

V&V Methodology for Socio-Cultural Models

Julie F. Bouchard
 Sandia National Laboratories
 jfbouch@sandia.gov

Abstract

The V&V Methodology for Socio-Cultural Models developed by Sandia National Laboratories [1] applies a graded approach to evaluate models that include human, social, and cultural behavior. The V&V Methodology assists users in assessing model maturity, focusing on “intended use”, and selecting V&V techniques for models of different types, size, and maturity. The methodology is tailorable and scalable to assist in identifying potential V&V techniques and down-selecting to an appropriate set of techniques.

1. Introduction

Across government and industry, particularly since 9/11, federal agencies have invested heavily in computational modeling and simulation software to help analysts and decision makers better understand human social, cultural, and behavioral phenomena. Despite current investment in computational modeling and simulation, there has been less effort to develop methods for systematically evaluating the fidelity and relevance of the computational simulations for the complex problem spaces they are intended to address [2].

The Human, Social, Cultural and Behavior (HSCB) Team from Sandia National Laboratories (herein referred to as Sandia) created a generalized verification and validation (V&V) methodology to address this emerging problem area [1]. This methodology builds upon techniques employed in a variety of fields, ranging from the physical sciences and engineering to the social sciences.

The purpose of a V&V program is to provide confidence that the model and the theory and data the model is based on are suitable for the *intended use*. The definition of validation that is the basis for this V&V Methodology builds a collection of objective evidence to demonstrate the requirements for the *intended use* have been satisfied.

2. Goals

Our goals for the V&V Methodology included the following:

1. Systematic methodology that allows better model evaluations designed to improve the decision-making and cultural understanding
2. Generalizable Methodology that is applicable to a wide variety of model types and variety of human, social, and cultural behaviors
3. Practical methodology that is detailed, tailorable, and useable.

The methodology that resulted is systematic, generalizeable, and tailorable. In Section 10 Future Work, some recommendations to improve useability are identified.

3. Definitions

The definitions of verification and validation from ISO 9000 are used in this V&V Methodology. ISO 9000 defines verification as the “confirmation, through the provision of objective evidence, that specified requirements have been fulfilled.” [4] Confirmation can be made from different activities including performing alternative calculations, tests, demonstrations, comparisons, and reviewing documentation.

The V&V Methodology builds on the concise definition of validation from ISO 9000 and the approach implied by this definition. Validation is “confirmation, through the provision of *objective evidence*, that the requirements for a specific *intended use* or application have been fulfilled”. [4]

Note that the ISO 9000 definition says “confirmation” not proof. This is a critical distinction. It is possible to understand the underpinnings of a model and build confidence in the model’s abstraction of the real world and still not *prove* model correctness. The concept of correctness of a theory or model is the subject of scientific inquiry (scientific method) and is not a rigid requirement for V&V. An incorrect theory may be totally suitable for the needs of a specific model if the model only needs an approximate answer and is not sensitive to the theoretical results. A

"proven" highly accurate theory may be unsuitable for a model if the domain of applicability of the theory is significantly different than the domain of intended use of the model.

Like the ISO 9000 definition, this V&V methodology focuses on the *intended use* of the model. Focus on the intended use means focus on the applicability of the model for one or more specific problem areas identified by the end user as required. Focus on the intended use also means that validation cannot be assumed to address problem areas outside the intended use.

Also like the ISO 9000 definition, our V&V methodology is based on the systematic assembly of *objective evidence* sufficient to provide a comprehensive evaluation of a model relative to intended use. During the research for and development of a model, conscientious scientists and software engineers document their efforts and decisions. This documentation, data accumulated, and the models themselves are objective evidence. This developer-created objective evidence is commonly referred to as "artifacts" in the V&V community. Additional objective evidence results from the analysis of developer artifacts during validation activities.

4. Illumination

Historically, V&V approaches have employed a variety of focuses. One focus that we find particularly valuable is the focus on model illumination. Using V&V to illuminate (shine a light on the theory foundation, model development process, and model behavior) can provide information without being prescriptive or evaluative in approach.

