

VALIDATION, VERIFICATION, AND TESTING TECHNIQUES THROUGHOUT THE LIFE CYCLE OF A SIMULATION STUDY†

Osman Balci

Department of Computer Science
Virginia Polytechnic Institute and State University
Blacksburg, Virginia 24061-0106, U.S.A.

ABSTRACT

Life cycle validation, verification, and testing (VV&T) is extremely important for the success of a simulation study. This paper surveys current software VV&T techniques and current simulation model VV&T techniques and describes how they can all be applied throughout the life cycle of a simulation study. The processes and credibility assessment stages of the life cycle are described and the applicability of the VV&T techniques for each stage is stated.

1. INTRODUCTION

Simulation is the process of constructing a model of a system which contains a problem and conducting experiments with the model on a computer for a specific purpose of experimentation to solve the problem. Credibility of simulation results not only depends on model correctness, but also is significantly influenced by accurate formulation of the problem. Therefore, validation, verification, and testing (VV&T) techniques must be employed throughout the life cycle of a simulation study starting with problem formulation and culminating with presentation of simulation results.

Model Validation is substantiating that the model, within its domain of applicability, behaves with satisfactory accuracy consistent with the study objectives. Model validation deals with building the *right* model. It is conducted by running the model under the "same" input conditions that drive the system and by comparing model behavior with the system behavior.

Model Verification is substantiating that the model is transformed from one form into another, as intended, with sufficient accuracy. Model verification deals with building the model *right*. The accuracy of transforming a problem formulation into a model specification or the accuracy of converting a model representation in micro flowchart into an executable computer program is evaluated in model verification.

Model Testing is demonstrating that inaccuracies exist or revealing the existence of errors in the model. In model testing, we subject the model to test data or test cases to see if it functions properly. "Test failed" implies the failure of the model, not the test. Testing is conducted to perform validation and verification. Some tests are

devised to evaluate the behavioral accuracy (i.e., validity) of the model, and some tests are intended to judge the accuracy of model transformation from one form into another (verification). Therefore, we commonly refer to the whole process as *model VV&T*.

Model VV&T is employed to prevent the occurrence of three major types of errors in conducting simulation studies (Balci 1990): *Type I Error* is the error of rejecting the model credibility when in fact the model is sufficiently credible. *Type II Error* is the error of accepting the model credibility when in fact the model is *not* sufficiently credible. *Type III Error* is the error of solving the wrong problem.

Every organization conducting a substantial simulation study should have a department or group called Simulation Quality Assurance (SQA). The SQA group is responsible for total quality management and closely works with the simulation project managers in planning, preparing test cases, and administering some of the VV&T activities throughout the simulation study. The SQA is a managerial approach which is critically essential for the success of a simulation study. Ören (1981, 1986, 1987) presents concepts, criteria, and paradigms which can be used in establishing an SQA program within an organization.

The purpose of this paper is to survey current software VV&T techniques and current model VV&T techniques and describe how they can all be applied throughout the life cycle of a simulation study. Section 2 presents the life cycle of a simulation study and provides guidelines for conducting its ten processes. The VV&T techniques are briefly described under a taxonomy in Section 3. Section 4 describes the credibility assessment stages of the life cycle and shows the applicability of the VV&T techniques for each stage. Concluding remarks and research directions are given in Section 5.

2. LIFE CYCLE OF A SIMULATION STUDY

The life cycle of a simulation study is presented in Figure 1 (Balci 1990; Nance 1994). The phases are shown by shaded oval symbols. The dashed arrows describe the processes which relate the phases to each other. The solid arrows refer to the credibility assessment stages. Banks *et al.* (1987) and Knepell and Arangno (1993) review other modeling processes for developing simulations.

The life cycle should not be interpreted as strictly sequential. The sequential representation of the dashed arrows is intended to show the direction of development

† This is a condensed version of a paper with the same title appeared in *Annals of Operations Research* volume on Simulation and Modeling [Balci 1994a].

