to insure that the data are representative of what one actually wants to model. For example, the data collected during a military field test (see Subsection 4.C) may not be representative of actual combat conditions due to differences in troop behavior and battlefield smoke. Schellenberger [14] discusses this and other aspects of data validity.)

- (iv) General knowledge - In building a model, one should seek out and use relevant results from similar models, so as not to "re-invent the wheel" with each new study.
- (v) Intuition - It will often be necessary to use one's intuition to hypothesize how certain components of a complex system operate. It is hoped that these hypotheses can be substantiated during the later steps of the validation process.

We believe that it is also very important for the modelers to interact with the decision-makers (or managers) throughout the course of the simulation study. When a study is first conceptualized, a decision-maker may not have an understanding of the system sufficient to know precisely the ultimate objectives of the study. Thus, as the study proceeds and a better understanding of the system is obtained, this information should be conveyed to the decision-maker by the modeler, who in turn might revise his objectives for the study. Not only will this approach increase the actual validity of the model but, in addition, the "perceived validity" to the decision-maker of the model will

be increased. A decision-maker is much more likely to accept as valid and to use a model in whose development he was actively involved.

B. Empirically Test the Assumptions of the Model

The goal of the second step of validation is to test quantitatively the assumptions made during the initial stages of model development. We now give some examples of techniques which can be used for this purpose, all of which are of general applicability. If a theoretical probability distribution has been fit to some observed data and used as input to the simulation model, then the adequacy of the fit can be assessed by use of the chi-square or Kolmogorov-Smirnov (K-S) goodness-of-fit tests. It should be mentioned, however, that these tests are often misstated in statistics books. See Breiman [2, p. 187] and Conover [4, p. 295] for good discussions of the chi-square and K-S tests, respectively.

As stated in Subsection 4.A, it is important to use "representative" data in building a model; however, it is equally important to exercise care when structuring these data. For example, if two or more sets of observed data have been merged and used for some purpose in a model, then whether this pooling is correct can be determined by use of the Mann-Whitney or Kruskal-Wallis tests of homogeneity of populations (see [2, p. 286]). (In a simulation study of a post office which we performed, it was found that the service time distributions of different postal clerks were not the same, since one clerk engaged in a conversation with each of his customers. Thus, it was not appropriate to fit one distribution to a pooled set of observed service time data.)

One of the most useful tools during the second step of validation is sensitivity analysis. For example, this technique can be used to determine how much the model output will vary with a small change in an input parameter. If the output is particularly sensitive to some parameter, then a better estimate of it should be obtained. Another important use of sensitivity analysis is to determine the level of detail at which a particular subsystem is to be modeled. Sometimes a simulation model is developed which is so detailed that one can only afford to run it for a few replications and, thus, a thorough analysis of the system under study is impossible. In this case, the modelers might determine what subsystem's model is resulting in an excessive running time and try to develop a simpler (and less expensive) representation of this subsystem. Both representations of the entire system are then run and the output data are compared for significant differences. If the simpler model produces "similar" results, then it can safely be used for a detailed study of the system. This use of sensitivity analysis was employed in the freeway simulation model discussed in Gafarian and Walsh [5] (and conveyed to us in a personal communication with the first author).

C. Determine How Representative the Simulation Output Data Are

Probably the most definitive test of the validity of a simulation model is to establish that the model output data closely resemble the output data that would be expected from the actual system. If a system similar to the one being studied now exists, then a

simulation model of the existing system is developed and its output data are compared to data from the actual existing system.

(These system data might be available from historical records or might have to be collected explicitly for validation purposes.

Furthermore, the time required to construct a model of the existing system will probably not be wasted, since such a model will be needed to compare definitively the present system to proposed system designs.) If the two sets of output data compare favorably, then the model of the existing system is modified so that it represents the proposed system. Although we cannot be sure that the model of the proposed system is "valid," we hopefully have more confidence than if the comparison had not been made.

The above idea will be used to validate a simulation model of a welfare office's operations which is being developed by HEW for the purpose of evaluating the effect of various proposed administrative policies, using such performance measures as applicant delay and accuracy of welfare payments. Here the "existing system" will be a welfare office in Massachusetts run under current administrative policy.

A number of statistical tests have been suggested in the validation literature for comparing the output data from a simulation model with those from the corresponding real-world system (see, for example, Shannon [15, p. 208]). However, the comparison is not as simple as it might appear, since the output processes of almost all real-world systems and simulations are nonstationary and autocorrelated.

