IDA

INSTITUTE FOR DEFENSE ANALYSES

Mobility Capabilities Study 2005: Use Case Study

Dr. Jack Jackson

Dr. Thomas Allen, task leader

January 2009

Approved for public release; distribution is unlimited. 3 February 2011.

IDA Document D-3710

Log: H 09-000001

The Institute for Defense Analyses is a non-profit corporation that operates three federally funded research and development centers to provide objective analyses of national security issues, particularly those requiring scientific and technical expertise, and conduct related research on other national challenges.

About This Publication

This work was conducted by the Institute for Defense Analyses (IDA) under contract DASW01-04-C-0003,Task BA-8-2911, "VV&A Use Case Study," for the Director, Program Analysis and Evaluation and the Modeling and Simulation Coordination Office, Office of the Deputy Under Secretary of Defense (Science and Technology). The views, opinions, and findings should not be construed as representing the official position of either the Department of Defense or the sponsoring organization.

INSTITUTE FOR DEFENSE ANALYSES

IDA Document D-3710

Mobility Capabilities Study 2005: Use Case Study

Dr. Jack Jackson

Dr. Thomas Allen, task leader

THIS PAGE LEFT INTENTIONALLY BLANK

Preface

This document was prepared under task order BA-8-2911, Analysis Community Verification, Validation and Accreditation Use Case, for the Director, Program Analysis and Evaluation and the Modeling and Simulation Coordination Office, Office of the Deputy Under Secretary of Defense (Science and Technology). It addresses Subtask D5 of the task order to provide a written assessment of the US Department of Defense's Mobility Capabilities Study based upon the principles determined in Task A of the task order.

The author would like to thank the reviewers at the Institute for Defense Analyses who contributed their time and expertise to enhancing this document: William L. Greer, System Evalutation Division; Thomas L. Allen, Joint Advanced Warfighting Division (JAWD); Laura M. Williams, Cost Analysis Research Division; Stuart Starr, JAWD; and James H. Kurtz of JAWD.

THIS PAGE LEFT INTENTIONALLY BLANK

Executive Summary

In 2005, the US Department of Defense (DoD) embarked on a major effort to determine Departmental mobility requirements using an innovative analysis technique—the Mobility Capabilities Study (MCS 2005). During its evaluation of the MCS 2005, the US General Accountability Office (GAO) identified several deficiencies that undermined the credibility of the analysis. In particular, the GAO acknowledged that while the DoD used an innovative approach to conduct the study, there were significant problems. The GAO concluded that:

- the DoD, while acknowledging some limitations, did not fully disclose how those limitations would affect its study, and
- measured against generally accepted research standards, there were limitations in the MCS 2005 that raised questions about the study's adequacy and completeness.

In general, the GAO determined that DoD practice did not match DoD policy. In response, DoD drafted new guidelines for the Verification, Validation, and Accreditation (VV&A) of models and simulations. This document is an application of those guidelines to evaluate the VV&A practices in MCS 2005 after the fact.

The results of this use case study indicate that DoD analysis teams typically perform adequate V&V but do not document their efforts by completing a V&V report. DoD analysis teams need to emphasize completing an Accreditation process with a risk assessment and documenting the results in an Accreditation report.

THIS PAGE LEFT INTENTIONALLY BLANK

Mobility Capabilities Study 2005: Use Case Study

A. Introduction

Essentially, all models are wrong, but some are useful.[1]

In the quote above, Professor Geroge Box addresses two key issues:

- All models are abstractions of reality. Because they are abstractions and do not intend to reproduce reality exactly, they are inherently "wrong."
- However, the skilled practitioner will endeavor to ensure the model adequately addresses key aspects of reality so it can be used to examine something of interest. If the skilled practitioner is successful then the model may be useful.[2]

Verification, Validation and Accreditation (VV&A) are the US Department of Defense (DoD) processes used to determine whether the skilled practitioner or modeling and simulation (M&S) team successfully produced an abstraction of reality suitable to its purpose. The range of intended purposes for DoD modeling includes: analysis, acquisition, training, testing, experimentation, and operational planning

In 2005, the DoD embarked on a major effort to determine Departmental mobility requirements using an innovative analysis technique, the Mobility Capabilities Study (MCS 2005).[3] During its evaluation of the MCS 2005, the US General Accountability Office (GAO) identified several deficiencies that undermined the credibility of the analysis. In particular, the GAO acknowledged that while the DoD used an innovative approach to conduct the study, there were significant problems. The GAO concluded that:

- the DoD, while acknowledging some limitations, did not fully disclose how those limitations would affect its study, and

---

1 Attributed to George Box, Professor Emeritus of Statistics, University of Wisconsin. George E.P. Box and Norman R. Draper, *Empirical Model-Building and Response Surfaces* (New York: Wiley Press, 1987), 424.

2 George E.P. Box, "Robustness in the Strategy of Scientific Model Building," in *Robustness in Statistics*, R.L. Launer and G.N. Wilkerson, eds. (New York: Academic Press, 1979).

3 *2005 Mobility Capabilities Study (MCS, 2005)*, Office of the Secretary of Defense, Program Analysis and Evaluation OSD/PA&E, December 2005.

- measured against generally accepted research standards, there were limitations in the MCS 2005 that raised questions about the study's adequacy and completeness.

The GAO's primary findings were:

> *Aspects of modeling and data were inadequate in some areas because data were lacking and the models used could not simulate all relevant aspects of the missions. The report did not explain how these limitations could affect the study results or what the impact on projected mobility capabilities might be. Generally accepted research standards require that models used are adequate for the intended purpose, represent a complete range of conditions, and that data used are properly generated and complete. For example, the MCS modeled hypothetical homeland defense missions rather than homeland defense demands derived from a well-defined and approved concept of operations for homeland defense, because the specific details of the missions were still being determined and the data used may have been incomplete. The MCS also was unable to model the flexible deterrent options/deployment order process to move units and equipment into theater because of lack of data, but the study assumed a robust use of this process. In addition, the MCS report contains more than 80 references to the need for improved modeling or data.*
>
> *While the MCS concluded that combined US- and host nation transportation assets were adequate, when describing the use of warfighting metrics in its analyses, the report [MCS 2005] does not provide a clear understanding of the direct relationship of warfighting objectives to transportation capabilities. Additionally, the report stated that further analysis is required to understand the operational impact of increased or decreased strategic lift on achieving warfighting objectives. Relevant generally accepted research standards require that conclusions be supported by analyses. The use of both warfighting and mobility metrics would allow decision-makers to know whether combat tasks were achieved and how much strategic transportation is needed to accomplish those tasks.*
>
> *In some cases, the MCS results were incomplete, unclear, or contingent on further study, making it difficult to identify findings and evaluate evidence. Relevant research standards require results to be presented in complete, accurate, and relevant manner. For example, the report recommends further studies and assessments, five of which are under way. However, DoD has no plans to report how these studies impact the MCS results once the studies are finished. In addition, the report contains qualified information that is not presented clearly, such as varying assessments of intra-theater assets in three different places."*[4]

---

[4] "Study Limitations Raise Questions about the Adequacy and Completeness of the Mobility Capabilities Study Report," US General Accountability Office, GAO-06-938, September 2006; and "Issues Concerning Airlift and Tanker Programs," US General Accountability Office, GAO-07-566T, 7 March 2007.

