Proceedings of the 2000 Winter Simulation Conference
J. A. Joines, R. R. Barton, K. Kang, and P. A. Fishwick, eds.

VERIFICATION, VALIDATION, AND ACCREDITATION IN THE LIFE CYCLE OF MODELS AND SIMULATIONS

Jennifer Chew

HQ, U.S. Army Developmental Test Command
ATTN: CSTE-DTC-TT-M
Aberdeen Proving Ground, MD 21005-5055, U.S.A.

Cindy Sullivan

U.S. Army Yuma Proving Ground
ATTN: CSTE-DTC-YP-CD, Building 2105
Yuma, AZ 85365, U.S.A.

ABSTRACT

Verification, validation, and accreditation (VV&A) activities should be an on-going process throughout the life cycle of models and simulations (M&S). It is important to note that there is no single set of VV&A tasks, events, or methods that would apply every time to every situation. The VV&A emphasis and methods used vary depending on the particular life cycle phase it is in, previous VV&A and use, the risks and uncertainty, its size and complexity, and of course, the resources available. For simplification, this paper discusses the activities and tasks during the early stages of model development and addresses each of the VV&A efforts separately, along with its associated activities. It outlines the specific VV&A activities and products that are appropriate to each phase of model development.

1 INTRODUCTION

In recent years, the Department of Defense (DoD) has aggressively applied M&S in wargaming, analysis, design, testing, etc., to support acquisition decisions. One caveat is that if the model is intended to be used by DoD, then the model must be verified and validated to ensure that the simulation outputs are sufficiently credible for its intended use(s). While the DoD is responsible for its own M&S, M&S that are developed and/or used by industry and academia in support of DoD acquisition activities must also comply with the DoD VV&A policy. The information presented herein has been compiled from a wide variety of sources, including DoD directives and instructions related to M&S management and VV&A, software industry standards and practices, and academic text and professional literature.

The VV&A activities contained herein are broadly applicable to all stand-alone models and federates which are used for supporting DoD acquisition decisions. Federates are individual M&S products that are capable of joining High Level Architecture—based federations. This paper does not cover the VV&A on a federation of models. VV&A of a federation must be completed after doing VV&A on each of its federates. The activities described in this paper are intended to be used for planning, producing, and documenting proper evidence to support the VV&A of M&S. This paper is also intended to help the reader to plan for and develop structured and organized VV&A activities; provide a systematic approach for preparing VV&A documentation; and give a better understanding of how VV&A can be an integral part of the M&S life cycle. It emphasizes activities that are crucial during each phase of M&S development and use.

Too often Verification and Validation (V&V) are considered separately from development and documentation. The V&V plans and process should begin on the first day of development and continue in such a manner that the same documentation used for requirements, design, development, and configuration control also serves to support V&V activities. Finding and resolving problems early via application of V&V can significantly reduce the subsequent cost of M&S design, development, and testing. There are many V&V tasks that the M&S developer should be doing before and during model development. As a matter of fact, VV&A activities should begin as soon as there is a decision to apply M&S to a problem. The planning effort for VV&A is as important as implementing it. The earlier we start the V&V planning, the easier it is to implement. It is always good practice to ensure that all pertinent information is documented along the way.

It is important to note that all the VV&A activities are tailorable to the specific requirements. Unless there is high impact given a failure (e.g., cost or safety) or it is a very large and/or complex developmental effort, we probably do not need to accomplish every task or method mentioned in this paper. There is no single set of VV&A tasks, events, or methods that applies exclusively every time to every situation. VV&A emphasis and methods used vary depending on the particular life cycle phase it is in, previous VV&A and use, the risks and uncertainty, its size and complexity, and resources available. The depth of analysis involved with the V&V of an established legacy model would be different from the development of a new M&S. Likewise, the available information for the accreditation of legacy model might be based more on historical performance than results from the detailed tasks outlined in this paper for a new M&S.

