To Main Proceedings Document

THE DOING OF MODEL VERIFICATION AND VALIDATION:

BALANCING COST AND THEORY

Geoffrey Back, Gregory Love, and Justin Falk
Project Performance Corporation, McLean, Virginia, USA

ABSTRACT

Much of the model verification and validation (V&V) guidance and literature is useful for explaining the principles of V&V and how V&V is ideally integrated into the simulation model development life cycle. There is less information available, however, on how to execute V&V, especially, as is often the case, when the resource commitment for V&V is limited. There are few examples that illustrate concrete application of the available V&V techniques or discuss the tradeoffs between theory and cost that are often made.

This paper describes the V&V approach used by Raytheon Company, C^3I Systems and Project Performance Corporation in developing several low-resolution multi-purpose simulations of integrated industrial facilities and industrial sectors for a government customer. These projects were characterized by (i) the need to deploy each simulation model within a 60- or 120-day period; (ii) the need to utilize a commercial-off-the-shelf system dynamics software application; and (iii) heavy reliance on subject matter expert input to assess real-world fidelity. Furthermore, V&V had to be performed with little guidance at the outset as to what the acceptability criteria would be and V&V budgets of no more than 8 to 10 percent of the total project cost. Consequently, while the V&V efforts conducted for these projects were built upon the "what and why" guidance outlined in the customer's policies and in such documents as the Defense Modeling and Simulation Office's *Verification, Validation, and Accreditation Recommended Practices Guide*, tradeoffs had to be made in developing an efficient "how to" approach.

April 30, 2000

Contact Information:

Geoffrey Back
Project Performance Corporation
7600 Colshire Drive, Suite 500
McLean, Virginia, 22102
USA
Email: gback@ppc.com
(703) 748-7070 (voice)
(703) 748-7001 (fax)

Gregory Love
Project Performance Corporation
7600 Colshire Drive, Suite 500
McLean, Virginia, 22102
USA
Email: glove@ppc.com
(703) 748-7055 (voice)
(703) 748-7001 (fax)

THE DOING OF MODEL VERIFICATION AND VALIDATION:

BALANCING COST AND THEORY

INTRODUCTION

Except, perhaps, for the simplest spreadsheet-based models, few model sponsors, users, and developers doubt the wisdom of investing in the verification and validation (V&V) of models and simulations. After all, the basic principles of V&V make common sense (see Table 1). However, while there is much in the literature concerning V&V principles, theory, and techniques, how much to invest in V&V and what should be done with that investment remain questions for which there seems to be very little specific guidance, especially for systems dynamics (SD) models. Therefore, these questions confronted our Raytheon Company, C^3I Systems-Project Performance Corporation team (the "Team") beginning in late 1998 as we began to develop a library of SD models for a government customer (the "Model Sponsor"). This paper traces the evolution of our V&V process for the range of SD simulation models that our Team has developed with an emphasis on (1) how, in concert with the Model Sponsor, we chose to focus a limited V&V budget and (2) the cost-efficient V&V documentation our Team devised to sustain the Model Sponsor's evolving model accreditation process.

Table 1. The Principles of Verification, Validation, and Accreditation.[1]

No.	Principles
1.	There is no such thing as an absolutely valid model.
2.	VV&A should be an integral part of the entire M&S life cycle.
3.	A well-formulated problem is essential to the acceptability and accreditation of M&S results.
4.	Credibility can be claimed only for the intended use of the model or simulation and for the prescribed conditions under which the model or simulation has been tested.
5.	M&S validation does not guarantee the credibility and acceptability of analytical results derived from the use of simulation.
6.	V&V of each sub-model or federate does not imply overall simulation or federation credibility and vice versa.
7.	Accreditation is not a binary choice.
8.	VV&A is both an art and a science, requiring creativity and insight.
9.	The success of any VV&A effort is directly affected by the analyst.
10.	VV&A must be planned and documented.
11.	V&V requires some level of independence to minimize the effects of developer bias.
12.	Successful VV&A requires data that has been verified, validated, and certified.

[1] Source: *Verification, Validation, and Accreditation Recommended Practices Guide (DMSO, 1996)*

STARTING WITH A BLANK SHEET OF PAPER

When our Team convened to develop its first SD simulation model for the Model Sponsor, a prototype or proof-of-concept SD model of a petrochemical industrial process, no monies had been earmarked for V&V; in fact, model V&V had not even been identified as a contractual requirement. Shortly thereafter, however, the Team helped the Model Sponsor realize that a prudent investment in model V&V offered a considerable potential payback. The final model was designed to represent a simplification of reality, but without the requisite confidence in its capabilities and limitations, no judgment could be made regarding its fitness for a particular purpose. Coyle and Exelby aptly make this conclusion in their observations about the concept of validity as it relates to practice of consultancy in SD (Coyle and Exelby, 2000). As a result, the Model Sponsor added a V&V requirement to its next simulation model development effort, an SD model of an integrated nitrogen fertilizer production complex. At the same time though, the Model Sponsor had yet to develop its policies and guidelines for model V&V and, on this occasion, no additional funding would be available to sustain the V&V effort. In short, our Team faced the dual challenges of fulfilling a new mandate with no specific sense of what the Model Sponsor required (or might require) and having to minimize the draw on the funds sustaining the other simulation model development tasks. Effectively, our Team was starting with a blank sheet of paper.