In particular, the GAO cited the following issues concerning VV&A:

> *The Validation, Verification and Accreditation (VV&A) of models and data was not complete because it was not done in accordance with DoD policy or relevant research standards.*
>
> *a. The MCS did not contain a signed VV&A accreditation report from the study director addressing the models and data certification.*
>
> *b. DoD officials did not comply with VV&A policy with respect to legacy M&S and asserted that the long use of legacy M&S constituted an alternate VV&A process.*
>
> *c. No documentation existed for any alternate VV&A procedures to establish the credibility of the results.*
>
> *d. No additional documentation existed to support key analytical and decisionmaking processes used by the senior DoD leadership.*[5]

This Use Case reviews DoD policies concerning VV&A in the next section and then suggests guidelines for conducting VV&A. In section C, the guidelines are used as a checklist to evaluate the MCS 2005 and to provide examples for improvement to DoD VV&A procedures. Section D contains Conclusions. It may be helpful to periodically refer to the GAO criticisms to highlight improved procedures for VV&A.

B. Verification, Validation and Accreditation

For two decades, the fundamentals of VV&A have been widely known and understood. The Military Operations Research Society (MORS) mini-symposium on Simulation Validation in 1994 highlighted the following definitions (excluding the words in parentheses):

Verification—the process of determining that a model implementation (and its associated data) accurately represents the developer's conceptual description and specifications.

Validation—the process of determining the degree to which a model (and its associated data) is (are) an accurate representation of the real world from the perspective of the intended uses of the model.

Accreditation—an official determination that a model (simulation, or federation of models and simulations and its associated data) is (are) acceptable for a specific purpose.[6]

---

[5] "Study Limitations Raise," September 2006.

[6] Julian I. Palmore, ed., "Simulation Validation (SIMVAL) 1994," Mini-Symposium Report (Alexandria, VA: Military Operations Research Society, 1994).

By adding the items within the parentheses to these definitions, amplifications were made in the current copy of the draft *DoD Instruction 5000.61, DoD Modeling and Simulation (M&S) Verification, Validation, and Accreditation (VV&A),* which governs the practice of VV&A for all DoD models and simulations. Notice the inclusion of the *model's data* within the V&V processes and the expansion from *model* to *model, simulation and/or federation of models* to complete the Accreditation process.

Figure 1 depicts the key relationships between VV&A activities and the normal practice of developing models and simulations to support DoD activities. The red arrows indicate the key aspects of VV&A that are considered essential by most of the M&S literature. The Recommended Practices Guide published by the Defense Modeling and Simulation Office (DMSO) in 1996 elaborated on the basic issues surrounding the terms Verification, Validation and Accreditation (see Table 1).[7]

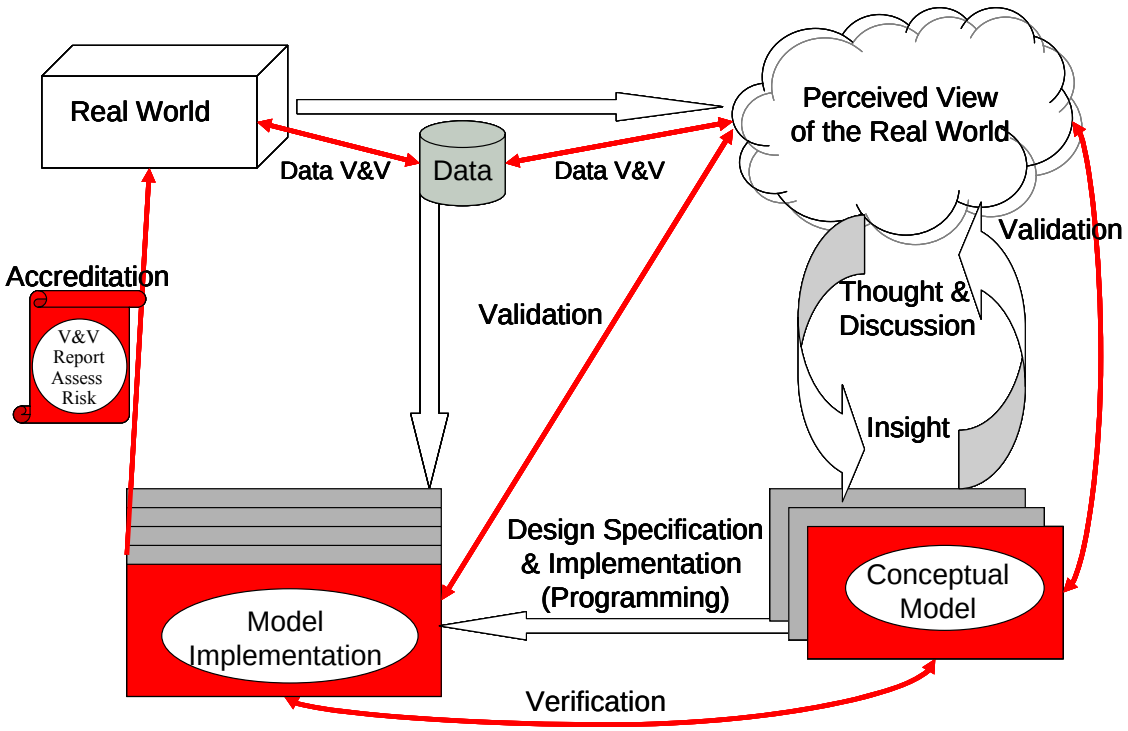

Figure 1. VV&A Relationships in M&S[8]

---

[7] *Verification, Validation and Accreditation (VV&A) Recommended Practices Guide*, Office of the Director of OSD/DDRE/DMSO, November 1996.

[8] Figure adapted from Dean S. Hartley III, "Verification and Validation in Military Simulations," Figure 1 in *Proceedings of the 1997 Winter Simulation Conference*, S. Andradottir, K. J. Healy, D. H. Withirs, and B. L. Nelson, eds.

Table 1. Basic issues concerning the terms VV&A

Validation	Was the right M&S built?	▪ Seeks to determine if the fidelity of the M&S is adequate for the intended purpose. ▪ Ensures a model or simulation conforms to a specified level of accuracy when its outputs are compared to the real-world system. ▪ Generally performed at two levels: conceptual model validation and results validation.
Verification	Was the M&S built right?	▪ Ensures a model or simulation meets the user's requirements and that it implements those requirements correctly in software.
Data V&V	Is the data used in the M&S credible and the most appropriate for its purpose?	▪ Ensures the data accurately represent key aspects of the real world and encompass the range of activities adequate to meet the intended use of the M&S.
Accreditation	Is this the right M&S to use for this purpose?	▪ Officially certifies that a model, simulation (or federation of models and simulations) and its associated data are acceptable for a specific purpose. ▪ Assesses the acceptability, suitability, and risk associated with the use of the M&S for this intended purpose.

DoD Instruction 5000.61 defines procedures associated with VV&A in Enclosure 3 stating that those procedures include the following:

Verification

Document M&S verification activities to record how well a model or simulation implementation and its associated data represent the developer's conceptual description and specifications as well as the application requirements, to include:

1. Identify the person or organization performing the verification activities.
2. Describe the model or simulation implementation version or release and identify the developing organization.
3. List or reference the M&S application specifications and conceptual description.
4. List and/or describe the verification activities.
5. Summarize the verification results.
6. Identify M&S limitations and assumptions.

Validation

Document validation activities, at both the conceptual model and M&S implementation phases, to record how well a model or simulation is an accurate representation of the real world from the perspective of the intended use, to include:

1. Identify the person or organization performing the validation activities.

2. Identify the model, simulation, or M&S federation version and/or release and its developing organization.
3. Describe, or reference a description of, the Simulation Conceptual Model.
4. List, describe and/or identify the validation activities.
5. Summarize the validation results.
6. Identify M&S limitations and assumptions.

