AN INTRODUCTORY TUTORIAL ON VERIFICATION AND VALIDATION OF SIMULATION MODELS

Robert G. Sargent

Department of Electrical Engineering and Computer Science
L. C. Smith College of Engineering and Computer Science
Syracuse University
Syracuse, NY 13244, USA

ABSTRACT

Model verification and validation are defined, and why model verification and validation are important is discussed. A graphical paradigm that shows how verification and validation are related to the model development process and a flowchart that shows how verification and validation is part of the model development process are presented and discussed. The three approaches to deciding model validity are described. Comments are made on the importance of model accuracy and documentation. An overview of conducting verification and validation is presented and a recommended procedure for verification and validation is given.

1 INTRODUCTION

An introduction to verification and validation of simulation models is given in this paper. Verification and validation are concerned with determining whether a model and its results are “correct” for a specific use or purpose. Formally, model verification is defined as “ensuring that the computer program of the computerized model and its implementation are correct” and model validation is defined as the “substantiation that a computerized model within its domain of applicability possesses a satisfactory range of accuracy consistent with the intended application of the model.” Our discussion of verification and validation of simulation models will be primarily concerned with simulation models that are used to predict system behaviors such as systems outputs. Two related topics are model credibility and model usability. Model credibility is concerned with developing in (potential) users the confidence they require in order to use a model and in the information derived from that model. Model usability is determining that the model and its user instructions are easy to use.

It is important that verification and validation of a simulation model be performed for each use or purpose of a model. If the purpose of a simulation model is to answer a variety of questions, the validity of the model needs to be determined with respect to each question. The developers and users of simulation models, the decision makers using information obtained from the results of these models, and the individuals affected by decisions based on such models are all rightly concerned with whether a model and its results are “correct” for each question being addressed.

Numerous sets of experimental conditions are usually required to define the domain of a model’s intended applicability. (A set of experimental conditions contains a set of values for the set of variables that define the domain of applicability.) A model may be valid for one set of experimental conditions and invalid in another. A model is considered valid for a set of experimental conditions if the model’s accuracy is within its acceptable range of accuracy, which is the accuracy required of the model for its
intended purpose. This usually requires that the model’s output variables of interest (i.e., the model variables used in answering the questions that the model is being developed to answer) be identified and then their acceptable range of accuracy specified. A model’s acceptable range of accuracy should be specified prior to starting the development of the model or very early in the model development process. If the variables of interest are random variables, then properties and functions of the random variable such as means and variances are usually what is of primary interest and are what is used in determining model validity. Several versions of a model are usually developed prior to obtaining a satisfactory valid model. The substantiation that a model is valid, i.e., performing model verification and validation, is generally considered to be a process and is usually part of the (total) model development process.

2 MODEL DEVELOPMENT PROCESS WITH VERIFICATION AND VALIDATION

In this section a graphical paradigm is presented in Subsection 2.1 that relates model verification and validation to the model development process. Then in Subsection 2.2 the model development process is described that includes verification and validation.

2.1 A Simple Graphical Paradigm

There are two common ways to view how verification and validation relate to the model development process. One way uses a simple view and the other uses a complex view. A simple graphical paradigm is presented in Figure 1 that was developed by this author called the Simplified View of the Model Development Process (Sargent 1981, 1982, 1983, 2001, 2011, 2013). A more complex paradigm developed by this author that includes both the “Simulation World” and the “Real World” is contained in Sargent (2001, 2013).

![Figure 1: Simplified version of the model development process.](image)

Consider the simplified version of the model development process in Figure 1. The problem entity is the system (real or proposed), idea, situation, policy, or phenomena to be modeled; the conceptual model is the mathematical/logical/graphical representation (mimic) of the problem entity developed for a particular study; and the computerized model is the conceptual model implemented on a computer. The conceptual model is developed through an analysis and modeling phase, the computerized model is
developed through a computer programming and implementation phase, and inferences about the problem entity are obtained by conducting computer experiments on the computerized model in the experimentation phase.

