Proceedings of the 1997 Winter Simulation Conference
ed. S. Andradóttir, K. J. Healy, D. H. Withers, and B. L. Nelson

SIMULATION MODEL VERIFICATION AND VALIDATION: INCREASING THE USERS' CONFIDENCE

Stewart Robinson

Operations and Information Management Group
Aston Business School
Aston University
Birmingham, B4 7ET, UNITED KINGDOM

ABSTRACT

This paper sets simulation model verification and validation (V&V) in the context of the process of performing a simulation study. Various different forms of V&V need to take place depending on the stage that has been reached. Since the phases of a study are performed in an iterative manner, so too are the various forms of V&V. A number of difficulties with verifying and validating models are discussed, after which a series of V&V methods are described. V&V is seen as a process of increasing confidence in a model, and not one of demonstrating absolute accuracy.

1 INTRODUCTION

A significant element of any simulation study should be the verification and validation (V&V) of the simulation model. Without thorough V&V there are no grounds on which to place confidence in a study's results. That said, V&V is far from straightforward and is often not performed as thoroughly as it might be.

The purpose of this paper is to give an understanding of how simulation models can be verified and validated. In particular it aims to show where V&V fit within the overall modelling process and to describe various methods of V&V. First, verification and validation are defined, including various different forms of validation. Following this there is a description of the difficulties that are encountered when trying to perform V&V. The final part of the paper describes some of the more commonly used methods of V&V.

What the paper demonstrates is that there is no such thing as absolute validity. The aim of V&V is not to prove that a model is correct, since this is not possible. Indeed, the aim is to try and prove that a model is in fact incorrect. If it cannot be proved that a model is incorrect, then V&V has served to increase confidence in the model and its results.

2 WHAT IS VERIFICATION AND VALIDATION (V&V)?

Verification is the process of ensuring that the model design (conceptual model) has been transformed into a computer model with sufficient accuracy (Davis, 1992); in other words, building the model right. Validation, on the other hand is the process of ensuring that the model is sufficiently accurate for the purpose at hand (Carson, 1986); in other words, building the right model. A key concept is the idea of sufficient accuracy. No model is ever 100% accurate, indeed, there are good reasons for not having completely accurate models (Pidd, 1996). However, in V&V the aim is to ensure that the model is sufficiently accurate. Further, this accuracy is with reference to the purpose for which the model is to be used. For instance, many demonstration models are highly inaccurate, but because they are built with the sole purpose of demonstrating simulation's potential, they can still be described as valid. As a consequence, the purpose, or objectives, of a model must be known before it can be validated. This purpose may have been determined at the start of the simulation study, or may be an alternative use for an existing model.

V&V can be understood further by mapping the V&V requirements onto the process of performing a simulation study. Many authors have attempted to describe how a simulation study should be performed. There is much similarity in these descriptions. In each case a series of steps (phases, tasks, activities or processes) are outlined. A study starts by developing an understanding of the *real world* and the problem to be tackled. Following this a model is developed, first as a *conceptual model* and then as a *computer model*. Once the computer model is complete it is then used to perform experiments through which a greater understanding of the real world is sought. Data is required in order to develop and use this model. There is general agreement among authors that although there is some logical sequence to these steps, they are not necessarily performed in a strictly sequential manner and that iteration through the steps is necessary.

Figure 1, which is adapted from Sargent (1992), shows how V&V map onto the modelling process. What this shows is that each stage in the modelling process requires, in parallel, a verification or validation process. What it also shows is that there are various forms of validation, which can be defined as follows:

- *Conceptual Model Validation*: determining that the scope and level of detail of the proposed model is sufficient for the purpose at hand, and that any assumptions are correct. The question being asked is: does the conceptual model contain all the necessary details to meet the objectives of the simulation study?
- *Data Validation*: determining that the data required for model building, validation and experimentation are sufficiently accurate.
- *White-box Validation*: determining that the constituent parts of the computer model represent the corresponding real world elements with sufficient accuracy. This is a detailed, or micro, check of the model, in which the question is asked: does each part of the model represent the real world with sufficient accuracy?
- *Black-box Validation*: determining that the overall model represents the real world with sufficient accuracy. This is an overall, or macro, check of the model's operation, in which the question is asked: does the overall model provide a sufficiently accurate representation of the real world system?