Accreditation

Document accreditation activities and results, specifically the acceptability of the model, simulation, and its associated data for the intended use, to include:

1. Identify the person or organization performing the accreditation activities.
2. Identify of the model, simulation, or federation version and/or release and the developing organization(s).
3. Identify the M&S Application Sponsor's intended use for the model, simulation, and/or federation.
4. List or describe the requirements addressed by the model, simulation (to include an M&S federation if appropriate), and associated data.
5. Identify the acceptability criteria.
6. Describe the accreditation methodology including V&V activities that support the accreditation, data verification and validation, and risk assessments.
7. Summarize the results of the accreditation assessment.
8. Identify the M&S Application Sponsor and document the accreditation decision.

The current Draft DoD Instruction 5000.61 also requires the use of MIL-STD-3022 templates or a suitable substitute for reporting.

1. VV&A Guidelines

Based on these items, the Institute for Defense Analyses developed guidelines to help the DoD M&S community methodically accomplish the tasks required by the Office of the Secretary of Defense (OSD). The guidelines can be viewed (or act) as a checklist for the senior analyst of each analytical study. The guidelines include:

1. Identify the intended purpose for the model, simulation, and/or M&S federation and their data.

2. Identify and describe the model, simulation, or M&S federation version and/or release and its developing organization.
3. Provide documentation and the configuration management plan for the M&S and/or M&S federation to support the following items.
4. Verify and Validate the Model (Identify the Point of Contact (POC)).
    - Validate the M&S conceptual design, based on the intended purpose, i.e., is the M&S abstraction of reality appropriate for the intended use?
    - Verify that the M&S algorithms, design specifications, and model implementation match the conceptual design and are sufficient for the intended purpose.
    - Validate the resultant M&S implementation, based on the intended purpose.
5. Verify and Validate the Data (Identify the POC).
    - Identify the authoritative source for the data.
    - Verify the internal consistency of the data, based on the intended purpose; were they collected consistently (conditions, limitations, etc.) with the intended purpose of the M&S?
    - Validate the data for use in this M&S, based on the intended purpose.
    - Were the data compared to corresponding known, real-world or best-estimate values?
    - If there are transformations of the data, how do you know if the aggregated, disaggregated, or transformed data are accurate for the intended purpose?
    - If there is a federation of models exchanging data, how do you know the data exchange is performed correctly and the assumptions between models/simulations are consistent?
6. List the known assumptions, limitations, and constraints for the M&S and/or M&S federation.
7. Sponsor decides whether to accredit the M&S and associated data (Identify the POC).
    - Describe the requirements addressed by the model, simulation (to include an M&S federation if appropriate), and associated data?
    - What V&V activities were performed?
    - What are the acceptability criteria? (i.e., why should the Sponsor believe the results from the M&S and its associated data)

- What are the known or projected risks (uncertainty and consequences) associated with the use of the M&S and their data based on the intended purpose?
- Has an accreditation report been completed and provided to the study Sponsor?
- Has the Sponsor reviewed this information and accredited the model and data for this use?

A list of basic V&V techniques from the combined work of Sargent and Department of the US Army Pamplet 5-11 is provided in Appendix C.[9] A more complete list of V&V techniques and VV&A principles are available from Balchi.[10]

C. Applying the VV&A Guidelines to the MCS 2005

Each portion of the VV&A Guidelines will be treated, in turn, as checklist items to examine the adequacy of the VV&A effort in MCS 2005 and to suggest how adherence to such guidelines could better meet the intent of VV&A policies in the Department. As references, the author consulted the MCS 2005 Final report, the report's Appendix M, which addresses the VV&A effort, and interviews with key personnel engaged in the MCS 2005 study to determine the quality of the VV&A effort.

1. What is the intended purpose for the model, simulation, and/or M&S federation and their data?

The intended purpose for M&S used in MCS 2005 is spelled out in numerous study documents and clearly established by the Objectives stated in the Executive Summary.

> The analysis conducted by the Mobility Capabilities Study addressed the following objectives:
>
> - Identify and quantify how variations in mobility capabilities support the Defense Strategy, from point of origin to point of use (e.g., tactical assembly areas) and return, using forces available in 2012. Examine alternatives that include variations in assets (land, air, sea), sources (military, civilian, and foreign) forward basing, sea basing, pre-positioning (afloat and ashore), air refueling capability, advanced logistics concepts, and destination theater austerity based on the new global footprint and global presence initiatives.

---

[9] Robert G. Sargent, "Validation and Verification of Simulation Models," *Proceedings of the 2004 Winter Simulation Conference*, R.G. Ingalls *et al.*, eds., 2004; and *Department of the Army Pamphlet 5-11, Verification, Validation, and Accreditation of Army Models and Simulations, 30 September 1999.*

[10] Osman Balci, "Verification and Validation and Testing of Models," *Encyclopedia of Operations Research and Management Science* (Kluwer Academic Publishers, 2001).

- Identify mobility capability gaps, overlaps, or excesses and provide associated risk assessments with regard to the ability to conduct operations. Recommend mitigation strategies where possible.
- Identify mobility capability alternatives that mitigate the potential impacts on the logistic system of irregular, catastrophic, or disruptive threats.
- Identify and quantify the combinations of air mobility, sealift, surface movements, prepositioning, forward stationing, seabasing, engineering, and infrastructure capabilities required to support theater and tactical deployment and distribution.
- Identify new metric(s) for assessing mobility capabilities.

The Executive Summary continues to outline the study scope and many of the key assumptions. The M&S purpose step was clearly addressed by the study effort.

2. Identify and describe the model, simulation, or M&S federation version and/or release and its developing organization.

Each M&S used during the MCS 2005 is described briefly in the study's Appendix M. This appendix contains the model name, proponent, a point of contact, a short model description, a list of other users, and occasionally some notes on VV&A status or VV&A activities previously accomplished. Except for the AMP and ELIST models, no mention is made of the model's version number, and none of the models are accompanied by their release dates. A check with the model proponents reveals that this information was available and could easily have been incorporated in the study documentation. The models cited in Appendix M include:

- AMP 10.0
- JICM
- EADSIM
- MIDAS
- THUNDER
- CMARPS
- ELIST 8.0
- ITEM
- ARCEM

No federations were used in this study, though AMP, MIDAS, and ELIST have migrated toward a federated configuration over time. In this study, these three models were used as if federated through the use of a "sneaker net" for manual transmission of output data from one model into input data for another model, a process for which no documentation exists. Appendix M describes the use of AMP-MIDAS as if a federation existed for these two models and was used for MCS 2005.

Of the nine models, only CMARPS has a documented VV&A history. The initial V&V process was completed by Strategic Air Command (SAC) in 1990–92, but all documentation was subsequently lost. Before MCS 2005, the Air Combat Command (ACC) initiated a new

VV&A effort for CMARPS. After completing MCS 2005, AF/AMC/A9 completed a VV&A activity for CMARPS and documentation of that effort is available on request from Maj Langhorne (AF/AMC/A9). CMARPS has a consistent configuration management plan and the proponent for the model has maintained an ongoing VV&A process since its first application in 2006. CMARPS is the only model in the list above which has documented a rigorous VV&A process in an attempt to meet current DoD standards. This documentation is being updated for the MCRS 2016 study to include better accreditation criteria and a current risk assessment.

3. Provide documentation and the configuration management plan for the M&S and/or M&S federation to support the following VV&A items.