We now relate model verification and validation to this simplified version of the model development process. (See Figure 1.) Conceptual model validation is defined as determining that the theories and assumptions underlying the conceptual model are correct and that the model representation of the problem entity is “reasonable” for the intended purpose of the model. Computerized model verification is defined as assuring that the computer programming and implementation of the conceptual model are correct. Operational validation is defined as determining that the model’s output behavior has a satisfactory range of accuracy for the model’s intended purpose over the domain of the model’s intended applicability. Data validity is defined as ensuring that the data necessary for model building, model evaluation and testing, and conducting the model experiments to solve the problem are adequate and correct.

2.2 Model Development Process

A model should be developed for a specific purpose or use and be a parsimonious model which means that it is as simple as possible yet meets its purpose. A simulation model is a structural model, which implies the model contains logical and causal relationships that occur in the systems. (Note: Structural models are different than empirical models which are developed purely from data as regression models are.) Developing a valid simulation model is an iterative process where several versions of a model are developed prior to obtaining a valid model.

The model development process should include model verification and validation. Following the paradigm given in Figure 1, the iterative process shown in Figure 2 can be used to develop a valid

![Figure 2: The model development iterative process.](image-url)
Sargent

simulation model (Sargent 1984). First a conceptual model is developed through initially analyzing the problem entity and then developing a model of the problem entity. The model should be a parsimonious model and need to have only the accuracy necessary to satisfy the purpose of the model, sometimes referred to as model fidelity. Then conceptual model validation is performed. This process is repeated until the conceptual model is satisfactory. Next a computerized model is developed of the (validated) conceptual model by developing a simulation model of the conceptual model and implementing it on a computer. Then computerized model verification is performed. This process is repeated until the computerized model is satisfactory. Lastly, operational validation is performed on the computerized model. Model changes required by conducting operational validity can be in either the conceptual model or in the computerized model. Verification and validation must be performed again when any model change is made. This process is repeated until a valid simulation model is obtained. As stated above, several versions of a model are usually developed prior to obtaining a valid simulation model. (The specifics of conducting conceptual model validation, computerized model verification, and operational validity are discussed in, e.g., Sargent (1984, 2011, 2013).)

It is often too costly and time consuming to determine that a model is absolutely valid over the complete domain of its intended applicability. Instead, tests and evaluations are conducted until sufficient confidence is obtained that a model can be considered valid for its intended purpose or use (Sargent 1982, 1984). If a test determines that a model does not have sufficient accuracy for any one of the sets of experimental conditions, then the model is invalid. However, determining that a model has sufficient accuracy for numerous experimental conditions does not guarantee that a model is valid everywhere in its applicable domain. Figure 3 contains two relationship curves regarding confidence that a model is valid (Confidence in Model) over the range of 0-100 percent as they would occur in most cases. (Note that these curves are qualitative conceptual curves and that they cannot be quantitatively calculated.) The cost curve (and a similar relationship holds for the amount of time) of performing model validation shows cost increases at an increasing rate as the confidence in the model increases. The value curve shows that the value of a model to a user increases as the confidence in a model increases but at a decreasing rate. (In some cases these curves may have a different shape for the lower confidence range but would usually be similar to what is shown in Figure 3 for the upper confidence range.) The cost of model validation is usually quite significant, especially when extremely high model confidence is required.

![Figure 3: Confidence that model is valid.](image)

**3 MODEL ACCURACY**

The accuracy required of a model for its intended use is needed in order to determine if a model is valid for that purpose or use. The accuracy required of a model is often not easy to obtain or determine. In almost all situations the model accuracy needs to be specified quantitatively; an exception is when a
The amount of model accuracy is being used to explore its qualitative behavior or to gain insights. (See, e.g., Pidd (2010) for a discussion on a spectrum of model use.) The accuracy needed is usually expressed as a range of the difference between the model’s output and the corresponding system output which is usually referred to as the model’s acceptable range of accuracy.