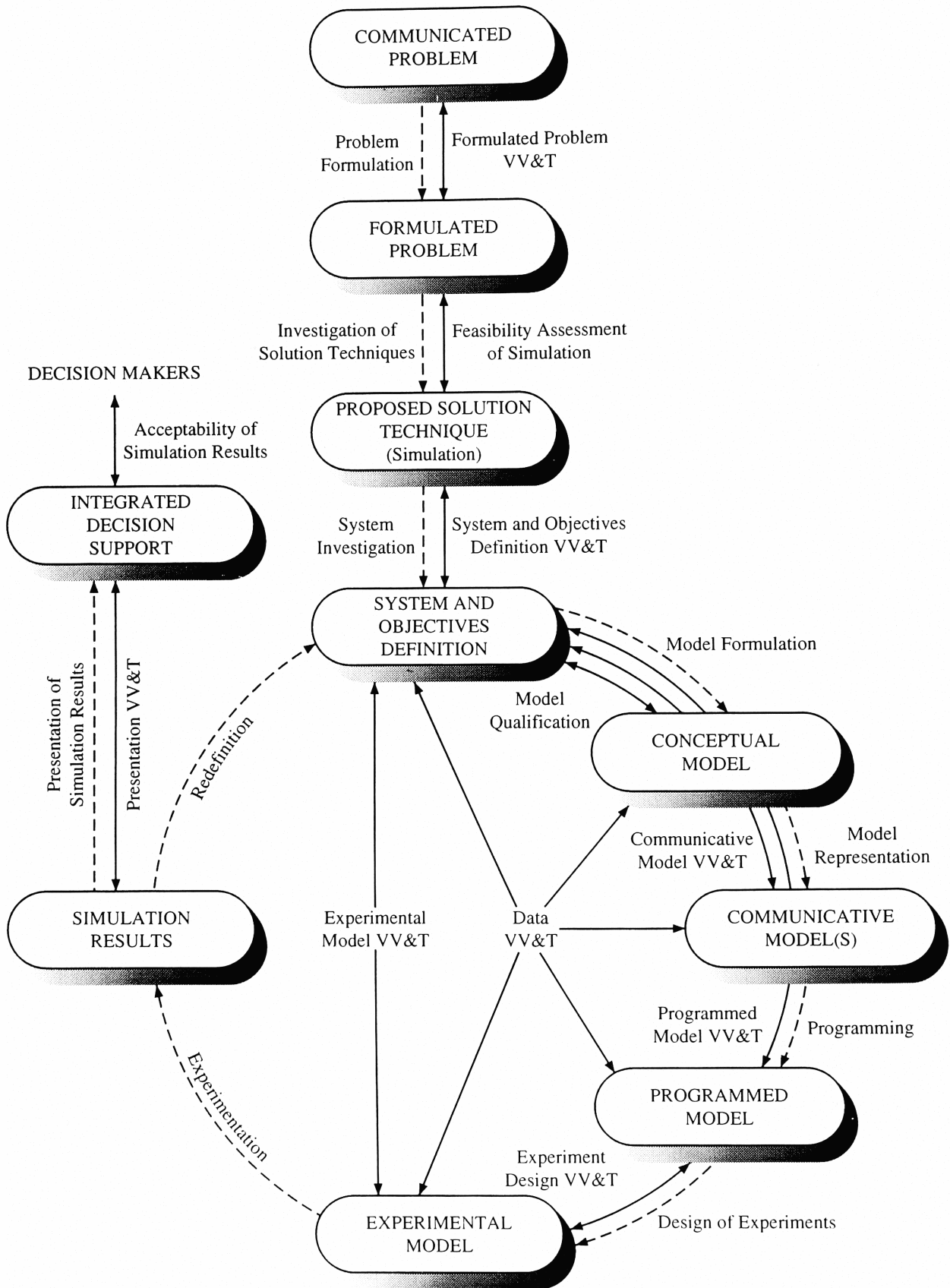


Figure 1: The Life Cycle of a Simulation Study

throughout the life cycle. The life cycle is iterative in nature and reverse transitions are expected. Every phase of the life cycle has an associated VV&T activity. Deficiencies identified by a VV&T activity may necessitate returning to an earlier process and starting all over again.

The VV&T is not a phase or step in the life cycle, but a continuous activity throughout the entire life cycle. Conducting the VV&T for the first time in the life cycle when the experimental model is complete is analogous to the teacher who gives only a final examination (Hetzel 1984). No opportunity is provided throughout the semester to notify the student that he or she has serious deficiencies. Severe problems may go undetected until it is too late to do anything but fail the student. Frequent tests and homeworks throughout the semester are intended to inform the students about their deficiencies so that they can study more to improve their knowledge as the course progresses.