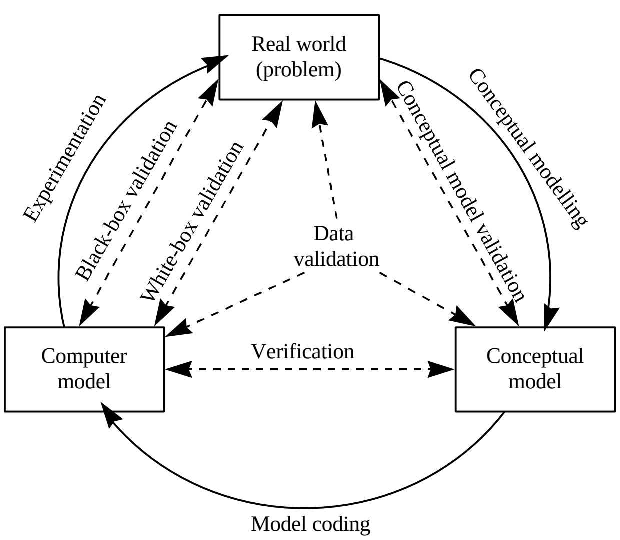


Figure 1: Simulation Model Verification and Validation in the Modelling Process

What should be apparent is that V&V is not just performed once a complete model has been developed, but that V&V is a continuous process that is performed throughout the life-cycle of a simulation study. In the same way that modelling is an iterative process, so too is V&V. It should be noted that one effect of modern Visual Interactive Modelling tools is to increase the level of iteration in simulation projects, which in turn increases the need to iterate and reiterate V&V activities.

At an early stage in a simulation project a conceptual model is developed. At this point this model should be validated. However, as the project progresses the conceptual model is likely to be revised as the understanding of the problem and modelling requirements change. As a consequence the conceptual model also needs to be revalidated. While the conceptual model is being transformed into a computer model, the constituent parts of the model (particularly those recently coded) should be continuously verified. Similarly, the details of the model should be checked against the real world throughout model coding (white-box validation). Black-box validation is the only process that requires a completed model, since it makes little sense to compare the overall model against the real world until it is complete. However, the identification of model errors and continued changes to the conceptual model (through changes to the project's requirements) necessitates model revisions and therefore further black-box validation. Although white-box validation and black-box validation are often lumped together under one heading, it is because they are performed as separate activities during the modelling process that a distinction is drawn between them here.

3 THE DIFFICULTIES OF VALIDATION

Before discussing specific methods of V&V it is important to recognise that there are a number of problems that arise in trying to validate a model.

3.1 There is No Such Thing as General Validity

A model is only validated with respect to its purpose. It cannot be assumed that a model that is valid for one purpose is also valid for another. For instance, a model of a production facility may have been validated for use in testing alternative production schedules, however, this does not mean that it is necessarily valid for determining that facility's throughput. A model could only be described as generally valid if it could be demonstrated that it was suitably accurate for every purpose to which it might ever be put. Not only is it unlikely that every potential purpose for a model could be determined, but also such a model would probably be very extensive, requiring vast amounts of code, data and run-time. This goes against the principle of keeping models as simple as possible for the task at hand (Robinson, 1994). Indeed, reality is the only 'model' which is generally valid.

3.2 There may be No Real World to Compare Against

Many models are developed of proposed real world systems, for instance, new production or service facilities. As a consequence there is no real world to use for comparison. Even if the model is of an existing system, it is likely that it is to be used to investigate alternative layouts or methods of operation, for which again no real world exists. The model may be valid when it is made to represent the existing operation, but this does not guarantee that it is valid once it is representing some change to that system.

3.3 Which Real World?

Different people have different interpretations of the real world, known as a world view or Weltanschauung (Checkland, 1981). An employee in a bank may see the bank as a means for earning money, while a customer may see it as a means for safely storing money, or as a means for borrowing money. Depending on who we speak to, we obtain different interpretations of the purpose and operation of the bank. Every day we can read multiple accounts of the same event in our newspapers, each with subtle (or not so subtle!) differences. The event was the same, but the reporters' interpretations vary.

This presents a problem when validating models. If people have different interpretations of the real world, which interpretation(s) should be used for developing and validating a model? A model that is valid to one person may not be valid to another.

3.4 Often the Real World Data are Inaccurate

Validation often involves a comparison of some facet of the model, for instance throughput, against real world data. The model is run under the same conditions as the real world to see if it performs in a similar manner. There are two difficulties that arise with this procedure. First, the real world data may not be accurate. Indeed, the purpose of data validation is to determine the accuracy of the data that is being used. If the data are not accurate, however, this creates problems in determining whether a model's results are correct or not.