The amount of model accuracy is also needed for other steps in the model development process. (See Figures 1 and 2.) The amount of model accuracy is needed to develop the conceptual model. Recall that the conceptual model should be a parsimonious model with the model accuracy only as required for its purpose. Simulation models allow any level of detail. Thus, simulation model developers must be careful to develop conceptual models at the highest level possible that satisfies the model accuracy required. If the model accuracy required has not been specified prior to starting the model development process it needs to be developed and specified during the development of the conceptual model. The higher the amount of model accuracy required the more detailed a conceptual model must usually be. Also a higher amount of model accuracy usually requires additional input data, more analysis of the input data, more computer programming to develop the computerized model, and an increased amount of verification and validation testing. If the amount of model accuracy required is found to be not satisfactory during operational validity of the model, the conceptual model usually needs to be modified by adding more model detail.

The cost of model development increases as the amount of model accuracy increases. The amount of model accuracy is one of the most critical factors in model development. The amount of model accuracy depends on how the model is going to be used. Unfortunately, how to determine the minimal amount of accuracy required for a model has received little attention.

4 DECISION-MAKING APPROACHES

There are three basic decision-making approaches for deciding whether a simulation model is valid. Each of these three approaches uses a different decision-maker. All of the approaches require the model development team to conduct verification and validation as part of the model development process, which was discussed in Section 2. One decision-making approach, and a frequently used one, is for the model development team itself to make the decision as to whether a simulation model is valid. The decision is based on the results of the various tests and evaluations conducted as part of the model development process. It is usually better, however, to use one of the next two decision-making approaches, depending on which situation applies.

A better decision-making approach is to have the user(s) of a simulation model decide the validity of the model. In this approach the users of the simulation model are heavily involved with the model development team when the team is conducting verification and validation of the model and the users determine if the model is satisfactory in each phase of verification and validation. This approach is generally used with a model development team whose size is not large. Also, this approach aids in model credibility.

Another decision-making approach, usually called “independent verification and validation” and often referred to as IV&V, uses a third party to decide whether the simulation model is valid. The third party (the IV&V team) is independent of both the simulation development team(s) and the model sponsor/user(s). The IV&V approach is generally used with the development of large-scale simulation models, whose development usually involves several teams. The IV&V team needs to have a thorough understanding of the intended purpose(s) of the simulation model in order to conduct IV&V. There are two common ways that the IV&V team conducts IV&V: (a) IV&V is conducted concurrently with the development of the simulation model and (b) IV&V is conducted after the simulation model has been developed.

In the concurrent way of conducting IV&V, the model development team(s) gives their model verification and validation test results to the IV&V team as the simulation model is being developed. The IV&V team evaluates these results and provides feedback to the model development team regarding
whether the model verification and validation is satisfying the model requirements and when not, what the difficulties are. When conducting IV&V this way, the development of a simulation model should not progress to the next stage of development until the model has satisfied the verification and validation requirements in its current stage. It is the author’s opinion that this is the better of the two ways to conduct IV&V.

When IV&V is conducted after the simulation model has been completely developed, the evaluation performed by the IV&V team can range from simply evaluating the verification and validation conducted by the model development team to performing a separate thorough verification and validation effort themselves. Performing a complete IV&V effort after the model has been completely developed is usually both extremely costly and time consuming. This author’s view is that if IV&V is going to be conducted on a completed simulation model then it is usually best to only evaluate the verification and validation that has already been performed.

When an IV&V team concludes that a model is valid, there is a much greater likelihood that others will accept the model as valid and results from the model as being “correct.” Cases where this decision-making approach is helpful are (i) when the problem associated with the model has a high cost or involves a high risk situation and (ii) when public acceptance of results based on the model is desired.

5 CONDUCTING MODEL VERIFICATION AND VALIDATION

There are numerous approaches, methods, tests, and validation techniques used for verification and validation of simulation models. The most important factor affecting verification and validation is whether the problem entity being modeled is observable, meaning can data be collected on it to be used in developing the model and for model validation.

This author suggests that the approach to be used for verification and validation of simulation models is to follow the approach covered in Section 2. Below we briefly discuss conceptual model validation, computerized model verification, and operational validity. For an in depth coverage of these three topics, data validity, and the various validation techniques, see Sargent (2013) and references within. Information on other approaches, methods, and tests can be found in the references given at the end of this paper.