The situation in the VV&T is exactly analogous. The VV&T activities throughout the entire life cycle are intended to reveal any quality deficiencies that might be present as the simulation study progresses from the communication of the problem until the implementation of the simulation results. This allows us to identify and rectify quality deficiencies during the life cycle phase in which they occur.

The ten processes of the life cycle are shown by the dashed arrows in Figure 1. Although each process is executed in the order indicated by the dashed arrows, an error identified may necessitate returning to an earlier process and starting all over again. Some guidelines are provided in (Balci 1994a) for each of the ten processes.

3. VALIDATION, VERIFICATION, AND TESTING TECHNIQUES

Figure 2 shows a taxonomy which categorizes the VV&T techniques into six distinct credibility assessment perspectives: informal, static, dynamic, symbolic, constraint, and formal. The level of mathematical formality of each category increases from very informal on the far left to very formal on the far right. Likewise, the complexity also increases as the category becomes more formal (Whitner and Balci 1989).

It should be noted that some of the categories presented in Figure 2 possess similar characteristics and in fact have techniques which overlap from one category to another. However, a distinct difference between each classification exists.

The techniques are described in (Balci 1994a).

Informal VV&T techniques are among the most commonly used ones. They are called informal because the tools and approaches used rely heavily on human reasoning and subjectivity without stringent mathematical formalism. The "informal" label does not imply any lack of structure or formal guidelines for the use of the techniques.

Static VV&T techniques are concerned with accuracy assessment on the basis of characteristics of the static model source code. Static techniques do not require machine execution of the model, but mental execution may be used. The techniques are very popular and widely used, with many automated tools available to assist the VV&T. The simulation language compiler is itself a static VV&T tool.