Second, even if 'accurate' real world data do exist it must be remembered that these are only a sample, which in itself creates inaccuracy. For instance, data may have been collected on the throughput of a production facility over a ten week period. If, however, data had been collected for a further ten weeks, this would no doubt have changed the mean and variance of the distribution. To exacerbate the problem, the simulation itself is providing only a sample - results of say ten weeks of operation. This means that the real world-to-model comparison is a comparison of two samples. Although statistical procedures can be used to determine whether these two samples are similar, these only provide a probabilistic and not a definitive answer.

3.5 There is Not Enough Time to Verify and Validate Everything

There is simply not enough time to verify and validate every aspect of a model. Those that develop software have experienced users breaking what was thought to be perfectly sound code. This is a problem that affects both verification and validation. The modeller's job is to ensure that as much of the model is verified and validated as possible, both in terms of the model details (conceptual model validity, verification, white-box validation and data validation) and the overall validity (black-box validation).

The conclusion of this is that it is not possible to talk of absolute model validity, but only of confidence in a model. The process of V&V is not one of trying to demonstrate that the model is correct, but in fact one of trying to prove that the model is incorrect. The more tests that are performed in which it cannot be proved that the model is incorrect, the more the user's (and the modeller's) confidence in the model grows. This is the very purpose of V&V, to increase confidence in the model and its results.

4 METHODS OF VERIFICATION AND VALIDATION

There are many methods of V&V available to simulation modellers. Here a summary of some of the more common techniques is provided. For a detailed review of V&V techniques see Balci (1994).

4.1 Conceptual Model Validation

In order to develop a conceptual model it is necessary for the modeller to acquire an in-depth understanding of the real world system and the problem to be tackled. This requires a great deal of interaction with those who have a detailed knowledge of the system. In this way the modeller is able to obtain a 'rich picture' consisting of the different interpretations of the real world, or Weltanschauung (Checkland, 1981). Then the modeller uses this information to develop a set of modelling objectives and a conceptual model. This conceptual model makes explicit those aspects of the real world system that are to be included in the model (and those that are to be excluded), and is expressed either formally (e.g. Activity Cycle Diagrams (Paul and Balmer, 1993)) or informally (e.g. a list of elements and assumptions).

In order to develop a conceptual model the modeller must interpret the information that is obtained and develop a consensus of opinion. Obviously there is much room for misinterpretation during this process. Therefore, it is important that the conceptual model, once developed, is validated.

There are no formal methods for validating a conceptual model. However, a project specification, or terms of reference, is a useful device (Robinson, 1994). This should outline the objectives of the project and the modelling approach (whether formally or informally expressed). The document can then be circulated among those who have a detailed knowledge of the system and feedback sought on whether it is thought appropriate. Presenting the information may be more effective and enables immediate feedback. Hopefully any errors can be identified and changes incorporated in a new project specification. Throughout the modeller continues to be involved in a process of interpretation and developing consensus among different world views (Weltanschauung).

4.2 Data Validation

Data are extracted from the real world for input to the conceptual model and the computer model, and are also used for model validation and experimentation. Obviously inaccurate data could be a significant source of inaccuracy in any simulation model. It is therefore important that strenuous effort is made to ensure that the data are as accurate as possible. The modeller should explore the source of any data to determine their reliability. Also, the data should be analysed for inconsistencies and any cause for concern investigated. Beyond this, much has to be put down to trust, especially when the modeller is external to the organisation that is responsible for the provision of data.

In any simulation study it is likely that some data are inaccurate, in the wrong format or totally unavailable. In these situations procedures need to be put in place for collecting or estimating these data and for performing sensitivity analysis to ascertain the effects of any inaccuracies (Robinson, 1994). It is also useful to store data separately from the simulation code, in data files or spreadsheets for instance, to facilitate easy identification of errors and updates as more accurate data become available.

The modeller should be aware of the way in which the software interprets the data that are input to it. This is of particular importance when a package is being employed. For instance, a package may ask for information on time between failures. The modeller needs to be aware of whether or not this is inclusive of the repair time.

4.3 Verification and White-box Validation

Although verification and white-box validation are conceptually different, they are treated together here because they are both performed continuously throughout model coding. Also, they are both micro checks of the model's content. Verification ensures that the model is true to the conceptual model, while white-box validation ensures that the content of the model is true to the real world (in this way it is an indirect form of conceptual model validation).