5.1 Conceptual Model Validation

Conceptual model validity is determining that (1) the theories and assumptions underlying the conceptual model are correct and (2) the model’s representation of the problem entity and the model’s structure, logic, and mathematical and causal relationships are “reasonable” for the intended purpose of the model. The theories and assumptions underlying the model should be tested using mathematical analysis and statistical methods on problem entity data. Examples of theories and assumptions are linearity, independence of data, and arrivals follow a Poisson process. Examples of applicable statistical methods are fitting distributions to data, estimating parameter values from the data, and plotting data to determine if the data are stationary. In addition, all theories used should be reviewed to ensure they were applied correctly. For example, if a Markov chain is used, does the system have the Markov property, and are the states and transition probabilities correct?

Each submodel and the overall model must be evaluated to determine if they are reasonable and correct for the intended purpose of the model. This should include determining if the appropriate detail and aggregate relationships have been used for the model’s intended purpose, and also if appropriate structure, logic, and mathematical and causal relationships have been used. The primary validation techniques used for these evaluations are face validation and traces. Face validation has experts on the problem entity evaluate the conceptual model to determine if it is correct and reasonable for its purpose. This usually requires examining the flowchart or graphical model (Sargent 1986), or the set of model equations. The use of traces is the tracking of entities through each submodel and the overall model to determine if the logic is correct and if the necessary accuracy is maintained. If errors are found in the conceptual model, it must be revised and conceptual model validation performed again.
5.2 Computerized Model Verification

Computerized model verification ensures that the computer programming and implementation of the conceptual model are correct. The major factor affecting verification is whether a simulation language or a higher level programming language such as FORTRAN, C, or C++ is used. The use of a special-purpose simulation language generally will result in having fewer errors than if a general-purpose simulation language is used, and using a general-purpose simulation language will generally result in having fewer errors than if a general purpose higher level programming language is used. (The use of a simulation language also usually reduces both the programming time required and the amount of flexibility, and increases the model execution times.)

When a simulation language is used, verification is primarily concerned with ensuring that an error free simulation language has been used, that the simulation language has been properly implemented on the computer, that a tested (for correctness) pseudo random number generator has been properly implemented, and that the model has been programmed correctly in the simulation language. The primary techniques used to determine that the model has been programmed correctly are structured walkthroughs (Law (2014) and Sargent (2013)) and traces.

If a higher level programming language has been used, then the computer program should have been designed, developed, and implemented using techniques found in software engineering. (These include such techniques as object-oriented design, structured programming, and program modularity.) In this case verification is primarily concerned with determining that the simulation functions (e.g., the time-flow mechanism, pseudo random number generator, and random variate generators) and the computerized (simulation) model have been programmed and implemented correctly.

There are two basic approaches for testing simulation software: static testing and dynamic testing (Fairley 1976). In static testing the computer program is analyzed to determine if it is correct by using such techniques as structured walkthroughs, correctness proofs, and examining the structure properties of the program. In dynamic testing the computer program is executed under different conditions and the values obtained (including those generated during the execution) are used to determine if the computer program and its implementations are correct. The techniques commonly used in dynamic testing are traces, investigations of input-output relations using different validation techniques, data relationship correctness, and reprogramming critical components to determine if the same results are obtained. If there are a large number of variables, one might aggregate the numerical values of some of the variables to reduce the number of tests needed or use certain types of design of experiments (Kleijnen 2008).

It is necessary to be aware while checking the correctness of the computer program and its implementation that errors found may be caused by the data, the conceptual model, the computer program, or the computer implementation. (See Whitner and Balci (1989) for a detailed discussion on model verification.)

5.3 Operational Validity

Operational validation is determining whether the simulation model’s output behavior has the accuracy required for the model’s intended purpose over the domain of the model’s intended applicability. This is where much of the validation testing and evaluation take place. Since the simulation model is used in operational validation, any deficiencies found may be caused by what was developed in any of the steps that are involved in developing the simulation model including developing the system’s theories or having invalid data.