Illumination of the model provides for informed decisions about the usefulness of the model for the ultimate users. Fundamental questions that are addressed include: does the model adequately support the intended uses; and does the model satisfy customer requirements. Illumination also provides information about all the different stages in the development of a socio-cultural model. Figure 1 shows a sample model development process for a socio-cultural model that combines social science research and theory development (in green) with modeling and simulation computational model development (in purple). The V&V Methodology is intended to illuminate all parts of this process.

5. Methodology Components

The V&V Methodology contains four components: selection of V&V methods, identification of the documentation artifacts, identification of relevant questions for each documentation artifact, and assessment of model maturity. These components,

described below in more detail, can be applied to a model either individually or in combination.

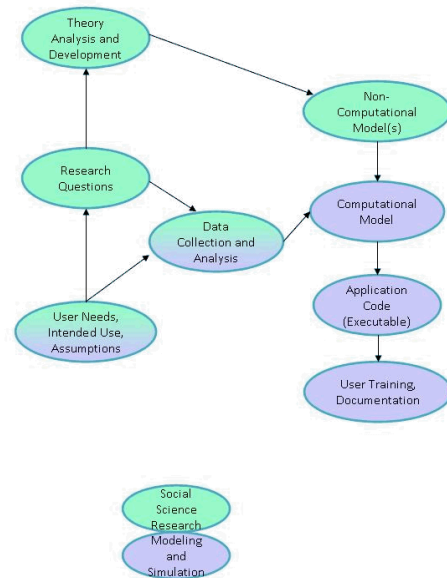


Figure 1: Example Socio-Cultural Model Development Process

5.1 V&V Methods

In this methodology, we identified approximately forty V&V methods. These methods include those previously employed in verifying or validating physical science and/or social science models and some new or research approaches that we developed into V&V methods. We identified as many methods as we could in a limited amount of time and chose not to down-select the methods so that others could understand the research approach.

These methods span the range from common techniques like Expert Review to more specialized approaches like Active Nonlinear Testing. The methods address different model development stages, from theory to the computational model, all the way to user documentation and training. Also, these methods include those that rely on artifacts from the developers as well as those that look into the code to determine the implemented behavior. Each method is documented and includes references to uses of the method.

Each stage of model development produces artifacts including documents, diagrams, or codes developed. The methodology provides a spreadsheet with artifacts in the rows and methods in the columns. See Figure 2. Methods are associated to artifacts by either solid or hollow bullets. Solid bullets indicate strongly recommended methods. Hollow bullets are generally applicable but more weakly recommended. The user

then can identify the methods associated with the model development activities and artifacts available and review the method documentation. With this information, it is much easier to select one or more methods to use for model V&V. This methodology spares development or V&V teams the significant time required to research the available methods.

Documentation related to Methods ○ possible method ● recommended	Conceptual Validation	Data Validation	Decision-Space Validation	Face Validation
Data Collection Plan		●		
Data Structure		●		○
Theory Traceability	●			
New/Modified Theory Documentation			●	●

Figure 2: Selection of V&V Methods

5.2 Documentation Artifacts

A spreadsheet approach is also used to identify documentation artifacts. Developers can use this spreadsheet to ensure the proper documentation is being developed in preparation for future V&V efforts. This spreadsheet has columns for development stages (research, technical development, and transition to operations) and for low/medium/high model risk or criticality levels.

A subset of the spreadsheet is shown in Figure 3, which shows the differences in low (L), medium (M) and high (H) risk/criticality model consequences and the methods that are appropriate. For example, a research stage model with low risk/criticality should have developed an initial theory paper addressing their research but would probably not have completed new/modified theory development or have developed use scenarios.