Various aspects of the model should be checked during model coding:

- timings e.g. cycle times, repair times and travel times
- control of elements e.g. breakdown frequency and shift patterns
- control of flows e.g. routing
- control logic e.g. scheduling and stock replenishment
- distribution sampling e.g. the samples obtained from an empirical distribution

Three methods of verification and white-box validation are now discussed.

4.3.1 Checking the Code

The modeller needs to read through the code to ensure that the right data and logic have been entered. This is especially true for areas of complex logic. A useful idea is to get someone else to read the code as a second check. If no modelling experts are available then most simulation software vendors offer a help desk service with which specific areas of code could be discussed. Alternatively, by expressing the code in a non-technical format (the documentation could be used for this purpose) a non-expert could check the data and the logic. This is especially useful for obtaining the opinion of those who have a detailed knowledge of the system being modelled.

4.3.2 Visual Checks

The visual display of the model proves to be a powerful aid for V&V. By running the model and watching how each element behaves both the logic of the model and the behaviour against the real world can be checked. Various ideas aid this approach:

- stepping through the model event by event
- stopping the model, predicting what will happen next, running the model on and checking what happens
- interactively setting up conditions to force certain events to take place
- isolating areas of the model so it runs faster, reducing the time to perform thorough V&V
- explaining the model to those knowledgeable about the real system in order to gain their opinion
- tracing the progress of an item through the model

It is useful simply to watch a model running for period of time. In so doing a lot can be learnt about the behaviour of the model. It is also useful to demonstrate the model, formally and informally, to those who have a detailed knowledge of the system. Not only does this enable them to identify any shortcomings in the model, but by involving them this should increase the credibility of the work (assuming that not too many errors are found!).

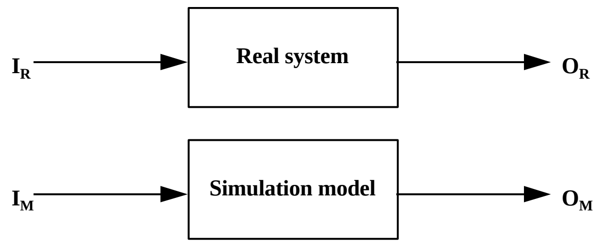
4.3.3 Inspecting Output Reports

By inspecting the reports from a simulation run, the actual and expected results can be compared. Of particular interest at this stage is the performance of individual elements, for example, their utilisation or work-in-progress. Graphical reports of samples from input distributions, for instance machine repair times, are an aid in checking that they are being modelled correctly.

A report which may be of some use is a "trace" of a simulation run. This is a blow-by-blow history, written to a text file, of every event which took place during the run. Inspecting this report can help to diagnose and rectify any problems.

4.4 Black-box Validation

In black-box validation the overall behaviour of the model is being considered. If confidence is to be placed in the model then when it is run under the same conditions (inputs) as the real world system, the outputs should be sufficiently similar (figure 2). As already stated the significant difficulty with this form of validation is that there may not, for whatever reason, be any accurate real world data with which to perform such a comparison. If this is the case then comparison can be made against the expectations and intuition of those who have a detailed knowledge of the real system, or against other models. Comparison against approximate real world data such as these may not give absolute confidence in the model, but it should help to increase confidence.



I_R - inputs to real system
O_R - outputs from real system
I_M - inputs to simulation model
O_M - outputs from simulation model

H_0: If $I_M = I_R$ then O_M O_R

Figure 2: Black-box Validation

The methods of black-box validation described here can be split into two categories. First, comparison against the real system, whether that be actual or expected, and second, comparison with other models. When performing this validation it must be remembered that both the real world data and the model results are only samples. Confidence in the model results can be improved by following proper experimental procedures, especially by performing multiple replications (Robinson, 1994).

4.4.1 Comparison with the Real System

Historic (or expected) data collected from the real system such as throughput and customer service levels can be compared to the results of the simulation when it is run under the same conditions. It is important to check not only the average levels of these data but also to compare their spread. This can be performed by judging how closely the averages from the model and the real world match, and by visually comparing the distributions of the data. Various statistical tests also lend themselves to such comparison (Law and Kelton, 1991; Kleijnen, 1995).

An alternative approach is to compare the relationships between inputs and outputs in the model and the real world. For instance, if it is known that when an input (e.g. a storage area) is increased by 20% in the real world there is a corresponding 10% increase in one of the outputs (e.g. throughput), a similar relationship should be obtained from the model.