Numerous validation techniques are applicable to operational validity. Which techniques and whether to use them objectively or subjectively must be decided by the model development team and the other interested parties. The major attribute affecting operational validity is whether the problem entity (or system) is observable, where observable means it is possible to collect data on the operational behavior of the problem entity. Table 1 gives a classification of the validation techniques used in operational validity.
based on the decision approach and system observability. “Comparison” means comparing the simulation model output behavior to either the system output behavior or another model output behavior using graphical displays and/or statistical tests and procedures. “Explore model behavior” means to examine the output behavior of the simulation model using appropriate validation techniques, including parameter variability-sensitivity analysis (Sargent 2013). Various sets of experimental conditions from the domain of the model’s intended applicability should be used for both comparison and exploring model behavior.

To obtain a high degree of confidence in a simulation model and its results, comparisons of the model’s and system’s output behaviors for several different sets of experimental conditions are usually required. Thus if a system is not observable, which is often the case, it is usually not possible to obtain a high degree of confidence in the model. In this situation the model output behavior(s) should be explored as thoroughly as possible and comparisons made to other valid models whenever possible.

Table 1: Operational Validity Classification.

<table>
<thead>
<tr>
<th>Decision Approach</th>
<th>Observable System</th>
<th>Non-observable System</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subjective Approach</td>
<td>• Comparison Using Graphical Displays</td>
<td>• Explore Model Behavior</td>
</tr>
<tr>
<td></td>
<td>• Explore Model Behavior</td>
<td>• Comparison to Other Models</td>
</tr>
<tr>
<td>Objective Approach</td>
<td>• Comparison Using Statistical Tests and Procedures</td>
<td>• Comparison to Other Models Using Statistical Tests</td>
</tr>
</tbody>
</table>

5.3.1 Explore Model Behavior

The simulation model output behavior can be explored either qualitatively or quantitatively. In qualitative analysis the directions of the output behaviors are examined and also possibly whether the magnitudes are “reasonable.” In quantitative analysis both the directions and the precise magnitudes of the output behaviors are examined. Experts on the system often know the directions and frequently know the “general values” of the magnitudes of the output behaviors. Many of the validation techniques can be used for model exploration. Parameter variability-sensitivity analysis should usually be used. Graphs of the output data can be used to display the simulation model output behavior (See, e.g., Sargent 1996, 2013). A variety of statistical approaches can be used in performing model exploration including metamodeling and design of experiments (See, e.g., Kleijnen 2008). Numerous sets of experimental frames should be used in performing model exploration.

5.3.2 Comparisons of Output Behaviors

There are three basic approaches used in comparing the simulation model output behavior to either the system output behavior or another model output behavior: (1) the use of graphs to make a subjective decision, (2) the use of confidence intervals to make an objective decision, and (3) the use of hypothesis tests to make an objective decision. (See Sargent (2014) for a recently developed method of hypothesis testing that uses intervals.) It is preferable to use confidence intervals or hypothesis tests for the comparisons because these allow for objective decisions. However, it is often not possible in practice to use either one of these two approaches because (a) the statistical assumptions required cannot be satisfied or only with great difficulty (assumptions usually required are data independence and normality) and/or (b) there is an insufficient quantity of system data available, which causes the statistical results to be
Sargent

“meaningless” (e.g., the length of a confidence interval developed in the comparison of the system and simulation model means is too large for any practical usefulness). As a result, the use of graphs is the most commonly used approach for operational validity. Extreme care must be used in using this approach. Each of these three approaches is discussed further in Sargent (2013).

6 DOCUMENTATION

Documentation on model verification and validation is usually critical in convincing users of the ‘correctness’ of a model and its results. This documentation should be included in the simulation model documentation and include both detailed and summary documentation. The detailed documentation should include specifics on tests used, evaluations made, data, results, etc. The summary documentation should include a separate evaluation table for data validity, conceptual model validity, computer model verification, operational validity, and an overall summary table. The summary results should contain the confidence the evaluators have in the results and conclusions which are often expressed as low, medium, and high. Examples of different tables are contained in Sargent (1991, 1996, 2013).