	Research Stage			Technical Development		
	L	M	H	L	M	H
Theory Selection and Development	●	●	●	●	●	●
				●	●	●
	●	●	●	●	●	●

Figure 3: Documentation identification

5.3 Relevant Questions

The V&V Methodology also provides questions to guide model development and V&V. These questions are focused on the specific documentation or artifact and there are five questions for each artifact. For instance, questions related to theory include the following:

- Are the relevant theories (including those researched but not selected) documented and bibliography maintained?
 - Are the selection criteria documented and applied to selection of theory?
 - Is new theory supported by documented research?
- For the computational model/executable code, the following questions are included:
- Are the design of the computational model and the data design documented?
 - Were any theoretical changes made to the social science model as a result of instantiation in a computational tool?
 - Is there traceability from model features to the intended use of the system?

5.4 Model Maturity

Determination of model maturity in this V&V methodology is a graded approach based on consequence level. There are currently four levels ranging from low consequence use for research/scoping activities to high consequence with a strong influence on high risk operations or decisions. This graded approach via consequence level is based on the Predictive Capability Maturity Model (PCCM) developed by Sandia [2]. The areas assessed goes beyond the PCCM to include areas often ignored in software development and modeling and simulation but critical to socio-cultural models (e.g., validation of statistical and behavioral modeling, data sufficiency for model use over time, and V&V of federated models).

Below are the top-level areas identified as part of this model assessment. Note that all areas may not be applicable to all models.

- Subject Matter Completeness
- Data Sufficiency for development, calibration, validation, model use
- Sub-model Validation of physical, statistical, behavioral modeling
- Federated Model Validation
- Uncertainty and Variability Quantification
- Verification
- Documenting and Archiving
- Transition and Deployment

6. Lessons Learned and Future Work

We conducted an initial test of the methodology on the computational model used for training overseas personnel about the potential impacts of their cultural decisions. The V&V approaches implemented for the test were found to be generally workable. As expected, some areas were found where improvements could be made.

The general philosophy is to illuminate potential weaknesses of the model and its backing

documentation. It is then up to the developers and users to evaluate whether these potential weaknesses need to be addressed in order to have enough confidence in the model for the way it is being used. The approaches used provide illumination of the model but may not validate its correctness. Lastly, the methods tested are subjective, so there would be some commonality if replicated by a different group, but there would also be differences.

Future possibilities have been identified for improving the V&V Methodology and extending it even further. These include, but are not limited to:

- automating the tailoring and selection processes to make them easier for the users to execute
- tailoring the V&V results for different types of users
- addressing uncertainty quantification in the methodology
- expanding the model to include federated models
- extending the method documentation to include step-by-step procedures that people with less V&V experience could execute

7. Summary

Sandia National Laboratories has developed a V&V Methodology to evaluate socio-cultural models of different types and sizes. We extended V&V concepts, methods and processes for computational and engineering sciences and integrated a wide variety of social science methods to develop a V&V Methodology that is intended for social science models. This methodology provides a structured, credible, and defensible analysis of HSCB model capability for the 'intended use'. The methodology is tailorable and scalable to assist users in identifying potential V&V techniques and down-selecting to an appropriate set of techniques.

We tested the methodology on the computational model used for training overseas personnel about the potential impacts of their cultural decisions. The goal of this test was to provide a partial assessment of the usability of the V&V methodology when applied to a social science model with a specific intended use scenario. This assessment provided the developers with valuable insights into their model, primarily because of the broad view the V&V Methodology takes of model development.

8. References

- [1] Nanco, A., E. Raybourn, A. Ames, D. Belasich, J. Bouchard, R. Heaphy, R. Hills, W. Shaneyfelt, M. Wills, M. Wilson, HSCB Generalized Validation & Verification Methodology, SAND2009-xxxx, Sandia National Laboratories, Albuquerque, NM, 2009.
- [2] Oberkampt, W.L., M. Pilch, and T.G. Trucano, Predictive Capability Maturity Model for Computational Modeling and Simulation, SAND2007-5948, Sandia National Laboratories, Albuquerque, 2007.
- [3] Zacharias, G., J. MacMillan, and S. B. Van Hemel, Committee on Organizational Modeling, Behavioral Modeling and Simulation: From Individuals to Societies, National Research Council, 2008.
- [4] ISO 9000:2000. Switzerland: International Organization for Standardization, 2000