Starting with this blank sheet of paper and limited funding for V&V, however, proved to be as much of a benefit as it was a challenge. It forced a focus on the essence of model V&V and on the need for cost efficiency, which allowed our Team to adapt to the Model Sponsor's ever changing funding constraints without sacrificing the principles of V&V. The result was a V&V process so well adapted to the low-resolution, multi-purpose simulations of integrated industrial facilities and industrial sectors being developed for the Model Sponsor that it helped shape the Model Sponsor's emerging V&V policy and procedures guidance due out in final form later this year.

Before proceeding to describe our tailored V&V process and documentation methods, it is important to ensure the reader is familiar with V&V terminology and to provide the reader with a sense of the types of SD simulation models our Team is developing for the Model Sponsor's library of analytical tools. Equally important for the reader to know is that these SD simulation models had to be delivered to the Model Sponsor in anywhere from 60 to 120 *days* after project start, not the many months or even years typical of many simulation models characteristic of large-scale efforts that are often the subject of applications in the V&V literature.

A BRIEF V&V TERMINOLOGY PRIMER

If the reader is unfamiliar with the basic concepts and terminology of model V&V, the authors suggest referring to such helpful resources as, among others, four of the references cited at the end of this paper (Coyle, 2000; DMSO,[2] 1996; Robinson, 1999;

[2] Defense Modeling Simulation Office, U.S. Department of Defense.

and Sargent, 1999). However, as a quick primer, model verification refers to a process aimed at answering the question "did we build the model right?" and model validation refers to a process aimed at answering the question "are the results credible?" Verification techniques are deployed to ensure that the requirements for the simulation model and its conceptual model design have been transformed into the computer model with sufficient accuracy. Validation techniques are deployed to reach an acceptable level of confidence that the simulation model's results are sufficiently accurate for its intended use and applicable to the real-world system being modeled. The basic goals of model V&V, therefore, are two-fold. First, demonstrate that the requirements established by the model sponsor can be traced through the structural design, functionality, and output(s) of the model. Second, build confidence in the modeled results by attempting to prove that the simulation model is, in effect, incorrect relative to the real-world system being modeled.

Ultimately, the V&V process furnishes the necessary information to make an informed accreditation decision. Accreditation refers to the decision by a model sponsor (or its accreditation authority) whether to use a specific model for a specific application. Without a sound and documented V&V process, those charged with making the accreditation decision lack the knowledge critical to certifying the model for the intended use.

In its draft policy guidance on model V&V, our Model Sponsor divides verification and validation activities across the simulation model development life cycle into eight functional events as outlined in Table 2. Five of these functional events relate to verification activities and the remaining three relate to validation activities.

Table 2. Eight Functional Events Associated with V&V.

Verification Activities	Validation Activities
1. Requirements Definition	1. Conceptual Model Validation
2. Conceptual Model Verification	2. Data Validation
3. Design Verification	3. Results Validation
4. Code Verification	
5. System Verification	

As shown in Table 3 on the next page, there are numerous specific V&V techniques available for use that can be grouped in a variety of ways (one such grouping scheme shown in Table 3). These techniques can be used in support of one or more of the V&V activities identified in Table 2. Typically, these activities and techniques are carefully planned and documented in the form of a V&V Plan, which is revised as the V&V activity it describes unfolds during the course of the model development life cycle.

Table 3. Categorization of V&V Techniques.[3]

Category	V&V Techniques	
Informal	Audit	Desk Checking
	Face Validation	Inspections
	Reviews	Turing Test
	Walkthroughs	
Static	Cause-Effect Graphing	Control Analysis (calling structure; concurrent process; control flow; state transition)
	Data Analysis (data dependency; data flow)	Fault/Failure Analysis
	Interface Analysis (model interface; user interface)	Semantic Analysis
	Structural Analysis	Symbolic Evaluation
	Syntax Analysis	Traceability Assessment
Dynamic	Acceptance Testing	Alpha Testing
	Assertion Checking	Beta Testing
	Bottom-Up Testing	Comparison Testing
	Compliance Testing (authorization; performance; security; standards)	Debugging
	Execution Testing (monitoring; profiling; tracing)	Fault/Failure Insertion Testing
	Field Testing	Functional (Black-Box) Testing
	Graphical Comparisons	Interface Testing (data; model; user)
	Object-Flow Testing	Partition Testing
	Predictive Validation	Product Testing
	Regression Testing	Sensitivity Analysis
	Statistical Techniques	Special Input Testing (boundary value; equivalence partitioning; extreme input; invalid input; real-time input; self-driven input; stress; trace-driven input)
	Structural (White-Box) Testing (branch; condition; data flow; loop; path; statement)	Sub-model/Module Testing
	Symbolic Debugging	Top-Down Testing
		Visualization/Animation
Formal	Induction	Inference
	Logical Deduction	Inductive Assertions
	Lambda Calculus	Predicate Calculus
	Predicate Transformations	Proof of Correctness