At first glance, the MIDAS model seems to be in violation of this guideline; further discussion of this problem is provided in the V&V section for MIDAS, below. However, further investigation reveals that users of the model were aware of an analyst's manual that existed prior to the study with complete documentation of the conceptual model. In addition, a configuration management plan existed for each model/simulation used by the study. Currently, a configuration control board meets regularly to guarantee coordination of these activities between the two locations in DoD (US Transportation Command (TRANSCOM) and the Pentagon) that use this model.

Appendix M refers to AMP-MIDAS as if a federation (or graphic user interface (GUI) and model) exists between these two models to provide end-to-end scheduling and simulation modeling for multiple strategic deployment scenarios. Documentation is being developed to explain this federation but did not exist during MCS 2005. Though the AMP GUI was undergoing considerable "script" testing beyond Appendix M's discussion (see section 4 below), no documentation explains the outcomes of those tests. Likewise, no documentation or discussion is available to address MIDAS reception of the AMP outputs to ensure that the data are received accurately by MIDAS and transformed correctly to a form useful for MIDAS. For MCS 2005, testing was performed to confirm these aspects of the M&S operated correctly and those tests are discussed in some detail. After completing these tests, validation of the AMP-MIDAS federation was performed using Face Validity checks.

All other M&S used for MCS 2005 had proponents, documentation, and maintained current configuration management plans.

4. Verify and validate the model (identify the POC)

- Validate the M&S conceptual design, based on the intended purpose.

- *Is the M&S abstraction of reality appropriate for the intended use?*

- Verify the M&S algorithms, design specifications, and model implementation match the conceptual design, based on the intended purpose.
  - *Are they sufficient for the intended use?*
- Validate the resultant M&S implementation, based on the intended purpose.

Based on the extensive review process used in the MCS 2005, the original participants as well as the author of this document believe that the third bullet above—Validation of the M&S implementation—occurred at numerous stages for all of the models used in the MCS (except perhaps for an examination of the AMP-MIDAS federation already mentioned). With the exception of CMARPS, (which followed an internally-generated and well-documented V&V process and for which an Accreditation Report was generated), none of the other models underwent a rigorous conceptual model validation. The general attitude was characterized by the statement, "The Services say this model is adequate for *X* purpose and that is good enough for us."

The following excerpts are from Appendix M and describe MCS and legacy attempts to V&V the other models. These attempts do not always address V&V for the intended purposes of MCS 2005.

AMP-MIDAS VV&A excerpt and discussion

AMP—Verification and Validation (V&V) of the AMP application is a continuing process that began with the release of AMP 10 and continues as new versions of the application are released to the user community. AMP software V&V includes verification of user inputs through the GUI, as well as verification of AMP generated input/output files and the integration of Federation model components. Through the use of scripts, regression testing, unit testing, and a testing team, AMP continues its verification and validation process as development continues.

Testing of the AMP GUI has primarily been accomplished through the use of scripts. USTRANSCOM has written scripts within Excel workbooks that look at individual cells and the data that is input into those cells. The tester will walk through the AMP application, following the script testing steps and the results are compared to the expected results. Results are given a pass or fail grading and the tester is given the option to document problems with the result from the script and/or general comments. Failing tests and documented problems and concerns are all reported within the web-based bug database. The database allows users and developers to coordinate and exchange detailed bug information, bug prioritization and bug status. Script testing includes:

- *verifying that upper and lower bound limits are in place and correct for each of the data input areas.*
- *testing AMP's response to 'out of bounds' data.*
- *verifying that changes made to one specific piece of data is filtered down to correctly change/impact other data fields within the GUI.*
- *verifying that all data needed to be manipulated by the user, is within the application and easy to use.*

MIDAS—To ensure that input data is translated correctly into MIDAS input files, regression testing and unit testing are completed. The regression test launches AMP, loads test scenarios, and produces the MIDAS input files. These MIDAS input files are compared to baseline test results to determine any deviations. Unit tests are also conducted and focus on newly coded functionality within the application where prior bugs were known to exist. If problems still exist within the code, the unit test fails and reports are generated, and if the problem no longer exists, the unit test will report success. Currently, regression testing consists of 3 different test scenarios and unit testing and consists of over 300 tests that help validate internal business logic and/or computation.

A testing team that includes technical developers and testers at the AMP development site, as well as functional testers at USTRANSCOM, continually test all portions of the AMP application to continually verify that data input is working correctly and to ensure that model integration is maintained. At a high level, the testing team performs basic integration testing to ensure that all of the modeling components within the AMP Federation continue to interoperate. The team then evaluates the model results and compares them to a previously established benchmark. This allows the team to determine if there are integration problems between AMP and its Federation model components.

On a more detailed level, the testing team finds and records bugs and reports them to developers, using a web-based bug database. Once software developers repair these bugs, the testing team validates the newly written code to ensure that the bug fix was correctly implemented.

In addition to the algorithmic VV&A, MCS conducted an extensive variable exploration analysis. This is described in detail in Appendix H – Intertheater Analysis, but involved investigating over 30 different parameters/data inputs into MIDAS. Alternative values were examined to determine those, which the mobility analysis was especially sensitive to. The outputs were analyzed by subject matter experts to determine if the model was behaving correctly and whether the outputs reflected what would be expected operationally.

Considerable work was accomplished to verify the data transfer to MIDAS and validate the implementation of the GUI and model interface. However, the available documentation provides no discussion of the underlying conceptual model. Face validation of the software implementation, GUI interface, and resulting federation output was more than adequate. Two teams were tasked with running MIDAS under the AMP GUI. Only recently have they agreed concerning model configuration, model management, conceptual model documentation, and a common configuration control board to synchronize their results. During MCS 2005, the absence of these agreements contributed to different results achieved by the two study teams and could have introduced errors to the study.

ELIST VV&A excerpt and discussion

ELIST—an existing system that was further developed using the rapid prototype approach. Rapid prototyping is a "design-code-test" methodology in which there is close coordination between the user and developer. It allows for development of a model when requirements are not defined completely at the beginning. Users test new functionality and provide feedback to the developer with each new release of the model. In an ideal situation, VV&A is performed concurrently with the development of a new simulation. In the case of existing software such as ELIST, the VV&A process identifies any critical deficiencies in the existing system by examining the requirements, conceptual model, design, code, and model results. Solutions to these deficiencies, and future modifications and enhancements to the model can then be designed and implemented using the DMSO recommended practices and subject to the VV&A process

During the development of ELIST 8.0, a detailed VV&A test plan was developed. Users (both at SDDCTEA [Surface Deployment and Distribution Command Transportation Engineering Agency] *and the Force Projection Battle Lab Support Element at FT Eustis) tested the model using many different scenarios and at many different levels of detail. The tests were designed to root out inconsistencies in the behavior and results. Many of the tests included extracting detailed cargo and personnel movement data from the various ELIST Oracle tables, sorting and grouping the results in ways to get at inconsistencies in the model. These have included tests to verify the following types of methodologies used in the model: theater asset allocation for all modes of travel, container and cargo loading, convoy-generation, road and rail clearance, ship-berthing, and crane usage. Countless tests have been conducted to verify the functioning of the network, scenario, and ETPFDD* [Expanded Time Phased Force Deployment Data] *editors. Reports have been verified using results obtained from the underlying ELIST Oracle tables.*

During MCS, subject matter experts (SME) reviewed the model set-up parameters and analyzed outputs for consistency and reasonableness. Where limiting factors

(LIMFACs) were identified, SMEs investigated their reasonableness. Data inconsistencies were addressed through working group discussions.