There are many ways and techniques to accomplish VV&A. Although there is an abundance of literature on VV&A advocating diverse methods, this paper compresses the information to provide a simplified process that focuses on the activities and tasks during each phase of the development. For simplification, this paper addresses the VV&A activities and products that apply to each M&S development phase.

2 VV&A IN THE LIFE CYCLE OF M&S

Figure 1 shows a typical life cycle of an M&S and its associated VV&A activities. These activities or tasks may be tailored and applied differently based on the depth of analysis, as required by the user or established acceptability criteria. The authoritative data source (ADS) library, as shown in Figure 1, contains DoD data sources used for supporting M&S which are cataloged through the M&S Resource Repository (MSRR). The ADS library is available through the Defense Modeling and Simulation Office website at <www.dmso.mil>.

The remainder of this paper examines each of the VV&A phase and discusses the activities associated with them.

2.1 Requirements Verification and Validation

The M&S development should begin with a clear and unambiguous statement of the problem that the M&S are intended to address. A good definition of the problem makes it easier to define M&S requirements such as simulation outputs, functions, and interactions. It is also important to specify, at least in general terms, how much like the real world the user needs these outputs, functions, and interactions to be. We believe that the most critical piece of the M&S development and V&V activities falls in the very beginning of the life cycle. If the requirements do not make sense or not well understood, then the M&S will not do what was originally intended.