Verification Activities

Requirements definition focuses on, first, defining the problem statement accurately, and then identifying the simulation model outputs, functions, and interactions required to answer the problem statement. Conceptual model verification entails reviewing the conceptual model and ensuring it meets all the specified requirements because it is the

[3] Source: *Verification, Validation, and Accreditation Recommended Practices Guide (DMSO, 1996)*

bridge to the model developer's design for the simulation model. Design verification refers to the process of reviewing the model developer's detailed design to ensure it conforms to the conceptual model. Code verification can utilize a wide variety of techniques to test the simulation model's source code to ensure the detailed design has been implemented correctly. Finally, system verification encompasses tests of the simulation model to determine if it accurately represents the functional design and is traceable to the conceptual model and all requirements.

Validation Activities

Validation of a simulation model is addressed through three activities listed in Table 2. In the early stages of the simulation model development life cycle, validation of the conceptual model ensures that the proposed conceptual model and its design satisfy the fidelity, accuracy, and/or credibility requirements imposed by the model sponsor's problem statement. The key to conceptual model validation is to confirm that the classical influence diagrams, also sometimes referred to as causal loop diagrams have a sound theoretical basis. Discovering and correcting (if possible) fidelity errors at this stage of the simulation model development life cycle is often far less expensive than doing so after the conceptual model has been translated into computer code. As the simulation model is translated into computer code, however, data validation and results validation take center stage. Data validation refers to determining that the data to be used in building and validating the simulation model are sufficiently accurate. Results validation encompasses techniques used to test/see if the modeled results are sufficiently accurate for its intended use. One such technique is termed "face validation," which relies heavily upon the knowledge and judgment of a subject matter expert (SME) familiar with the real-world system, interrelationships, and influences being modeled. "Face validation" refers to the techniques described by Forrester (Forrester, 1961), the most important of which is an evaluation of scenario results (data and graphs) relative to the real system. In some cases this evaluation consists of a comparison to historical behavior from the actual system, in other cases this a reasonable test on the normalcy of the decision policies and resultant actions. The essence of the "face validation" technique is the process whereby scenarios (both planned and impromptu) are conducted in the presence of the SME who evaluates the results and opines on the outcome. To date, each of the SD simulation models our Team has developed has utilized one or more SMEs as a key resource contributing to the V&V process we describe in this paper.

It is important to emphasize here that any model V&V effort and the accreditation decision it supports is linked to the simulation model's particular requirements, design, and intended use. If the requirements, design, and/or the intended use of the simulation model are changed, the V&V effort must be repeated to some degree if not entirely from scratch.

A SPECTRUM OF SD MODEL TYPES

Figure 1 illustrates the spectrum of SD simulation model types that has characterized the tasks assigned to our Team to date by the Model Sponsor. As noted above, our initial proof-of-concept model dealt with a specific industrial process, which is represented by the far left column in Figure 1. Our next three SD simulation models moved to the right along the spectrum shown in Figure 1 to model the manufacture of chemicals produced at certain integrated industrial complexes (two models) and one other type of manufacturing facility. These first four SD simulation models, all developed over a one year-period, were examples of what we term “facility-level” models. Each high-level simulation model presented its unique modeling challenges, but each emerged with the functionality to match almost every possible facility or integrated complex configuration (including the associated utility elements) so that model users could explore multiple “what if” scenarios of interest. On the whole, these low resolution “facility-level” simulation models were very quantitative in that we could make use of known reaction chemistry and energy/material balances to quantify the flows of and interrelationships between the various chemical intermediates and final products produced by the modeled systems.