Though VV&A efforts are mentioned associated with each new version, no documentation exists to demonstrate such efforts. Significant mention is made of work performed to check new scenarios akin to regression testing across a wide range of inputs and scenario data files. Though a short description of the ELIST features is provided, the conceptual model is not described nor is there any discussion of its validity. As with AMP-MIDAS, the user community performed considerable Face Validity checks on the results of model implementation, and MCS working groups extensively reviewed the results for MCS 2005.

JICM VV&A excerpt and discussion

JICM—was originally commissioned in the 1980s as RSAS (RAND Strategy Assessment System) and evolved into its present form in the early 1990's under the auspices of OSD Net Assessment.

When the US Army Center for Army Analysis (CAA) was considering adopting the model in 1999, they hired an independent firm, Coleman Research, to perform VV&A of the model. A report was delivered at the conclusion of their evaluation. After their evaluation, CAA was satisfied the model met their standards, and adopted the model as their principal land combat model, replacing several other campaign models with it. It has been continuously evaluated by CAA since its adoption. The USMC Combat Development Command contracted with Group W, to evaluate JICM for its use. An evaluation was delivered in early 2005.

The JICM discussion provides a short history of the model's use, there is no mention of the conceptual model, no mention of any attempt to V&V for MCS 2005, and no mention of even Face Validity checks of the results for MCS 2005 as a basis for accreditation. As previously mentioned, the review of results was so extensive during MCS 2005 that it is unlikely that JICM results were not reviewed in detail as part of the ongoing study effort. However, to date attempts to locate the documents mentioned in Appendix M have failed. Taken together, these two paragraphs underline the belief by many in the DoD analytic community that long-term application of legacy M&S maintained by the Services is sufficient for V&V. While many of the guidelines suggested in this paper may have been followed over time, the missing ingredient is specific documentation to that effect. Even a short explanation of what a previous study may have done to validate conceptual models and specific applications would provide increased confidence in the model's application in new studies, as well as guide additional V&V activities to ensure the M&S meets the need of its intended use.

THUNDER VV&A excerpt and discussion

THUNDER V&V Status:

Verification

All elements of the core THUNDER simulation have been verified at their inception through a rigorous process involving requirements-to-design tracing, walkthroughs and formal reviews of the code, component and integrated testing, and alpha and beta release test phases. In addition, significant modifications or enhancements incorporated in authorized baseline releases entail revisitation and evaluation of specific THUNDER submodels. This repeated scrutiny of THUNDER's infrastructure and logic should provide the foundation for confidence in the integrity of the simulation on the part of a prospective user. The frequency and extent of review for individual elements of the overall THUNDER simulation are illustrated in the version chronology shown below.

As with the earlier model discussions, the THUNDER description is also inadequate; again there is no mention of the conceptual model, no mention of any attempt to V&V for MCS 2005 and no mention of even Face Validity checks of the results for MCS 2005. The author does have first hand knowledge of the extensive attempts to verify the THUNDER implementation in software code mentioned above. No documentation exists to indicate THUNDER has completed a rigorous V&V process and/or Accreditation process. An extensive list of studies completed using THUNDER is provided in the MCS 2005 Final Report, Appendix M.

ITEM VV&A excerpt and discussion

ITEM—VV&A is accomplished in two basic ways when programming new capabilities into the model. First, all new conceptual algorithms are forwarded to all members of the ITEM steering committee for a review of methodology. Second, the lead contractor, SAIC, uses the CMMI Level 5 processes to control our software development. SAIC is ISO 9001-2000 registered, using these processes to document trouble reports and follow the changes through testing. SAIC has a standard set of test procedures and regression tests that executed before each release. These tests include running customer provided databases. These are all internally documented in our software library.

The ITEM model is complicated enough, however, that the steps above will not achieve a zero error rate in coding. The SAIC programmers address this through Beta test releases. In general, the member of the ITEM steering committee who sponsored or requested the change performs the Beta testing. Over the past two years, most of the changes have been in support of USPACOM, and USPACOM J8

has been doing the Beta testing. Problems identified in Beta test are corrected prior to final release of an update.

In OSD PA&E [Program Analysis & Evaluation] *SAC, a new version of the ITEM model is subjected to VV&A prior to acceptance by rerunning an analysis done in the previous version. The outputs of the old and new version of the model are compared using the video playback capability of the model, and by rerunning the key Measures of Effectiveness used for the old analysis for the output of the new model version. Differences are explored and explained prior to accepting the new ITEM model version. SAIC provides rapid analysis and code corrections for identified problems under the DTRA contract.*

There have been two VV&A studies on ITEM. Both conducted by Center for Naval Analysis (CNA). The first was done in 1994 and is out of date. In the fall of 2004, CNA did a top-level review of ITEM 8.6 and briefed results to CNO N81. Lyntis Beard (703-824-2860) at CNA was the principal author.

The ITEM discussion provides no mention of the conceptual model, no mention of any attempt to V&V for MCS 2005 and no mention of the Face Validity checks of the results performed for MCS 2005. The information provided by SAIC indicates a rigorous verification process and rigorous check of model results for a range of scenarios and mentions the use of regression testing techniques. Two VV&A studies are cited including a high-level review which was recently conducted. Again no copies of the ITEM VV&A reports were available or located at OSD/PA&E.

With respect to the warfighting models used in this study—JICM, THUNDER, and ITEM—several members of the Joint Staff J8/WAD indicated that "we do not perform VV&A—that is not our job, we simply rely on the Services to do it." Had the guidelines used by this report been implemented for MCS 2005, much of the V&V information the Services used would have been reviewed by the MCS analytic team. In those instances where specific applications may not have been addressed (such as mobility assumptions associated with warfighting results), additional V&V activities, to include the development and articulation of the conceptual model governing that aspect of the warfight, would have been developed, tested, and documented.

EADSIM VV&A excerpt and discussion

EADSIM—Verification

The general process model shown in Figure 1 is used to perform M&S verification activities. This process was adapted from DA Pamphlet 5-11 and other DOD guidance, and has been used effectively on previous efforts. This process first investigates the software design and verifies that it properly implements the developer's concept. The

verification against the design concept may include some activities defined as logical verification, but there may be other types of logical verification, such as sensitivity analysis, that may also be appropriate. Thereafter, the process transitions to a code review to ensure proper implementation of the design of the M&S.

Generally, the M&S design will be verified against the design concept just before the critical design review and before the signing of the version description document. Additional logical and code verification may be performed throughout the development cycle.