Thus, classical statistical tests based on i.i.d. observations are not directly applicable. Furthermore, we question whether hypothesis tests, as compared to constructing confidence intervals for differences, are even the appropriate statistical approach. Since the model is only an approximation to the actual system, a null hypothesis that the system and model are the "same" is clearly false. We believe that it is more useful to ask whether or not the differences between the system and the model are significant enough to affect any conclusions derived from the model. For a discussion of statistical procedures which can be used for comparing system and model output data, see [10].

In addition to statistical procedures, one can use a Turing test to compare the output data from the model to those of the system. People knowledgeable about the system are asked to examine one or more sets of system data and one or more sets of model data without knowing which sets are which. If these "experts" can differentiate between the system and model data, then their explanation of how they were able to do it is used to improve the model. Even if a similar existing system exists but definitive output data are not readily available, the same "experts" can be asked to evaluate how reasonable the simulation output data are, and this information might be used to improve the model. This idea was put to good use by the developers of the ISEM simulation model of the Air Force Manpower and Personnel System. (This model was designed to provide Air Force policy analysts with a system-wide view of the effects of various

proposed personnel policies.) The model was run under the Air Force's baseline personnel policy and the results were shown to Air Force analysts and decision-makers who subsequently identified some discrepancies between model and perceived system behavior. This information was used to improve the model and, after several additional evaluations and improvements, a model was obtained which appeared to approximate closely current Air Force policy.

If the decisions to be made with a simulation model are of particularly great importance, then field tests are sometimes used (primarily by the military) to obtain system output data for validation purposes. For example, suppose some military organization is thinking of purchasing a weapons system for which it is infeasible or too expensive to perform a complete set of evaluational tests. As an alternative, a simulation model of the system is developed and then a prototype of the actual system is field tested on a military reservation for one or more specified scenarios. If the model and system output data compare closely for each of the specified scenarios, then the "validated" simulation model is used to evaluate the system for scenarios for which system field tests are not possible. For a further discussion of field tests, see [15, p. 231].

Up to now we have discussed validating a simulation model relative to past or present system output data; however, a perhaps more definitive test of a model is to establish its ability to predict future system behavior. Since models often evolve over time

and are used for more than one application, there is often an opportunity for such prospective validation. For example, if a model is used to decide which version of a proposed system to build, then after the system has been built and sufficient time has elapsed for output data to be collected, these data may be compared with the predictions of the model. If there is reasonable agreement, then we have increased confidence in the "validity" of the model. On the other hand, discrepancies between the two data sets are hopefully used to update the model. Regardless of the accuracy of a model's past predictions, a model should be carefully scrutinized before each new application, since a change in purpose or the passage of time may have invalidated some aspect of the existing model.

5. Additional Considerations in Validation

In Subsection 4.C we discussed comparing the output data from a simulation model with those from a corresponding existing system. However, if the system input and output data are complete enough and in the proper form, then it may be possible to perform the suggested comparison in a statistically more efficient manner. Since this idea is best illustrated by means of an example, consider the multi-teller bank discussed under (iii) in Subsection 4.A. Suppose that it is desired to validate a simulation model of the bank relative to the criterion of average delay of a customer between 12 and 1 P.M. (the busiest period in the bank). Suppose further that in collecting data from the bank, it is possible to observe the number of customers in the bank at 12 and, more importantly, the interarrival time, the

service time, and the delay in queue corresponding to each customer who arrives (and completes his delay) between l_2 and l_1 . Then, rather than running the model by generating the required interarrival and service times from the fitted theoretical distributions, it is preferable to drive the model with the actual observed interarrival and service times (i.e., no r.v.'s are generated) and to initialize the model at l_2 with the number of customers actually observed in the bank. (For this simple example we are, in effect, validating our model of jockeying, while for a more complex model, we would be validating everything in the model but the fitted theoretical distributions.) By comparing the bank and the model under a similar "statistical environment," we reduce the variance of the difference between the average delay in the bank and the average delay in the model, resulting in a more precise assessment of the difference between the model and the bank.

The idea of comparing a model and the corresponding system under the same statistical conditions is similar to the use of the variance reducing technique "common random numbers" in simulation (see Kleijnen [8, p. 200] and the use of "blocking" in statistical experimental design (see Box, Hunter, and Hunter [1, p. 102]). It should be mentioned, however, that we don't recommend using historical system input data to drive a model for the purpose of making production runs.

Sometimes one uses historical input data to build a model and then compares the model output data with the corresponding historical output data. If the agreement is not good, then the parameters or the structure of the model are "manipulated" and the resulting output

data are again compared to the historical output data. This procedure, which we call *calibration* of a model, is continued until the two data sets closely agree. However, we must ask whether this procedure produces a valid model for the system, in general, or whether the model is only representative of the particular set of input data. To answer this question (in effect, to validate the model), one can use a completely independent set of historical input and output data. The "calibrated" model might be driven by the second set of input data (in a manner similar to that described above) and the resulting model output data compared to the second set of historical output data. This idea of using one set of data for calibration and another independent set for validation, seems to be fairly common in economics and the biological sciences. In particular, it was used by the Crown Zellerbach Corporation in developing a simulation model of tree growth. Here the historical data were available from the U. S. Forest Service.