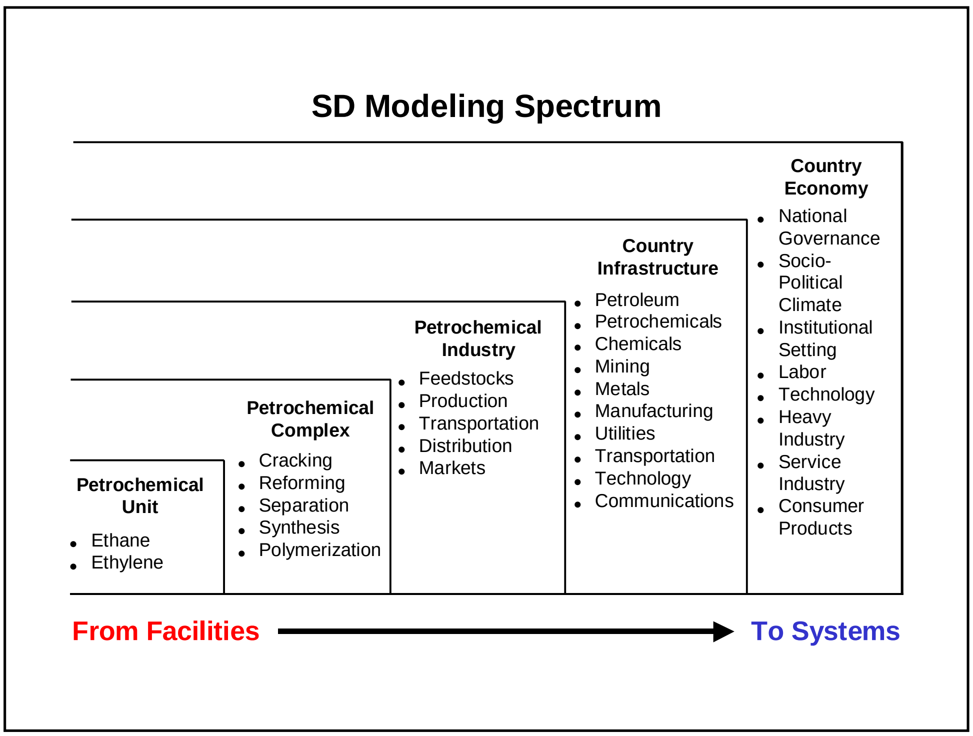

Figure 1. Schematic of SD Models Developed to Date.

Currently, we are developing a SD simulation model more closely resembling the right half of the spectrum illustrated in Figure 1, which we refer to as “system-level” models. This high-level “systems-level” simulation model will explore the interrelationships and influences within a specific commodity industry, which will require our Team to model the influence of certain economic and governmental policies, influences that did not need to be considered in the four “facility-level” models developed previously.

We offer this backdrop concerning the types of SD simulation models our Team has developed and is developing because when we turned to the V&V literature for guidance pertinent to developing the specifics of our V&V process, we found very little applicable "how to" or "how we did it" guidance. We did find many references, like the *Verification, Validation, and Accreditation Recommended Practices Guide* (DMSO, 1996), which spoke to the important principles of V&V, emphasized that V&V needed to be blended into the entire simulation model development cycle, and provided lists of V&V techniques. With few exceptions, however, the available V&V guidance we found either (1) fit modeling efforts of a size, budget, and duration that did not match our circumstance, (2) seemed better suited to acquisition models or training simulations, and/or (3) provided no sense of how to balance the theory of V&V with the cost of V&V.

One of the important V&V principles, of course, is that any V&V effort must be tailored to reflect many considerations, especially, as noted earlier, the intended use of the simulation model and each model sponsor's acceptability and accreditation criteria. Therefore, to a certain degree, the absence of "how to" specifics to guide a balancing of V&V theory and cost for our SD simulation models is an understandable limitation. Similarly, the balance struck between V&V theory and cost for a particular simulation model does not necessarily translate to the next model. Consequently, a V&V process that works well for one simulation model may need to be recast to work as well for another simulation model. Nevertheless, while not an ideal process when measured against the theory of V&V, the authors believe that the V&V process and documentation methods described below have broad "how to" applicability to a range of simulation models.

A BALANCED V&V INVESTMENT: ONE EXAMPLE

As our Team contemplated the appropriate balance between V&V theory and costs for the types of SD simulation models the Model Sponsor has asked us to develop, we chose to focus the available V&V investment on three critical stages of the simulation model development life cycle: requirements definition and tracking, conceptual model verification and validation, and results validation. We prioritized these three stages based on our determination that these were the most value-adding V&V activities after matching the available V&V techniques listed in Table 3 to the appropriate functional events presented in Table 2. In assessing the value-added contributions from the various V&V activities, we adhered to Coyle's notion that the goal of V&V is to provide confidence that the SD model is "well-suited to its purpose and soundly constructed" (Coyle, 1996). This is not to say that no investments were made in such activities as design verification, code verification, system verification, or data validation, but only that the investment in these areas was more measured and the documentation was reduced to a minimum to meet the budgetary constraints.

Figure 2 presents the key V&V activities as captured in the schedule of the major deliverables that occur throughout our generalized model development cycle. For simplicity and as an anchor for the temporal sequence of the milestone V&V activities, we have rolled the simulation model development process into three phases: requirements definition, model development (primarily translation), and testing &