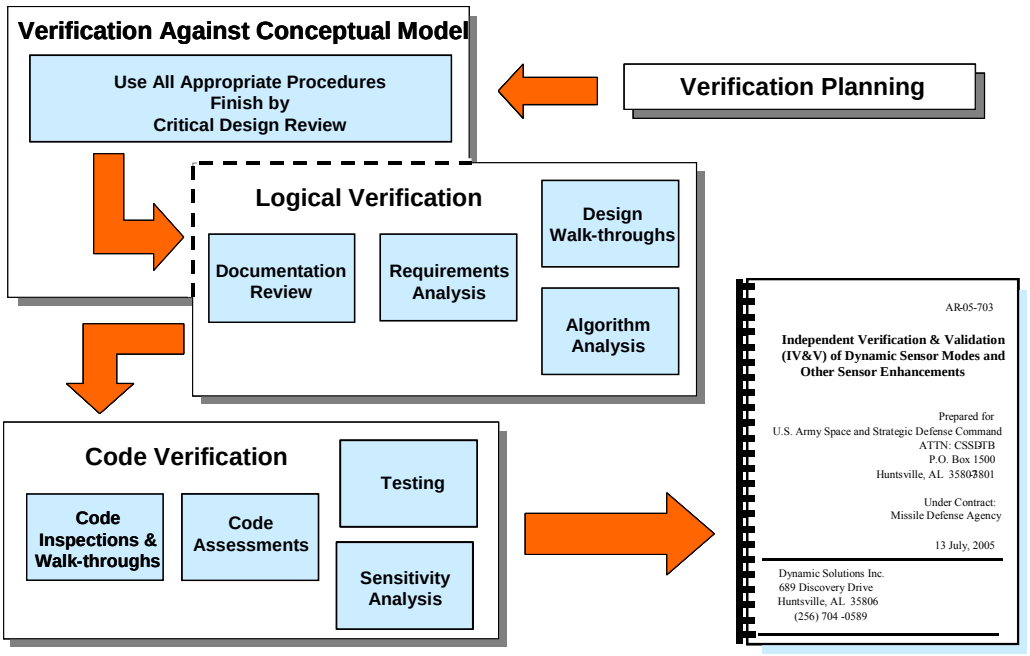

Figure 2. EADSIM Model Verification Process

Validation

According to DOD 5000.59, validation is the rigorous and structured process of determining the extent to which modeling and simulation accurately represents the intended "real world" phenomena from the perspective of the intended use of the model and simulation. Validation has two main components: structural validation and output validation (also called conceptual model validation and results validation). Structural validation focuses on the internal portion of the model and simulation which includes examination of model and simulation assumptions and review of the model and simulation architecture and algorithms in the context of their intended use. Output validation answers questions on how well the simulation results compare with the perceived real world.

The Validation Process used for is illustrated in Figure 2 and comprises four main tasks: (1) problem definition, (2) structural validation, (3) output validation, and (4) preparation of a validation report.

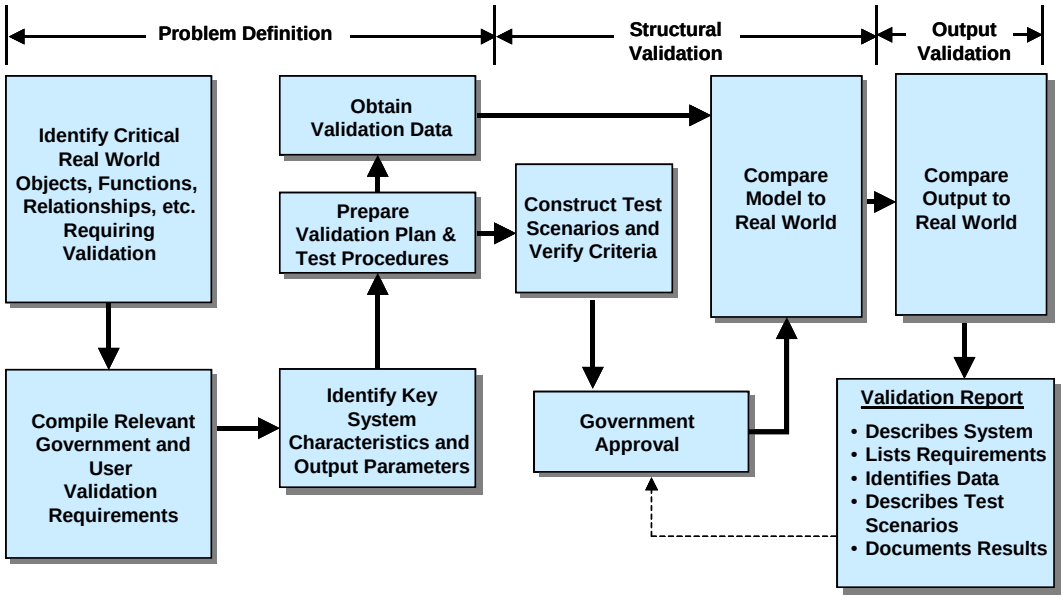

Figure 3. EADSIM Model Validation Process

The EADSIM discussion indicates that the US Army Space and Missile Command conducts a rigorous VV&A process in accordance with DA Pamphlet 5-11. Figure 3 describes a validation of the conceptual model, a rigorous verification of the code in comparison to the conceptual model and Face Validity checks with verification and system criteria specified. The EADSIM discussion indicates it is normal for a Validation Report to be produced during the version control process. As with the other models, the author is unable to verify that available EADSIM reports were reviewed by OSD/PA&E personnel or that any of the guideline items were met as part of MCS 2005.

ARCEM V&V excerpt and discussion

ARCEM—ARCEM tanker performance modeling was compared to flight plans from the Portable Flight Planning System, a tool used by operational aircrews to plan their missions. Fuel consumption for comparable mission profiles was within 1%. For MCS, ARCEM was cross checked against the Employment Mating and Ranging Program (EMARP), a sub component of CMARPS and results were comparable for representative days from the MCS employment analysis. ARCEM average sortie duration was higher than EMARP (17%), but the number of tankers used was less than EMARP for an equivalent number of tanker sorties (-4%).

ARCEM has also been used to model air tasking orders (ATOs) from Operation Iraqi Freedom with very close (within 5%) agreement in the number of tanker sorties needed to support the ATO.

The ARCEM discussion contains no information about the conceptual mode or verification procedures to demonstrate the software implementation is adequate. It does cover checks made with CMARPS, another model used in the study.

In summary, discussion of the conceptual models underlying the M&S used in MCS 2005 was almost non-existent. However, the analysts involved in this study had considerable experience with mobility analyses and with the models used to perform MCS 2005. They were well aware of the models' and simulations' performance characteristics, previous usage, and lineage. Users knew which models should address the MCS requirements, those models were selected, and the study was completed. Though considerable work is apparent in the areas of verification and validation of the M&S results, most mention is made of previous attempts to verify the software code by other parties outside the MCS 2005 study team. Documentation to support claims of previous work particularly by the Services is almost non-existent. Considerable evidence exists that demonstrates Face Validity checks were made at every level of the MCS 2005 effort.

5. Verify and Validate the Data (Identify the POC)

- Identify the authoritative source for the data
- Verify the internal consistency of the data, based on the intended purpose; were they collected consistently (conditions, limitations, etc.) with the intended purpose of the M&S?
- Validate the data for use in this M&S, based on the intended purpose
- Were the data compared with corresponding known, real-world or best-estimate values?
- If there are transformations of the data, how do you know the aggregated, disaggregated, or transformed data are accurate for the intended purpose?
- If there is a federation of models exchanging data, how do you know the data exchange is performed correctly and the assumptions between models/simulations are consistent?

Appendix M of the MCS 2005 contains a section discussing the data sources which is reproduced here:

Data Sets used in MCS

This component provides metadata on the data sets used in MCS and their accreditation processes. These data sets range from department-wide data sources to study-specific excerpts and include previously-approved data sets from accredited analytical baselines.

Future Forces Data Base, (FFDB)

- Proponent: OSD-PAE JDS [Joint Data System]
- Force Year Horizon: 7 years
- Sources: Services
- Updates: Semi annually
- Time since Inception: 2 years
- Vetting method: Common DOD user website
- Comment: DOD standard for US FY Forces, Units and equipment configurations

Joint Country Force Assessment, (JCOFA)

- Proponent: NGIC (National Ground Intelligence Center)
- Force Year Horizon: Current year plus 20 years (i.e., 2005-2025)
- Sources: Intelligence Production Centers (DIA, MCIA, MSIC, NASIC, NGIC, ONI, etc.)
- Vetting Method: NGIC website
- Updates: Every two years for priority countries
- Comment: DOD standard for projected future Red and Green Order of Battle information. Database consists of a scenario-independent 20-year forecast of air, ground, marine, naval, IADS, and space force structure and equipment. Narrative modules (economic, technology, equipment modernization, etc.) provide context for analysis included in database and equipment configuration tables.