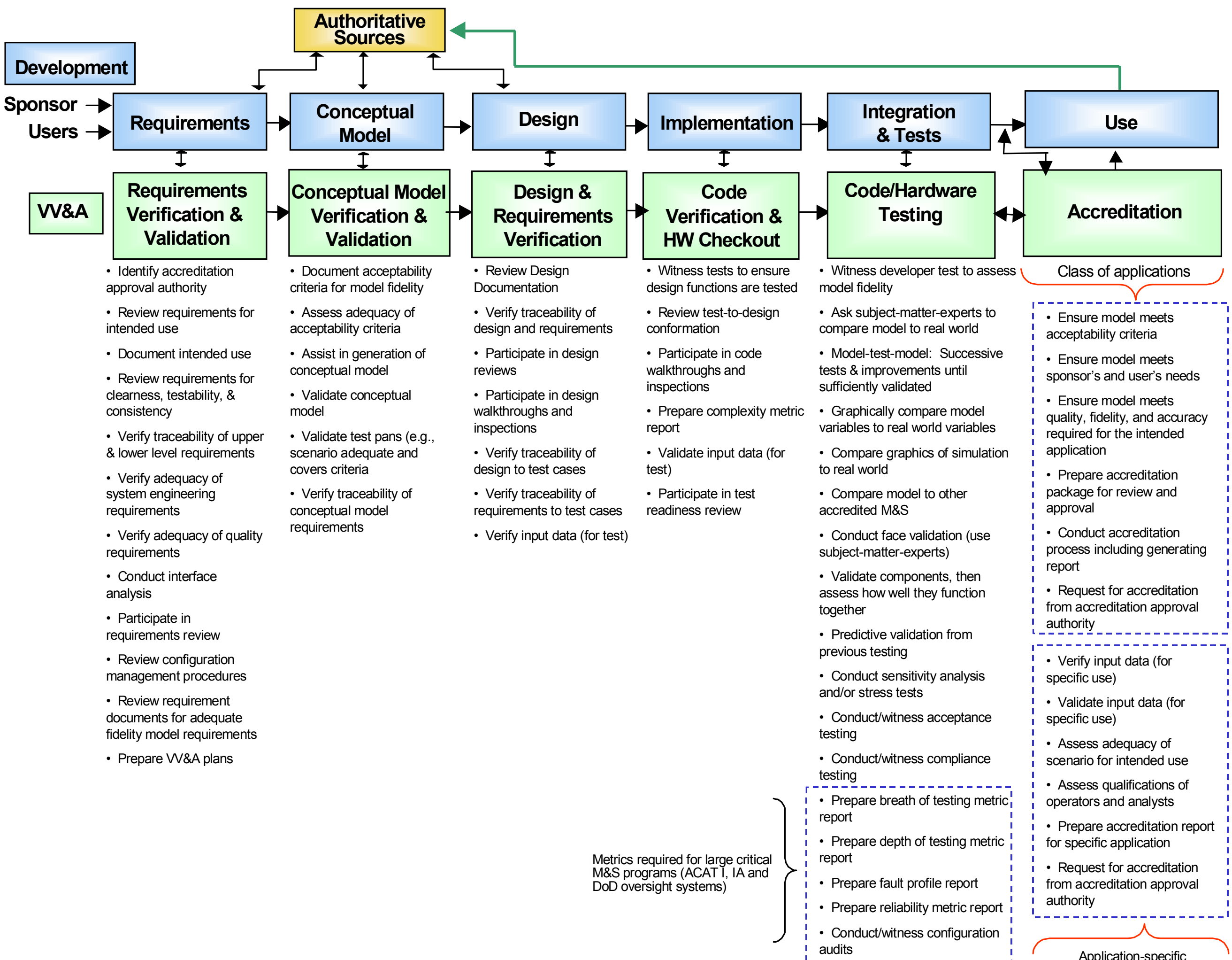

Figure 1: VV&A in the Life Cycle of M&S

Basically, this phase of the process is primarily involved in reviewing the requirement documentation and in documenting all findings. The review focuses on the intended use, acceptability criteria for model fidelity, traceability, quality, configuration management, and fidelity of the M&S to be developed. This is done to ensure that all the requirements are clearly defined, consistent, testable, and complete.

The first step is to gather information. Any information related to the M&S and its requirements increases understandability of the requirements and making the right decisions. It may not be obvious that one of the most critical V&V effort is to review all the information gathered and document all the findings. This could include:

- Requirements
- Interface requirements
- Developmental plans
- Previous V&V plans and results
- Configuration Management Plan
- Quality Assurance Plans
- Studies and Analyses

Documenting all the findings, assumptions, limitations, etc., from reviewing every piece of related information about the M&S, is extremely important. We review the requirement documentation, determine the risk areas, and assess the criticality of specific factors that need the most attention. Again, we document the assessment and highlight the areas that may need further analysis. We report all the findings to the sponsor/user and have all the discrepancies resolved before continuing with any further major efforts.

The following should be considered when tailoring. If the intended use is not adequately documented, the V&V team may need to talk to the users and document the intended use themselves. If the model has interfaces, these need to be verified to determine if the interface structure is adequate. User interfaces need to be analyzed to determine how accurately the interface is integrated into the overall M&S and for human factors engineering, for example, requirements to accommodate the number, skill levels, duty cycles, training needs, or other information about the personnel who will use or support the model. If this is a developmental effort or the developers are available, the V&V team may be able to participate in requirements review and ask the developers questions face-to-face. The following system engineering factors may be important to assess for adequacy:

- adaptation of installation independent data
- safety (prevent/minimize hazards to personnel, property, and physical environment)
- security and privacy
- for software, the computer hardware and operating system
- for hardware, the environment during transportation, storage, and operation, e.g., wind, rain, temperature, geographical location, motion, shock, noise, and electromagnetic radiation
- computer resources used by the software or incorporated into the hardware
- design and construction constraints
- logistics
- packaging

The requirements V&V phase culminates with the documentation of the intended use, requirements traceability matrix, unsupported requirements, acceptability criteria for model fidelity, risk assessment, and model fidelity.