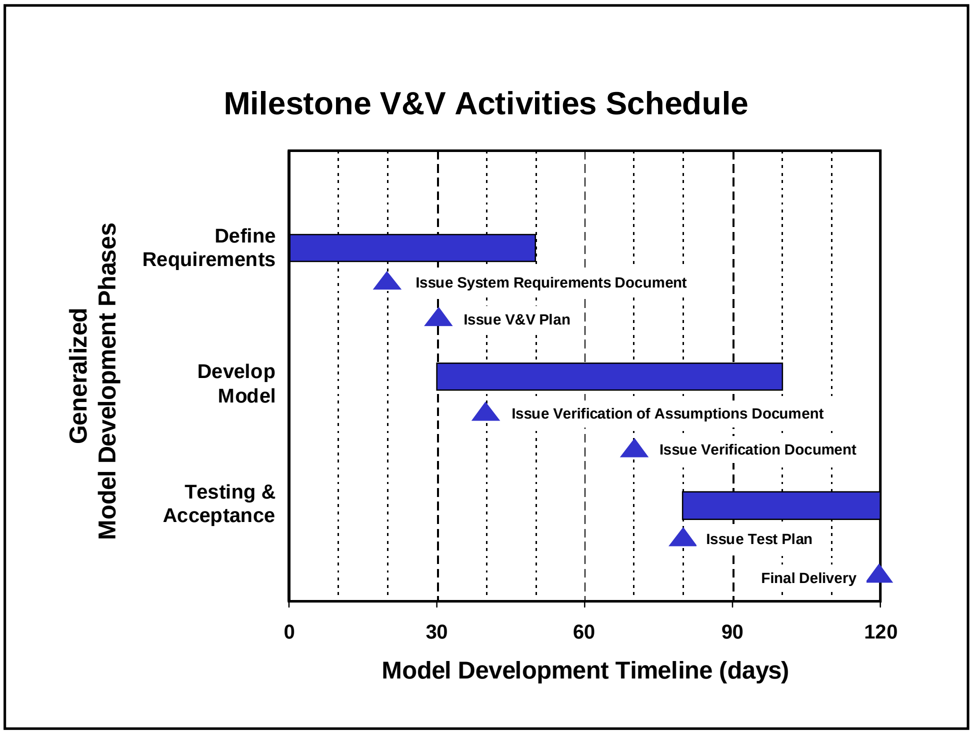

Figure 2. Sequence of Key V&V Activities in the Development Cycle.

acceptance. There are many more phases to our model development cycle than are identified in Figure 2, but our intent is to clarify when these three key V&V activities would occur in the overall development process.

Figure 2 identifies the initial issue date for each of our major V&V-related documents, but there is actually a period preceding each issuance date to draft, review, and finalize each document. After the initial version is issued, it is subsequently revised and updated as needed throughout the model development cycle. As an example, we have needed to issue revisions to the System Dynamics Requirements Document (SDRD) described on the next page late in the testing & acceptance phase in response to validation results prompting new requirements for the model. Although neither the Team nor the Model Sponsor are thrilled with the impacts of model changes identified late in the process, the fact these last-minute changes do occur occasionally attests to the fact that V&V documents are dynamic and evolve with the development process.

It should be immediately clear from reviewing Figure 2 that V&V must be fully integrated into the model development cycle to be successful. V&V begins early in the requirements definition phase and continues up to the conclusion of the development cycle when the SD simulation model is delivered to the Model Sponsor for accreditation. The specific V&V activities undertaken, however, evolve in accordance with the eight functional events identified in Table 2. Three of these stages formed the core of our V&V plan and are more fully described on the pages that follow.

Requirements Definition and Tracking

It is our view that all V&V activity flows from (a) understanding the problem statement for the simulation model and then (b) defining and then tracking the Model Sponsor's requirements for the simulation model through to the final product. Poorly defined simulation models, even if their results are judged credible, are basically worthless. Consequently, the V&V Plan we have developed for our SD simulation models has emphasized capturing the Model Sponsor's requirements as they are developed and refined, and then tracing each requirement through to an element or functionality in the completed model to ensure each requirement is met. Several requirements definition meetings with the Model Sponsor and the SME have been built into the early stages of our simulation model development life cycle (although model requirements can emerge at any stage of the life cycle), and the output of these meetings has been recorded in iterations of the SDRD. Table 4 presents the outline of our SDRD and Table 5 is an excerpt of the model requirements tracking table contained within that document that has been used to track changes to and eventual acceptance of each stated requirement. The SDRD eventually becomes one of the documents incorporated into the V&V Plan by reference.

Table 4. Outline of SD Requirements Document (SDRD).

System Dynamics Requirements Document	
1.0.0	............Introduction
1.1.0	.... Purpose of this Document
1.2.0	.... Intended Audience
2.0.0	............Project Summary
2.1.0	.... Project Name
2.2.0	.... Project Statement
2.3.0	.... Project Team Organization
3.0.0	............Model Requirements
3.1.0	.... General Description
3.2.0	.... Base Requirements
3.2.1	....................................Enumerated Requirement 1
3.2.2	....................................Enumerated Requirement 2
3.2.3	....................................Enumerated Requirement 3
3.2.x	....................................Enumerated Requirement X
4.0.0	............Requirements Tracking Table

The reader will note that the content of our SDRD and format of the requirements tracking table are not particularly inventive or innovative. However, the V&V literature and guidance often lacks such "how to" examples altogether or offers an example that

Table 5. Abridged Requirements Tracking Table.