Mobility Planning Factors Database, (MPFD)

- Proponent: USTRANSCOM J5/A
- Force Year Horizon: Future year 2012
- Sources: USTRANSCOM Components Commands
- Updates: Annually/as required to support MCS
- Vetting method: Through USTRANSCOM J5/A to OSD PA&E and JSJ4 for review.

- Comment: Selected port capabilities are projected for Out year; provides Militarily useful capacities for sea ports.

Integrated Scenario TPFDDs and TUCHA products

- Proponent: Joint Staff J4 and OSD PAE PFD (depending on scenario)
- Force Year Horizon: FY2012
- Sources: Services
- Updates: Throughout the study
- Vetting method: MCS Study website (JDS), JS J4 FTP site
- Comment: Key informing document and data input for planned deployment of unit equipment, personnel, and sustainment. Developed by each service, integrated at JS J4 and OSD PAE SAC. Processed also at SDDC-TEA. Derived from OA-04 and OA-05 study products.

MCS Mobility Data Assumptions Appendix.

- Proponent: MCS study Director (OSD PA&E/Joint Staff J4)
- Force Year Horizon: Current Study
- Sources: Mobility and Warfight Communities
- Vetting method: Study appendix, Study Website
- Updates: Compiled for each study
- Comment: Compilation of study-specific parameters, aggregation values and context for key assumptions made throughout MCS. Provides baseline Study data only.

Analytical Baseline data products.

- Proponent: Joint Analytical Data Steering Committee
- Force Year Horizon: Current and Future
- Sources: SPG-directed Department studies
- Vetting method: Formal Coordination for approval (see section 2.4.2)
- Updates: As directed by JADMSC
- Comment: These included both model data sets and source data sets

MCS 2005 data discussion

Each of the major data products is identified with a proponent, a general list of sources, a list of methods for vetting the data, methods for updating the data, and comments on the quali-

ty of the data. Elsewhere in the MCS 2005, there is discussion of the Analytical Baseline and Multi-Service Force Deployment (MSFD) processes that produce much of the scenario data, information about the working group structure used to produce the Concept of Operations and vet all the data used in the study. Additionally, a 65-page appendix covering all known data assumptions and limitations is included. The study meeting notes provide further evidence that the data were extensively reviewed at all stages during the MCS 2005.[11]

The GAO took exception to the MCS results based on model and data deficiencies. In particular, the GAO noted numerous references in the MCS 20005 final report to deficiencies in the data and caveats to the results based on those deficiencies. An example from their summary is useful to consider:

> *Data Collection. MCS analysis revealed several deficiencies in operational data collection, which hindered analysts' ability to assess system performance. Procedures for improving collection and retention of operational data must be identified and implemented. This effort includes collecting and retaining data dealing with dimensions and weight of cargo, identity of capability packages/units moved, mode selection criteria and rationale, timeliness metrics, and denied movement requests (regrets).*

While the DoD concurred with many of the comments concerning the improvement of data collection, the GAO did not seem to fully understand campaign analysis and data used to assess future situations. Since these events have never occurred, there is no real-world, historical data available for comparison. The data used in the MCS 2005 largely represent the DoD's best estimates and are produced by well-documented processes. The processes rely on the judgment of subject matter experts and were vetted by several panels of senior reviewers. Where possible, the MCS team noted the problems and performed additional sensitivity analysis to ensure they knew the impact of relying on data that might be suspect. These were reviewed by several levels of senior decision-makers.

If there is a single GAO criticism that might be accurate, it may lie with the ability of any campaign-level analysis to guarantee that all of the data used were collected in accordance with the intended purpose of the study (defined by the Objectives of MCS 2005). Most of the data—especially those captured by the MSFD and Analytic Baseline processes—meet these criteria.

The data are archived in the JDS registry and available for examination or reuse if needed. The study team adequately dealt with the issues related to transformations and federating AMP and MIDAS. Extensive sensitivity analysis and regression testing occurred to ex-

---

[11] Captured by Mr. Steve Ross, OSD/PA&E; interview with the author, Washington, DC, 4 April 2008.

amine study data issues. All 65 pages of Appendix B of the MCS 2005 discuss the data assumptions, limitations, and constraints in considerable detail. Improved reporting of the data V&V efforts would not likely have occurred within specific V&V reports and an Accreditation report discussed in item 7. In the author's opinion, the MCS team's lone failure was not concluding the data limitation appendix with the statement "The data is the best available."

Considering the enormity of the MCS task, data issues were handled excellently and exceed the guideline requirements in this area.

6. List the known assumptions, limitations, and constraints for the M&S and/or M&S federation.

While there were considerable attempts to address issues related to the data, the MCS 2005 final report contains one section that summarizes problems related to the models, excerpted here:

> *Analysis Tools. MCS analysis revealed several deficiencies in existing mobility models. To facilitate analyses of evolving logistics issues, modeling tools should be enhanced to:*
>
> - Permit better understanding of C2 impacts on system efficiency.
> - Better depict POL deployment system.
> - Synchronize units' forward movement and sustainment/ammunition resupply during offensive operations.
> - Account for random/stochastic events (e.g., weather, diplomatic clearances, aircraft breakdowns).
> - Allow for assets to move between inter-theater and intra-theater roles.

The Executive Summary also contains a single paragraph rehashing these same deficiencies with the models:

> *Mobility Models. MCS analysis also revealed several deficiencies in existing mobility models. Models should be enhanced to permit better understanding of command and control impacts on system efficiency, better depict the deployment system for petroleum, oils, and lubricants (POL), provide improved synchronization of units' forward movement and resupply during offensive operations, and better account for random/stochastic events (weather, breakdowns, etc.)*

Considering the number of M&S models used and the known issues of each particular model, this discussion is far from adequate. M&S limitations and assumptions should have merited its own appendix to justify the use of these results. In particular and as the GAO later notes,

there is no discussion justifying the innovative analysis technique linking mobility models and their metrics to the warfighting models and their outputs. At a minimum, this area should have merited the same attention to detail that was observed to address data issues. Had the study team produced a final V&A report for MCS 2005, that document would have served this purpose well.

7. Sponsor decides whether to Accredit the M&S and associated data.

- *Describe the requirements addressed by the model, simulation (to include an M&S federation if appropriate), and associated data.*
- *What V&V activities were performed?*
- *What are the acceptability criteria? (i.e., why should the Sponsor believe the results from the M&S and its associated data)*
- *What are the known or projected risks (uncertainty and consequences) associated with the use of the M&S and their data based on the intended purpose?*
- *Has an accreditation report been completed and provided to the study Sponsor?*
- *Has the Sponsor reviewed this info and accredited the model and data for this use?*

There was no attempt to accredit MCS 2005. Because of the analysts' experience with the models and simulations selected for this study, M&S requirements for MCS 2005 were not specified. The MCS study objectives were clearly specified and model requirements or acceptability criteria can be inferred from these requirements. Simple decision rules based on analyst experience determined which was the correct tool to apply (e.g., for campaign-level, ground combat use JICM; for campaign-level, air combat, use THUNDER). No acceptance criteria are discussed in MCS 2005 documentation. Because it was assumed that the use of legacy simulations with a long history of use within DoD for similar purposes was sufficient, there was little to no documented conceptual model discussion or validation.