Validation, Verification, and Testing Techniques

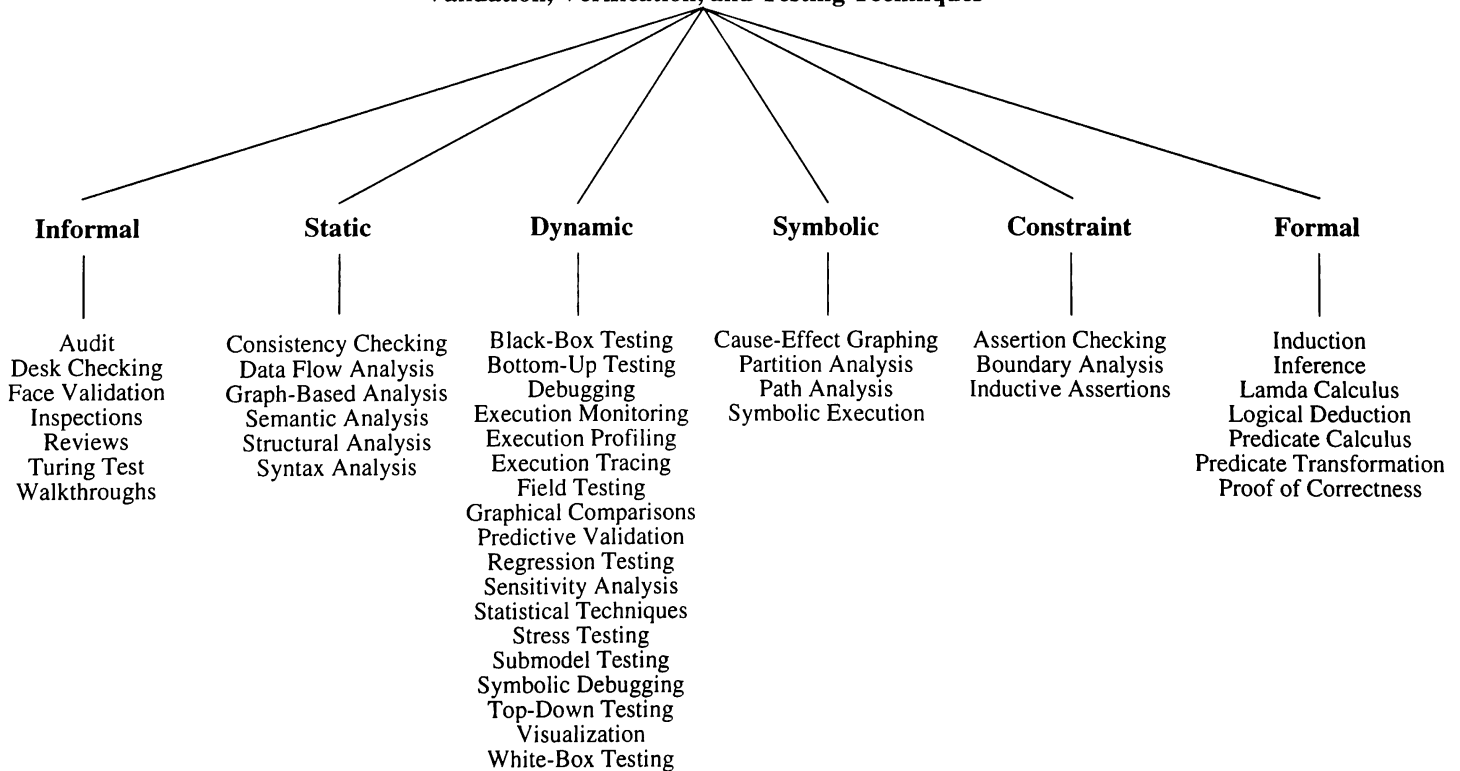


Figure 2: A Taxonomy of Validation, Verification, and Testing Techniques

Static VV&T techniques can obtain a variety of information about the structure of the model, coding techniques and practices employed, data and control flow within the model, syntactical accuracy, and internal as well as global consistency and completeness of implementation (Whitner and Balci 1989).

Dynamic VV&T techniques require model execution and are intended for evaluating the model based on its execution behavior. Most dynamic VV&T techniques require model instrumentation.

The insertion of additional code (probes) into the executable model for the purpose of collecting information about model behavior during execution is called *model instrumentation*. Probe locations are determined manually or automatically based on static analysis of model structure. Automated instrumentation is accomplished by a preprocessor which analyzes the model static structure (usually via graph-based analysis) and inserts probes at appropriate places.

Dynamic VV&T techniques are usually applied using the following three steps. In Step 1, the programmed or experimental model is instrumented. In Step 2, the instrumented model is executed, and in Step 3, the model output is analyzed and dynamic model behavior is evaluated.

Much research has been conducted in applying statistical techniques for dynamic VV&T. Balci (1994a) presents the statistical techniques proposed for model validation and lists related references.

The statistical techniques generally require that the system being modeled is completely observable, i.e., all data required for model validation can be collected from the system. Model validation is conducted by using the statistical techniques to compare the model output data with the corresponding system output data when the model is run with the "same" input data that derive the real system. Due to the multiple response problem (Shannon 1975), the comparison of model and system outputs must be carried out by using a multivariate statistical technique to incorporate the correlations among the output variables.

A validation procedure based on the use of simultaneous confidence intervals is presented in (Balci 1994a). Whenever possible, a multivariate statistical technique should be used to conduct model validation.

Symbolic VV&T techniques, like dynamic VV&T techniques, are used to evaluate the dynamic behavior of the model during execution. In symbolic analysis, symbolic inputs are provided to a simulation model as input and expressions are produced as output which are derived from the transformation of the symbolic data along model execution paths.

Constraint VV&T techniques are employed to assess model correctness using assertion checking, boundary analysis, and inductive assertions.

Formal VV&T techniques are based on formal mathematical proof of correctness. If attainable, formal proof of correctness is the most effective means of model VV&T. Unfortunately, "if attainable" is the overriding point with regard to formal VV&T techniques. Current state-of-the-art formal proof of correctness techniques are simply not capable of being applied to even a reasonably complex simulation model. However, formal techniques serve as the foundation for other VV&T techniques.

4. CREDIBILITY ASSESSMENT STAGES

It is very important to understand the principles of simulation model VV&T when applying the VV&T techniques throughout the entire life cycle of a simulation study. Balci (1994b) presents 15 principles that help the researchers, practitioners and managers better understand what model VV&T is all about. These principles serve to provide the underpinnings for the VV&T techniques described in Section 3. Understanding and applying these principles is crucially important for the success of a simulation study.

Table 1 marks the VV&T techniques that are applicable for each of the ten credibility assessment stages described in (Balci 1994a). The more of these techniques we apply the more confidence we gain in the credibility of a life cycle phase. The VV&T activities should continue until a sufficient level of confidence is achieved.

5. CONCLUDING REMARKS AND RESEARCH DIRECTIONS

The life cycle application of VV&T is extremely important for successful completion of complex and large-scale simulation studies. This point must be clearly understood by the sponsor of the simulation study and the organization conducting the simulation study. The sponsor must furnish funds under the contractual agreement and require the contractor to apply VV&T throughout the entire life cycle.