No.	Original Requirement	Proposed Revision	Revision Accepted/ Rejected/ Revised	Date	Current Requirement
3.2.1	Model a co-located, conventional integrated phosphate fertilizer complex that makes phosphoric acid and derivatives for the fertilizer industry. The major raw materials include phosphate ore, natural gas, and sulfur.	The operating units will be centrally located but not necessarily co-located with the mine.	Accepted	10/27/99	Model a conventional integrated phosphate fertilizer complex that makes phosphoric acid and derivatives for the fertilizer industry. The major raw materials include phosphate ore, natural gas, and sulfur.
3.2.2	The complex size will be primarily based on the user-selected capacity for the phosphoric acid unit. The user will also have the capability to specify the production rates for all the other process units. The model will have the capability to set the unit configuration in one of three ways: (1) entire complex, (2) phosphoric acid production only, (3) phosphoric acid with any allowable combination of derivatives (MAP, DAP, HF, SSP, and TSP).	Drop the option to produce phosphoric acid only. Add AlF_3 as an additional derivative to be consistent with the proposed revisions to requirement No. 3.2.4.	Reviewed and Revised	10/27/99	The complex size will be primarily based on the user-selected capacity for the phosphoric acid unit. The user will also have the capability to specify the production rates for all the other process units. The model will have the capability to set the unit configuration in one of two ways: (1) entire complex, and (2) phosphoric acid with any allowable combination of derivatives. Certain combinations of process units will not be allowed. The user must select either a HF or AlF_3 unit, but not both. The user must select either a MAP or DAP unit, but not both. In addition, the user may choose neither an HF/AlF_3 nor MAP/DAP unit(s) in the unit configuration.
		Certain combinations of process units will not be allowed. The user must select either a HF or AlF_3 unit, but not both. The user must select either a MAP or DAP unit, but not both. In addition, the user may choose neither an HF/AlF_3 nor MAP/DAP unit(s) in the unit configuration.	Revised Version Accepted	11/18/99	
3.2.3	A simulation cycle will be four months (120 days) long with a resolution down to one day.		Accepted	10/13/99	

appears to require a considerable investment of resources. In contrast, our "how to" example of a SDRD accomplishes the basic V&V objectives for the requirements definition stage simply and cost effectively.

Conceptual Model Verification and Validation

The conceptual models our Team has developed for the "facility-level" and "system-level" SD simulation models have ranged from fairly basic and high-level block-flow diagrams to causal loop diagrams. As the "blueprints" for the model translation process that follows in the simulation model development life cycle, the V&V of these conceptual models is a pivotal investment. As noted earlier, the nature of the SD simulation models our Team works on for our Model Sponsor requires the use of one or more SMEs as the key resource involved in conceptual model verification and validation. At these conceptual model meetings, the SME assists our Team in ensuring that the proposed conceptual model (1) speaks to all of the stated requirements as tracked in the SDRD (conceptual model verification) and (2) satisfies the fidelity requirements imposed by the problem statement (conceptual model validation).

Because we rely on a series of meetings with a SME and the Model Sponsor at which the conceptual model for our SD simulation models is reviewed and eventually "approved," we have found it necessary to develop two forms of documentation for the conceptual model V&V efforts. First, we develop very detailed narrative minutes of each of SME meeting. Among other things, these minutes, which are recorded by our designated V&V Lead on the model development team, track the ebb and flow of the conceptual model discussions. The minutes not only record the consensus decisions reached by the meeting participants, but also record the proposed ideas and competing opinions often expressed at brainstorming meetings. The minutes are then shared with all the meeting participants and subjected to a formal review and approval process. As with the SDRD, these minutes are incorporated by reference in the V&V Plan.

It has been our experience, however, that there is more to the conceptual model development process that takes place outside of the scheduled meetings and that initially tends to occur among the model developers only. Often, a model developer receives a set of fairly high-level model requirements from the model sponsor that must be fleshed out to a greater level of detail or resolution in order to meet the thrust of the requirement or to effect the desired functionality in the simulation model's design and source code. In addition, it may be necessary to develop certain simplifying assumptions or surrogate measures, particularly when there are data gaps or only a partial understanding of the relationship or influence to be modeled. This process of fleshing out or adding necessary details to the model sponsor's high-level requirements usually entails making a series of assumptions that are not captured in the SDRD per se and which are not evident in an influence diagram. Again, the literature speaks to the necessity of this V&V step in the development of conceptual models, but offered us no concrete example of "how to" capture the results of this step. Therefore, our Team developed the Verification of Assumptions Document (VAD), an excerpt from which is shown in Table 6.

Table 6. Abridged Verification of Assumptions Document.

Verification of Assumptions	Date	Revision
Plant Configuration. The user may select any desired configuration of available plants with the following exceptions: 1. A complex cannot have both a MAP and DAP plant or both a hydrofluoric acid and aluminum fluoride plant. 2. If either a MAP, DAP, HF, AF, or TSP plant is present, then there must be a phosphoric acid plant with adequate capacity.	10/27/99	
Intermediate Products. The only intermediates that can be received from outside sources are phosphate rock, sulfuric acid, and ammonia. Inventory will be managed for these materials and coordinated with offsite ordering.	10/13/99	
Capacity. The user will have the flexibility to establish the plant capacities for most plants. However, a warning message will notify the user when the phosphoric acid plant cannot adequately meet phosphoric acid feed requirements for the TSP, MAP, and DAP plants. This exception is necessitated by the inability to receive phosphoric acid from offsite.	11/18/99	The capacity of the hydrofluoric acid and aluminum fluoride plants will be calculated based on the steady state fluosilicic acid output of the phosphoric acid plant. This exception is necessitated by the inability to receive fluosilicic acid from offsite.
Onsite Storage of Feed Materials. Each plant, and the complex as a whole, will have limited capacities to store raw materials. The user will have the ability to alter most of these constraints. Normal (or target) operating inventory levels will be less than the scenario cycle time (120 days) to allow the user to examine the dynamics of inventory shortages.	10/13/99	