Acknowledgements

The author would like to thank David Kelton for reading the entire manuscript and making a number of valuable editorial suggestions. The author also gratefully acknowledges the following people who were kind enough to share their validation experiences: Yen Chao (FEDSIM), Claude Delfosse (CACI, Inc.), Andy Gafarian (University of Southern California), Murray Geisler (Logistics Management Institute), James Lin (Crown Zellerbach Corp.), Alan Pritsker (Purdue University and Pritsker and Associates, Inc.), Edward Russell (CACI, Inc.), Robert Sargent (Syracuse University), and Mitchell Spiegel (FEDSIM).

References

1. Box, G. E. P., W. G. Hunter, and J. S. Hunter, Statistics for Experimenters, Wiley, New York, (1978).
2. Breiman, L., Statistics: With a View Toward Applications, Houghton Mifflin, Boston, (1973).
3. Cinlar, E., Introduction to Stochastic Processes, Prentice-Hall, Englewood Cliffs, N.J., (1975).
4. Conover, W. J., Practical Nonparametric Statistics, Wiley, New York, (1971).
5. Gafarian, A. V. and J. E. Walsh, "Methods for Statistical Validation of a Simulation Model for Freeway Traffic Near an On-Ramp," Transpn. Res. 4, 379-384 (1970).
6. Gass, S. I., "Evaluation of Complex Models," Computers and Operations Research 4, 25-37 (1977).
7. Gross, D. and C. M. Harris, Fundamentals of Queueing Theory, Wiley, New York, (1974).
8. Kleijnen, J. P. C., Statistical Techniques in Simulation, Part 1, Marcel Dekker, New York, (1974).
9. Law, A. M., "Statistical Analysis of Simulation Output Data with SIMSCRIPT II.5," CACI, Inc., (1979).
10. Law, A. M., "Validation of Simulation Models, II: Comparison of Real-World and Simulation Output Data," Technical Report No. 78-15, Dept. of Industrial Engineering, University of Wisconsin (Fall 1979).
11. Law, A. M. and W. D. Kelton, "Confidence Intervals for Steady-State Simulations, I: A Survey of Fixed Sample Size Procedures," Technical Report No. 78-5, Dept. Of Industrial Engineering, University of Wisconsin (1979)

12. Law, A. M. and W. D. Kelton, "Confidence Intervals for Steady-State Simulations, II: A Survey of Sequential Procedures," Technical Report No. 78-6, Dept. of Industrial Engineering, University of Wisconsin (1978).
13. Naylor, T. H. and J. M. Finger, "Verification of Computer Simulation Models," Management Sci. 14, 92-101 (1967).
14. Schellenberger, R. E., "Criteria for Assessing Model Validity for Managerial Purposes," Decision Sciences 5, 644-653 (1974).
15. Shannon, R. E., Systems Simulation: The Art and Science, Prentice-Hall, Englewood Cliffs, N.J., (1975).
16. U. S. General Accounting Office, "Guidelines for Model Evaluation," Report PAD-79-17, Washington (1979).
17. Van Horn, R. L., "Validation of Simulation Results," Management Sci. 17, 247-258 (1971).

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 78-14	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Validation of Simulation Models, I: An Overview and Survey of Real-World Practice		5. TYPE OF REPORT & PERIOD COVERED Technical
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Averill M. Law		8. CONTRACT OR GRANT NUMBER(s) N00014-76-C-0403
9. PERFORMING ORGANIZATION NAME AND ADDRESS University of Wisconsin 750 University Avenue Madison, Wisconsin 53706		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR 047-145
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Arlington, Virginia 22217		12. REPORT DATE August 1979
		13. NUMBER OF PAGES 21
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) Office of Naval Research Branch Office 536 S. Clark St., Room 286 Chicago, Illinois 60605		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release, distribution unlimited		
17. DISTRIBUTION STATEMENT (of the Abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Validation, Verification, Calibration		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) We consider the problem of assessing the degree to which a model constructed for use in a simulation study faithfully represents the corresponding real system. Our focus here is to survey general approaches and methods which have been used in practice, or could be implemented, including good programming techniques, verifying that the program itself is correct, overall goals of the validation of a model, and a general framework for the total validation effort. Several concrete examples are discussed to illustrate the methods as well as possible pitfalls.		