Assessing credibility throughout the life cycle of a simulation study is an onerous task. Applying the VV&T techniques throughout the life cycle is time consuming and costly. In practice, under time pressure to complete a simulation study, the VV&T and documentation are sacrificed first. Computer-aided assistance for the VV&T is required to alleviate these problems. More research is needed to bring automation to the application of the VV&T techniques.

Integration of VV&T with model development is crucial. This integration is best achieved within a computer-aided simulation software engineering environment (Balci 1986; Balci and Nance 1987). More research is needed for this integration.

How much to test or when to stop testing depends on the study objectives. The testing should continue until we achieve sufficient confidence in credibility and acceptability of simulation results. The sufficiency of the confidence is dictated by the study objectives.

Establishing a simulation quality assurance (SQA) program within the organization conducting the simulation study is extremely important for successful credibility assessment. The SQA management structure goes beyond VV&T and is also responsible for assessing other model quality characteristics such as maintainability, reusability, and usability (human-computer interface). The management of the SQA program and the management of the simulation project must be independent of each other and neither should be able to overrule the other (Schach 1993).

Subjectivity is and will always be part of the credibility assessment for a reasonably complex simulation study. The reason for subjectivity is two-fold: modeling is an art and credibility assessment is situation depen-

Table 1: Applicability of the VV&T Techniques for the Credibility Assessment Stages

	FP VV&T	FA of Sim.	S&OD VV&T	Model Qual.	CM VV&T	PM VV&T	ED VV&T	Data VV&T	EM VV&T	Pres. VV&T
Assertion Checking						✓	✓	✓	✓	
Audit	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Black-Box Testing						✓	✓		✓	
Bottom-Up Testing						✓			✓	
Boundary Analysis						✓			✓	
Cause-Effect Graphing	✓				✓	✓			✓	
Consistency Checking	✓		✓	✓	✓	✓	✓	✓	✓	✓
Data Flow Analysis					✓	✓	✓	✓	✓	
Debugging						✓	✓		✓	
Desk Checking	✓		✓	✓	✓	✓	✓	✓	✓	✓
Execution Monitoring						✓	✓		✓	
Execution Profiling						✓	✓		✓	
Execution Tracing						✓	✓		✓	
Face Validation	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Field Testing									✓	
Graph-Based Analysis					✓	✓	✓		✓	
Graphical Comparisons						✓	✓		✓	
Induction						✓			✓	
Inductive Assertions						✓			✓	
Inference						✓			✓	
Inspections	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Lamda Calculus						✓			✓	
Logical Deduction						✓			✓	
Partition Analysis						✓			✓	
Path Analysis						✓	✓		✓	
Predicate Calculus						✓			✓	
Predicate Transformation						✓			✓	
Predictive Validation									✓	
Proof of Correctness						✓			✓	
Regression Testing						✓			✓	
Reviews	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Semantic Analysis						✓	✓		✓	
Sensitivity Analysis						✓	✓		✓	
Statistical Techniques								✓	✓	
Stress Testing						✓			✓	
Structural Analysis	✓		✓		✓	✓	✓		✓	✓
Submodel Testing						✓			✓	
Symbolic Debugging						✓	✓		✓	
Symbolic Execution						✓	✓		✓	
Syntax Analysis						✓	✓		✓	
Top-Down Testing						✓			✓	
Turing Test									✓	
Visualization						✓			✓	
Walkthroughs	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
White-Box Testing						✓			✓	

dent. A unifying approach based on the use of indicators measuring qualitative as well as quantitative aspects of a simulation study should be developed.