Each VAD, which is also incorporated by reference into our V&V Plans, is a compilation of the model developer's translational assumptions and is continuously updated as the conceptual model V&V process unfolds and the simulation model translation activity proceeds. The VAD is reviewed periodically with the SME(s) and the Model Sponsor to secure their agreement with these assumptions before the coding of the simulation model begins in earnest. In keeping with our other "how to" examples, the VAD is a simple cost-effective means of tracking these key assumptions that "gets the job done" for the minimal investment.

Results Validation

In the case of documenting results validation, the V&V literature actually contains numerous examples of test plans and the documentation generated in executing these test plans. It was evident, however, that these test plan examples were quite elaborate, very quantitative-oriented, and better suited to the acquisition model environment than to the types of SD simulation models our Team was developing. Again, we opted for a simpler cost-effective approach combining a brief narrative test plan describing the rationale behind the selected test scenarios and the results validation-tracking table an excerpt of which is shown in Table 7.

Initially, test scenarios designed to test the desired functionality and fidelity of the simulation model are solicited from the model developers, the Model Sponsor, model users, and the SME(s). The test scenarios are closely linked to the requirements identified in the SDRD. The first priority is to demonstrate the steady state or baseline scenario(s) to build confidence in the model fundamentals. The second priority is to investigate the model behavior that results from a series of perturbations to the baseline scenario. Each perturbation is designed to exercise a specific model capability that is identified in the SDRD. Each test scenario is categorized and then assigned its own row in the tracking table. The SME's and the Model Sponsor's reactions to the results of each test scenario are then recorded in the tracking table at the results validation meeting (as well as in the meeting minutes). As necessary, test scenarios are added and/or repeated at another results validation meeting if the SME's and Model Sponsor's reactions to the initial test results dictate a revision to the simulation model. Eventually, the tracking table shows the date at which the results of the test scenarios are accepted by the SME(s) and the Model Sponsor as sufficiently credible for the simulation model's intended purpose. This "final" results validation-tracking table is then incorporated into the "final" version of the V&V Plan.

Table 7. Abridged Validation Scenario Tracking Table.

Test No.	Category	Description	Date	Accepted/ Revised	Comments
1a.	Configuration/ Steady State	100 tpd ammonia unit; 1,200 tpd phosphoric acid unit; a MAP unit and an AlF_3 unit, but no SSP or TSP unit; default sizes for all the other “enabled” units in the complex; default feedstock and utilities priorities; and all default values on Configuration Screen No. 2	11/18/99	Revised	Add production rate adjustment logic in the sulfuric acid and phosphoric acid units that automatically adjusts the rate whenever the supply of these intermediates exceeds demand.
2a.	Offsite Ordering/ Inventory Management	200 tpd ammonia unit; 1,200 tpd phosphoric acid unit; a 1,025 tpd ore beneficiation unit; a DAP unit and an HF acid unit, but no SSP or TSP unit; default sizes for all the other “enabled” units in the complex; default feedstock and utilities priorities; and all default values on Configuration Screen No. 2, except the complex was unable to order phosphate rock for 20 days and the target and maximum inventories for phosphate rock were set at 15 days and 30 days of coverage, respectively	11/18/99	Accepted	
3a.	Utilities Shock	200 tpd ammonia unit; 1,200 tpd phosphoric acid unit; a DAP unit and an AlF_3 unit, but no SSP or TSP unit; default sizes for all the other “enabled” units in the complex; default feedstock and utilities priorities; all default values on Configuration Screen No. 2; and the model is run in steady state for 30 days followed by disabling the (i) steam turbine for 90 days and (ii) around day 60, disabling the off-site electrical power for 30 days.	11/18/99	Revised	Although most units shut down in the face of a loss of electrical power, the ammonia and sulfuric acid units continued to operate, even after the off-site power grid was lost. Both the ammonia and sulfuric acid units require electrical power to operate.

METRICS

What's the true cost of V&V? Certainly, it is an activity that we cannot afford not to do, but one that can consume a lot of resources if not kept in perspective. Here are the performance metrics on how much our Team has spent to date to address the gap between theory, guidance, and practice.