2.2 Conceptual Model Verification & Validation

A conceptual model is a preliminary or proposed design framework that is based on the outputs, functions, and interactions defined during the requirements V&V described in Section 2.1. A conceptual model typically consists of a description of how the M&S requirements are broken down into component pieces, how those pieces fit together and interact, and how they work together to meet the requirements specified. It should also include a description of the equations and algorithms that are used to meet the requirements, as well as an explicit description of any assumptions or limitations made or associated with the theories, concepts, fidelity, derivatives, logic, interfaces, or solution approaches. The process of determining the adequacy of the conceptual model and ensuring that it meets the specified requirements and intended use(s) is called conceptual model V&V.

One of the initial tasks for conceptual model V&V is to come to finalize and agree with the acceptability criteria for model fidelity and to define the criticality of data inputs and outputs. The importance of data is discussed in Section 2.6. Acceptability criteria and data requirements are used to ensure that each step of the conceptual model framework is traceable to the requirements, and ultimately to these criteria. These criteria are established by the accreditation approval authority defining the terms and conditions of the M&S that will be considered acceptable for the application. Therefore, a set of test cases must be defined to ensure that all the simulation scenarios and trials will adequately address the requirements and satisfy the acceptability criteria. It is crucial that we verify and validate the conceptual model adequately from which the code is generated and/or hardware is built.

The products of conceptual model V&V are model characteristics, input/output data items, interface issues, measure of model fidelity, potential weaknesses and limitations, perceived strengths, and traceability between conceptual model and requirements.

2.3 Design Verification

After the conceptual model is verified and validated, the developer produces a detailed design that describes exactly how the conceptual model will be coded or fabricated. It defines the components, elements, functions, and specifications that will be used to produce the simulation based on the conceptual model. Before a single line of software code is written or hardware is fabricated, we should review the detailed design to ensure it conforms to the conceptual model. This step is called Design Verification. It involves a mapping of the proposed design elements back to the conceptual model and requirements to ensure that there is traceability between those requirements and the proposed design. We should also develop test cases that can be traced back to the design and requirements.

Although traceability is the main focus during the design verification, other activities such as participating in design reviews, audits, walkthroughs, and inspections are important. For software, it is also important to verify input data; determine computer-aided software engineering tools and design methodology; conduct internal software testing; and perform software metrics analysis. For hardware, it is important for subject matter experts to review the adequacy of drawings (e.g., schematic drawings), interface control drawings, and, as appropriate, the adequacy of the electrical design, mechanical design, power generation and grounding, electrical and mechanical interface compatibility, and mass properties.

This phase culminates with the traceability matrix (detailed design to requirements, to conceptual model, and to test cases), design and requirement cross reference matrix, design walkthrough or inspection report, input data verification, software metric and test reports, and CASE tools.

2.4 Code Verification and Hardware Checkout

After the design is verified, the conceptual model and its associated design are converted into code or hardware by the developer. Code verification and hardware checkout ensure that the detailed design is being implemented correctly in the code or hardware respectively.

Code verification normally entails detailed desk checking and software testing of the code, comparing it to the detailed design, documenting any discrepancies and fixing any problems discovered. Other important activities include participating in code testing, audits, walkthroughs, and inspections; validating input data; preparing complexity report; conducting code analysis; and verifying code structure.

Hardware checkout entails reviews, audits and inspections, comparing the hardware to its design, documenting any discrepancies and fixing any problems.

This phase culminates with the design functionality, code walkthrough or inspection report, complexity metric report, input data validation, coding/interface/logic errors, and syntax and semantics.

2.5 Code and/or Hardware Testing

After the design and the initial implementation are completed, the developer integrates the code and/or hardware together and tests it. These tests are intended to verify and validate the M&S. Verification tests the correctness of the M&S to ensure that it accurately represents the developer's requirements, conceptual description, and design. Validation tests the extent to which an M&S accurately represents the real world from the perspective of the intended use of the M&S.