REFERENCES

- Balci, O. 1986. Requirements for Model Development Environments. *Computers & Operations Research* 13:53-67.
- Balci, O. 1990. Guidelines for Successful Simulation Studies. In *Proceedings of the 1990 Winter Simulation Conference*, ed. O. Balci, R. P. Sadowski, and R. E. Nance, 25-32. IEEE, Piscataway, New Jersey.
- Balci, O. 1994a. Validation, Verification, and Testing Techniques Throughout the Life Cycle of a Simulation Study. *Annals of Operations Research*, to appear.
- Balci, O. 1994b. Principles of Simulation Model Validation, Verification, and Testing. Technical Report TR-94-24, Department of Computer Science, Virginia Tech, Blacksburg, Virginia.
- Balci, O. and R. E. Nance. 1987. Simulation Model Development Environments: A Research Prototype. *Journal of Operational Research Society* 38:753-763.
- Banks, J., D. Gerstein, and S. P. Searles. 1987. Modeling Processes, Validation, and Verification of Complex Simulations: A Survey. In *Methodology and Validation*, ed. O. Balci, 13-18. SCS, San Diego, California.
- Hetzl, W. 1984. *The Complete Guide to Software Testing*. QED Information Sciences, Wellesley, Massachusetts.
- Knepell, P. L. and D. C. Arango. 1993. *Simulation Validation: A Confidence Assessment Methodology*. Monograph 3512-04, IEEE Computer Society Press, Los Alamitos, California.
- Nance, R. E. 1994. Conical Methodology: An Evolutionary Convergence of Systems and Software Engineering. *Annals of Operations Research*, to appear.
- Ören, T. I. 1981. Concepts and Criteria to Assess Acceptability of Simulation Studies: A Frame of Reference. *Communications of the ACM* 24:180-189.
- Ören, T. I. 1986. Artificial Intelligence in Quality Assurance of Simulation Studies. In *Modelling and Simulation Methodology in the Artificial Intelligence Era*, ed. M. S. Elzas, T. I. Ören, and B. P. Zeigler, 267-278. North Holland, Amsterdam, Holland.
- Ören, T. I. 1987. Quality Assurance Paradigms for Artificial Intelligence in Modelling and Simulation. *Simulation* 48:149-151.
- Schach, S. R. 1993. *Software Engineering*. Second Edition, Irwin, Homewood, Illinois.
- Shannon, R. E. 1975. *Systems Simulation: The Art and Science*. Prentice-Hall, Englewood Cliffs, New Jersey.
- Whitner, R. B. and O. Balci. 1989. Guidelines for Selecting and Using Simulation Model Verification Techniques. In *Proceedings of the 1989 Winter Simulation Conference*, ed. E. A. MacNair, K. J. Musselman, and P. Heidelberger, 559-568. IEEE, Piscataway, New Jersey.

AUTHOR BIOGRAPHY

OSMAN BALCI is an Associate Professor of Computer Science at Virginia Polytechnic Institute and State University (VPI&SU). He received B.S. and M.S. degrees from Boğaziçi University in Istanbul, Turkey in 1975 and 1977, and M.S. and Ph.D. degrees from Syracuse University (N.Y.) in 1978 and 1981. He joined the Computer Science faculty at VPI&SU in 1981. Dr. Balci is the Editor-in-Chief of *Annals of Software Engineering*—an international journal; Guest Editor of a special volume of *Annals of Operations Research* journal on Simulation and Modeling; Associate Editor of *Transactions on Operational Research*; Associate Editor of *ORSA Journal on Computing*; Quality Assurance Area Editor of *ACM Transactions on Modeling and Computer Simulation*; Simulation and Modeling Category Editor of *ACM Computing Reviews*; Editor of *ORSA CSTS Newsletter*; and Member of the Editorial Advisory Board of Operations Research/Computer Science Interfaces book series of Kluwer Academic Publishers. He has served as: Chairman (1993-94) and Vice Chairman (1992-93) of the ORSA Computer Science Technical Section; General Chairman of the 1992 ORSA CSTS conference; Proceedings Editor of the 1990 Winter Simulation Conference; Program Chairman and Proceedings Editor of the SCS conference on Simulation Methodology and Validation, 1986-87; Vice Chairman of ACM SIGSIM, 1985-87; and an Associate Editor of ACM SIGSIM's *Simuletter*, 1983-86. Dr. Balci received the College of Arts and Sciences Certificate of Teaching Excellence in 1993 at VPI&SU; Outstanding Teacher Award in 1985 and Teaching Excellence Award in 1988, 1991, and 1992 in his department at VPI&SU. He received a plaque for Distinguished Service from the Board of Winter Simulation Conference in 1990. He has published numerous technical articles and edited/co-edited four books. Dr. Balci has been a Principal Investigator, a Co-Principal Investigator, or an Investigator on research grants and contracts sponsored by the U.S. Navy, NSF, and IBM with a total funding of \$1.773 million. His current research interests center on software engineering, simulation and modeling, expert systems, and performance evaluation. Dr. Balci is a member of Alpha Pi Mu, Sigma Xi, Upsilon Pi Epsilon, ACM, IEEE CS, and ORSA.