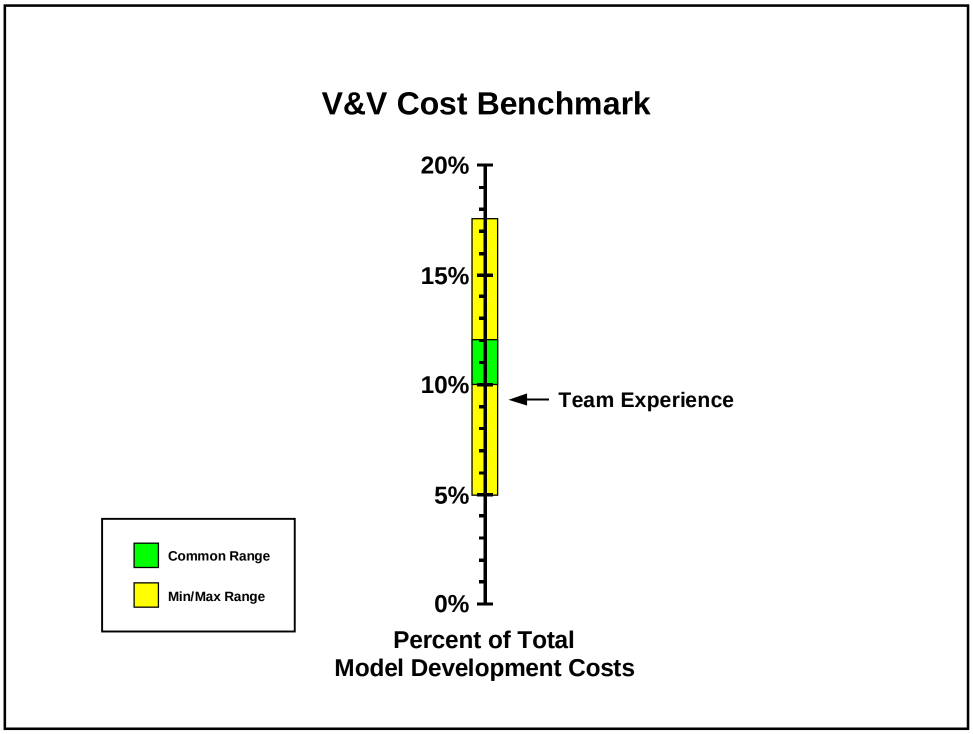

Figure 3. Experience rating for V&V Costs.

The *Verification, Validation, and Accreditation Recommended Practices Guide* (DMSO, 1996) offers some benchmark ranges. Using historical data, it reports that V&V activities expressed as a percentage of the total model and simulation budget ranges from 5 percent to 17.5 percent, with most efforts concentrated in the narrower range of between 10 and 12 percent. These benchmarks are presented in Figure 3. It is not clear to us what V&V costs are actually included in these cited historical benchmarks. If we assume that the narrow range captures some measure of the average practice that would include direct labor associated with executing V&V activities, then we can infer that it is primarily the time spent by the designated V&V Lead. Using this definition, our historical V&V investment of 9.5 percent of total funding falls just below the narrow benchmark range. Our experience has been averaged over a combination of "facility-level" and "system-level" models previously described.

One can argue that the benchmark understates the true cost of V&V because it excludes the cost of the SME. For example, if we were to include the SME costs involved in our past SD simulation efforts, the actual V&V investment as a percentage of total funding would increase by 8.5 percent to a total of 18 percent. Clearly, this analysis points to the weaknesses of what is or can be counted as part of the V&V investment.

Setting the SME costs aside, however, we attribute our V&V cost performance to successfully integrating V&V into every step in the model development process. V&V should be a mindset executed by everyone on the model development team with each team member embracing the V&V principles as an everyday part of his or her work process. This total team commitment has the benefit of lowering direct V&V labor and shifting the focus away from the perception that V&V is only an audit.

SUMMARY

One of the key V&V principles stated in the *Verification, Validation, and Accreditation Recommended Practices Guide* (DMSO, 1996) is that "V&V is both an art and a science, requiring creativity and insight. We believe an appropriate corollary to this principle would be "no one type or size fits all." It is not too surprising, therefore, that the V&V literature tends to be devoid of specific "how we did it" examples. It is still disappointing, however, to turn to such guidance and come away feeling it would have been helpful to see some concrete "how to" and "how we did it" examples. Apparently, others have had similar reactions to the available V&V guidance as we understand that some of the available guidance is being revised to offer more "how to" guidance in addition to discussions of V&V theory and individual V&V techniques. In this paper, our Team provided a synopsis and specific examples of how we balanced V&V theory and the resources we were allocated for the V&V of several low-resolution multi-purpose SD simulation models toward achieving a cost-effective and documented V&V investment.

REFERENCES

Coyle, G., 1996. *System Dynamics Modelling: A Practical Approach.* Chapman and Hall: London, pp. 12 & 155.

Coyle, G. and Exelby, D., 2000. "The Validation of Commercial System Dynamics Models." *System Dynamics Review*, Volume 16, No. 1, pp. 27-41.

Defense Modeling and Simulation Office (DMSO), November 1996. *Verification, Validation, and Accreditation Recommended Practices Guide*. U.S. Department of Defense. Office of the Director of Defense Research and Engineering.

Forrester, J. W., 1961. *Industrial Dynamics.* MIT Press: Cambridge, MA, pp. 119-123.

Robinson, S., June 1999. "Simulation Verification, Validation, and Confidence: A Tutorial." *TRANSACTIONS of the Society for Computer Simulation International.* Volume 16, No. 2, pp.63-69.

Sargent, R., 1999. "Validation and Verification of Simulation Models." *Proceedings of the 1999 Winter Simulation Conference*, pp. 39-48.