Verification tests that the M&S requirement, conceptual model and design are implemented as documented in the previous phases. Acceptance testing determines whether all requirements are satisfied. Compliance testing determines if the simulation meets required security and performance standards. Test cases should be traceable to the documented requirements and design to ensure that all were met. Metrics that may be used, if this is a large software development, include breadth and depth of testing, fault profiles, and reliability metrics. The breadth of testing metric (% requirements tested and % requirements addressed) address the degree to which required functionality has been successfully demonstrated as well as the amount of testing that has been performed. The depth of testing metric (% tested and passed testing) measures the amount of testing achieved on the software architecture, that is, the extent and success of testing the possible control and data paths and conditions within the software. Automated tools may be used to compute this measure. Fault profiles (open versus closed anomalies) provides insight into the number and type of deficiencies in the current baseline, as well, as the developer's ability to fix known faults. The reliability metric (mean time between failures) expresses the contribution to reliability.

The two issues that must be addressed during validation testing are to identify the real world being modeled and to identify the key structural characteristics and output parameters that are to be used for comparisons. In other words, validation has to do with the fidelity of the M&S. Fidelity is normally defined by the sponsor/user and is judged by several factors, one of which is its ability to predict the known behavior, or best estimate, of the real system when subjected to the same stimuli. The fidelity level is actually defined when the sponsor/user establishes the acceptability criteria for model fidelity. If the M&S is designed with these criteria in mind, then very likely the M&S will fall within the defined fidelity boundary and be acceptable by the sponsor/user. Otherwise, there is a chance of going back to the drawing board. Defining the acceptability criteria up-front is crucially important.

In those cases where there is no user or the user simply cannot come up with a set of criteria, we should make sure that all pertinent information about the M&S and the assumptions are documented every step of the way. As a user, validation, by far, is the most important phase of the M&S life cycle. Validation gives solid evidence to help analyze the extent to which the M&S are representing the real world. It is also critical that we assess the degree of detail that must be represented in the simulation to provide acceptable results and the degree of correspondence with real world phenomena that will be sufficient for use with high confidence. If the significant parameters of a real system have been properly incorporated into a model, a simulation experiment should reflect the behavior of a real system down to some level of detail commensurate with that description.

Many validation techniques such as using subject matter experts, comparison techniques, and face validation to just name a few. Validation based upon direct comparison of model results to the real world provides more credibility than other validation methods. Selection of techniques is based on the user's needs, M&S types, intended uses, and other factors.

Despite of the techniques used, the following products should be generated as part of the testing: model fidelity assessment; traceability between requirements, design, and test cases; subject matter expert opinions; M&S and real world comparison; model limitation and impact statement; sensitivity analysis report; test results; and metric report.

2.6 Accreditation

Accreditation is the official determination by the user that the capabilities of the M&S fit the intended use and that the limitations of the M&S will not interfere in drawing the correct conclusions. Accreditation planning should not wait until after the development is completed. It should begin when the requirements were being verified and validated because the first task, when preparing the accreditation plan, is to develop the acceptability criteria. Acceptability criteria established in the accreditation plan are what the user has identified as key characteristics for use in deciding whether or not to grant an accreditation for the particular M&S. Accreditation occurs at two levels: Class of Applications and Application-specific.

Accreditation at the Class of Applications level accredits an M&S for a generic set of purposes or applications and includes reviewing a complete audit trail of the development and use of the M&S. The audit trail includes reviews of M&S documentation, V&V documentation, configuration control, M&S assumptions, previous successful uses, and recognition of users' acceptances.

Accreditation of Application-specific level M&S includes data certification, scenarios, and the qualification/ training of the operator-analysts who will use the M&S.