In a Turing Test (Schruben, 1980) the model reports are made to look exactly the same as the reports provided by the real system. One or more reports from the model and from the real world are given to someone who is knowledgeable about the system. He/she is then asked to try and distinguish between the two. If he/she is unable to detect any difference this is indicative of model validity and increases confidence in the model. Even if real world reports are not available it is still worth asking an expert to review the model reports.

4.4.2 Comparison with Other Models

This group of methods is particularly useful when no real system data are available. However, this does not preclude their use when data are available. Indeed, using these in addition to real world comparison can only serve to increase confidence.

One approach is to compare the simulation model against a mathematical model. It is unlikely that a mathematical model is able to predict the outcome of the simulation exactly, otherwise the simulation would probably not have been built. However, for the purposes of comparison a mathematical model may be able to give a crude approximation of the outputs of the real system. Examples of mathematical models that might be used are paper calculations, spreadsheet analysis and queuing theory.

In order to aid comparison it is sometimes useful to simplify the simulation model to the extent that a mathematical model can predict exactly, or at least more exactly, the outcome of the model. One specific, and extreme, case of this is the use of deterministic models. This is a simulation model from which all the random events have been removed. In many cases it is possible to determine mathematically the exact outcome of such a model.

Comparisons can also be made against other simulation models of the same or similar systems. For instance, a more detailed model of the system may have been developed for some other purpose. This presupposes, of course, that the other model is itself valid.

5 CONCLUSION

It is not possible to prove that a model is absolutely correct. Therefore, model V&V is concerned with creating enough confidence in a model for the results to be accepted. This is done by trying to prove that the model is incorrect. The more tests that are performed in which it cannot be proved that the model is incorrect, the more confidence in the model is increased.

Of course, the modeller and the users may have different thresholds for confidence. Some users may derive their confidence simply from the model's display, others may require more in-depth V&V before they are willing to believe the results. The modeller is responsible for guiding the users and ensuring that sufficient V&V is performed.

This paper describes various forms of V&V as well as some V&V methods that should be employed throughout the life-cycle of a simulation study. V&V should be carried out in an iterative manner to reflect the iterative nature of the processes in a simulation project. By performing as much V&V as possible during a simulation study, the users should gain sufficient confidence in the model to accept the results. To conclude, the general rule for V&V is: the more the better!

REFERENCES

Balci, O. 1994. Validation, Verification, and Testing Techniques Throughout the Life Cycle of a Simulation Study. *Annals of Operations Research* 53: 121-173.

Carson, J. S. 1986. Convincing Users of Model's Validity is Challenging Aspect of Modeler's Job. *Industrial Engineering* 18 (6): 74-85.

Checkland, P. 1981. *Systems Thinking, Systems Practice.* Chichester, UK: Wiley.

Davis, P. K. 1992. Generalizing Concepts of Verification, Validation and Accreditation (VV&A) for Military Simulation. R-4249-ACQ, October 1992, RAND, Santa Monica, CA.

Kleijnen, J. P. C. 1995. Verification and Validation of Simulation Models. *European Journal of Operational Research* 82 (1): 145-162.

Law, A. M., and W. D. Kelton. 1991. *Simulation Modeling and Analysis*. 2d ed. New York: McGraw-Hill.

Paul, R., and D. Balmer. 1993. *Simulation Modelling.* Bromley, UK: Chartwell Bratt.

Pidd, M. 1996. *Tools for Thinking: Modelling in Management Science.* Chichester, UK: Wiley.

Robinson, S. 1994. *Successful Simulation: A Practical Approach to Simulation Projects.* Maidenhead, UK: McGraw-Hill.

Sargent, R. G. 1992. Validation and Verification of Simulation Models. In *Proceedings of the 1992 Winter Simulation Conference,* ed. J. J. Swain, D. Goldsman, R. C. Crain, and J. R. Wilson, 104-114. Institute of Electrical and Electronics Engineers, Piscataway, NJ.

Schruben, L. W. 1980. Establishing the Credibility of Simulations. *Simulation* 34 (3): 101-105.

AUTHOR BIOGRAPHY

STEWART ROBINSON lectures in Operations and Information Management at the Aston Business School in the UK. He was previously employed in simulation consultancy, offering simulation services to companies throughout Europe and the rest of the world. He is author of the book *Successful Simulation* (McGraw-Hill), a practical guide to simulation projects. His research interests are in finding ways to improve the use of simulation within industry and he has recently received an award from the United Kingdom Operational Research Society for his work. He is also investigating the use of simulation for modelling rail yard operations.