7 RECOMMENDED PROCEDURE

This author recommends that the following eight steps be performed in model verification and validation:

1. An agreement be made prior to developing the model between (a) the model development team and (b) the model sponsors and (if possible) the users that specifies the decision-making approach and a minimum set of specific validation techniques to be used in determining model validity.
2. Specify the acceptable range of accuracy required of the simulation model’s output variables of interest for the model’s intended application prior to starting the development of the model or very early in the model development process.
3. Test, wherever possible, the assumptions and theories underlying the simulation model.
4. In each model iteration, perform at least face validity on the conceptual model.
5. In each model iteration, at least explore the simulation model’s behavior using the computerized model.
6. In at least the last model iteration, make comparisons, if possible, between the simulation model and system behavior (output) data for at least a few sets of experimental conditions, and preferably for several sets.
7. Prepare the verification and validation documentation for inclusion in the simulation model documentation.
8. If the simulation model is to be used over a period of time, develop a schedule for periodic review of the model’s validity.

Some simulation models are developed for repeated use. A procedure for reviewing the validity of these models over their life cycles needs to be developed, as specified in Step 8. No general procedure can be given because each situation is different. For example, if no data were available on the system when a simulation model was initially developed and validated, then revalidation of the model should take place prior to each usage of the model if new data or system understanding has occurred since the last validation.

8 SUMMARY

Model verification and validation are critical in the development of a simulation model as a model and its results need to be “correct.” Every simulation project presents a new and unique challenge regarding model verification and validation. This introduction to verification and validation presented what it is, why it is important, how it relates to the model development process through the use of a graphical
paradigm and a flow chart, who the decision-maker is, comments on the importance of model accuracy and documentation, an overview on how to conduct verification and validation, and a procedure for performing it.

There is considerable literature on model verification and validation. There are conference tutorials and papers (e.g., Sargent 2011, 2014a), journal articles (e.g., Gass 1983, Landry, Malouin, and Oral 1983, Sargent 2013), discussions in textbooks (e.g., Banks et al. 2010; Law 2014; Robinson 2004; Zeigler, Praehofer, and Kim 2000), U.S.A. Government Reports (e.g., DoDI 5000.61 (2009) and U.S. General Accounting Office 1987), and books (Knepell and Arangno 1993, Oberkampf and Roy 2010) that can be used to further your knowledge on model verification and validation.

Research continues on verification and validation of simulation models. This includes such topics as advisory systems (e.g., Balci 2001, Rao and Sargent 1988, Wang 2013), cost of validation (e.g., Szabo and Teo 2012), and new approaches, procedures, and techniques (e.g. Balci 2004; Balci et al. 2002; Gore and Diallo 2013, Ruess and de Moura 2003, Sargent 2014b).

REFERENCES


Sargent


Sargent


AUTHOR BIOGRAPHY

ROBERT G. SARGENT is a Professor Emeritus of Syracuse University. He received his education at The University of Michigan. Dr. Sargent has published extensively and has served his profession in numerous ways including being the General Chair of the 1977 Winter Simulation Conference (WSC), serving on the WSC Board of Directors for ten years and chairing the Board for two years, being a Department Editor for the Communications of the ACM, holding the Presidency and other offices of what is now the INFORMS Simulation Society, serving as the Founding President of the WSC Foundation, and initiating the Simulation Archive. He has received several awards and honors for his professional contributions including the INFORMS Simulation Society’s Lifetime Professional Achievement Award and their Distinguished Service Award, a WSC 40th anniversary landmark paper award, the WSC Wilson Award, the ACM SIGSIM Distinguished Contributions Award, service awards from ACM and IIE, and being elected a Fellow of INFORMS. His current research interests include model validation, performance evaluation, and the methodology areas of modeling and of discrete event simulation. Professor Sargent is listed in Who’s Who in America and in Who’s Who in the World. His e-mail is <rsargent@syr.edu>.