All M&S are driven by data, either as direct inputs or as embedded values that drive simulation characteristics. As perfect as the equations, algorithms, and software design of an M&S may be after conceptual model validation and design verification, it will probably fail results validation if the data that drive the simulation are inaccurate or inappropriate for the task at hand. A relationship clearly exists between producer data V&V activities and user data V&V requirements throughout the M&S life cycle. However, there is a distinction between data V&V activities performed by the producer and by the user. Producer data V&V determine data quality in terms of correctness, timeliness, accuracy, completeness, relevance, and accessibility that make data appropriate for the purpose intended and values are within the stated criteria and assumptions. User data V&V ensure that the data are transformed and formatted correctly and that the data meet user specified constraints. Data accreditation is an integral part of the M&S accreditation procedures to ensure that M&S data are verified as correct, and validated as appropriate and reasonable for the intended application.

3 CONCLUSIONS

VV&A may sound challenging or even impossible. This should not be the case if proper VV&A activities are conducted throughout the M&S life cycle, especially during the early stages. Early VV&A planning can reduce or even eliminate many concerns that may arise at later stages. In fact, early planning can also allow you more flexibility in selecting the right V&V techniques and activities to fit the specific needs. However, many situations exist during the M&S planning stage. For example,

- Model acceptability criteria and V&V requirements/planning must be established and agreed upon by all parties concerned before any activities are defined.
- V&V activities can be very labor-intensive and must be focused and carefully scoped according to specific accreditation requirements.
- V&V plan changes as the M&S project matures. V&V planning should not be considered final until after V&V has actually been accomplished.
- Validation depends on the intended use and fidelity of the M&S, and it will likely change as new users are identified.
- V&V should begin on day one of the M&S development, should be an integral part of the M&S development, and should be a continuous process.

- When planning for V&V activities, alternate methods should be included to facilitate schedule driven events and to adjust as new techniques are developed.
- V&V efforts require an experienced and well-trained team.

ACKNOWLEDGMENTS

The authors would like to recognize Mr. Bob Lewis, Tecmaster, Inc., for his support to the development of the VV&A activities in the Life Cycle of M&S. His significant contributions have made this paper possible.

REFERENCES

Knepell, P.L. 1999. VV&A of Models and Simulations (A Five-Day Workshop) Participant Guide. Peak Quality Services, Colorado Springs, CO.

Department of Defense. 1996. Department of Defense Verification, Validation and Accreditation (VV&A) Recommended Practices Guide. Defense Modeling and Simulation Office, Alexandria, VA. (Co-authored by: O. Balci, P.A. Glasow, P. Muessig, E. H. Page, J. Sikora, S. Solick, and S. Youngblood).

Department of the Army. Army Regulation 5-11. 1997. Management of Army Models and Simulations, Washington, DC.

U.S. Army Developmental Test Command (DTC). 1998. Developmental Test Command Verification, Validation, and Accreditation (VV&A) Methodology. DTC Pamphlet 73-4, Aberdeen Proving Ground, MD.

Department of the Army. 1999. Verification, Validation, and Accreditation of Army Models and Simulations. Pamphlet 5-11, Army Modeling and Simulation Office, Cystal City, VA.

AUTHOR BIOGRAPHIES

JENNIFER CHEW is an Electronics Engineer in the Technology Management Division, HQ U.S. Army Developmental Test Command (DTC), Aberdeen Proving Ground, Maryland. She supports the development of the DTC Virtual Proving Ground program and has the lead in developing the DTC VV&A process and methodology. She received her B.S. in Chemical Engineering from University of Maryland and M.S. in Electrical Engineering Science from Loyola College. She is a graduate of the Army Management Staff College and Quality and Reliability Engineering program. Her email address is `<chewj@dtc.army.mil>`.

CINDY L. SULLIVAN is an Operations Research Analyst and manages the Yuma Proving Ground Virtual Proving Ground Program. She received her B.S. in Computer Science from Freed-Hardeman College and M.S. in Industrial Engineering from the University of Missouri–Columbia. She has 14 years of experience working with Army M&S and earned two Army Achievement Medals. She was the primary author of DTC Pam 73-4 M&S VV&A Methodology. Her email address is `<Cindy.Sullivan@yuma-exch1.army.mil>`.