Considerable evidence exists for software verification activities and for validation of model results. Sensitivity analysis was performed to address many of the risks from using the MCS data, metrics, and suite of models. Of the 16 techniques identified in Appendix A, evidence exists to demonstrate that only two were not widely used throughout this study. Prediction is not possible where the purpose is to examine future scenarios, and no techniques emphasizing prediction were used. Though Subject Matter Experts were widely used in MCS 2005, the Turing Test approach was never used as a V&V technique in MCS 2005. All other

techniques were widely used. However, there was no explicit evaluation of the risks for using these models nor were such risks mentioned in the MCS final report.

Finally, the MCS 2005 study team produced no specific V&V or Accreditation reports nor was the study Sponsor asked to accredit the models and data for use in this activity. Appendix M to the MCS 2005 final report was produced to help address the GAO's concerns and is the single MCS document addressing VV&A issues.

D. Summary

MCS 2005 had clearly stated objectives that could have formed the basis for requirements for the M&S tools selected for this study. Validation efforts of the conceptual models were not documented, while verification efforts for the software codes appear to have been strong and could have been leveraged. Results validation and data V&V appear to have been very strong. Configuration management of the models (except for AMP-MIDAS as noted) met all Departmental guidelines. Documentation of these efforts was the weakest aspect of the V&V effort. As a result, accreditation was assumed and not pursued.

The following items are suggested to fill the gaps revealed by the checklist discussion of MCS 2005:

1. Validate and document validation of the conceptual models of the M&S and ensure they are appropriate for the Sponsor's intended purpose. A review of the M&S requirements based on the study objectives and a review of the conceptual models selected for the study would not have changed which models were used. However, conducting such a review and documenting the results might have made the team more aware of any limitations of the model suite as they conducted the study.
    - For instance, a review of the M&S' conceptual models shows they assume perfect command and control (C2) and this perfect C2 is applied to handling of the assets in the mobility/logistics system. Knowledge of this limitation of the M&S suite leads the author of this use case paper to believe the analysis from MCS 2005 typically represents a "best case analysis." Some participants in MCS 2005 were fully aware of this limitation and this guided the innovative feedback loop analysis performed in MCS 2005; whereby after intra-theater, mobility models determined arrival schedules; warfighting models determined outcomes and usages of key stockpiles of ammo, POL, food and water; inter-theater, mobility models were run to ensure delivery profiles could meet warfighting demands; warfighting models were rerun to determine enemy leaker rates and attrition to BLUE stockpiles; and finally,

inter-theater, mobility models were rerun to determine final delivery profiles and estimate mobility/logistics shortfalls.

- While impressed with the innovative approach taken in MCS 2005, the GAO was concerned that there was a lack of *"clear understanding of the direct relationship of warfighting objectives to transportation capabilities."*[12] Within the MCS 2005 report, there is no discussion to justify the innovative analysis technique linking the mobility model outputs as inputs to the warfighting models and use of the warfighting model outputs as metrics to measure the effectiveness of the mobility/logistics system. Discussions with some MCS 2005 participants indicate they viewed this use of the models as a system as common sense and obviously correct.
- A white paper or appendix should have been produced to discuss the innovative analysis technique linking transportation capabilities to the warfighting model metrics. The paper's purpose would be to generate the necessary discussion and permit the conceptual model validation of this innovative analysis technique or the revelation that the technique is flawed. If flawed, this discussion might reveal appropriate alternatives or improvements. This aspect of the study is not covered by any known Service efforts to V&V their models. Some MCS 2005 participants were unclear as to how this innovative approach was used to solve the mobility-to-warfigth linkage problem. In light of the lack of conceptual model validation documentation and this confusion, this GAO comment (italicized above) appears to have some merit.

2. Document the V&V effort as it progresses. A lot of excellent work was done during the MCS 2005 study and documenting this work at the time would have provided a better basis for responding to audits as well as improving future M&S applications.
3. Execute an Accreditation process, define acceptance criteria, and evaluate the risk of the M&S use.
4. Document the Accreditation effort via a report covering the items in Guideline 7 (subsection 7 above).

Requiring an accreditation decision and producing the documentation to support it would have highlighted the excellent work performed in this study. At the same time, such a documented could have illuminated any problems and enabled them to be addressed during the study. It likewise would have better informed the Sponsor and his/her staff regarding key issues associated with the representations used within the study.

---

[12] "Study Limitations Raise Questions," September 2006.

Appendix A: Acronyms & Abbreviations

ACC	Air Combat Command
ATO	air tasking order
C2	command and control
CAA	US Army Center for Army Analysis
CNA	Center for Naval Analysis
DIA	Defense Intelligence Agency
DMSO	Defense Modeling and Simulation Office
DoD	US Department of Defense
EMARP	Employment Mating and Ranging Program
ETPFDD	Expanded Time Phased Force Deployment Data
GAO	US General Accountability Office
GUI	graphic user interface
IADS	Integrated Air Defense System
JDS	Joint Data System
M&S	modeling and simulation, or models and simulations
MCIA	Marine Corps Intelligence Activity
MCS 2005	Mobility Capabilities Study
MORS	Military Operations Research Society
MSFD	Multi-Service Force Deployment
MSIC	Missile and Space Intelligence Center
NASIC	National Air and Space Intelligence Center
NGIC	National Ground Intelligence Center
ONI	Office of Naval Intelligence
OSD	Office of the Secretary of Defense
PA&E	Program Analysis & Evaluation
POC	point of contact
POL	petroleum, oils, and lubricants
SAC	Strategic Air Command
SDDCTEA	Surface Deployment and Distribution Command Transportation Engineering Agency
SME	subject matter experts
TPFDD	Time Phased Force Deployment Data
TUCHA	Type Unit Characteristics
US TRANSCOM	Transportation Command
VV&A	verification, validation and accreditation

THIS PAGE LEFT INTENTIONALLY BLANK

Appendix B: References

2005 Mobility Capabilities Study (MCS, 2005), Office of the Secretary of Defense, Program Analysis and Evaluation OSD/PA&E, December 2005.

Balci, Osman, "Verification and Validation and Testing of Models," *Encyclopedia of Operations Research and Management Science* (Netherlands: Kluwer Academic Publishers, 2001).

Box, George E.P. and Norman R. Draper, *Empirical Model-Building and Response Surfaces* (New York: Wiley Press, 1987).

Box, George E.P., "Robustness in the Strategy of Scientific Model Building," in *Robustness in Statistics*, R.L. Launer and G.N. Wilkerson, eds. (New York: Academic Press, 1979).

Department of the Army Pamphlet 5-11, Verification, Validation, and Accreditation of Army Models and Simulations, 30 September 1999.

"Issues Concerning Airlift and Tanker Programs," US General Accountability Office, GAO-07-566T, 7 March 2007.

Palmore, Julian I., ed., "Simulation Validation (SIMVAL) 1994," Mini-Symposium Report (Alexandria, VA: Military Operations Research Society, 1994).

Sargent, Robert G., "Validation and Verification of Simulation Models," *Proceedings of the 2004 Winter Simulation Conference*, R.G. Ingalls *et al.*, eds., 2004.

"Study Limitations Raise Questions about the Adequacy and Completeness of the Mobility Capabilities Study Report," US General Accountability Office, GAO-06-938, September 2006.

Verification, Validation and Accreditation (VV&A) Recommended Practices Guide, Office of the Director of OSD/DDRE/DMSO, November 1996.

THIS PAGE LEFT INTENTIONALLY BLANK

Appendix C. Examples of V&V Methods

A number of methods can be used to perform V&V tasks. The following list is compiled from an article by Sargent and by referring to the Department of the Army Pamplet 5-11.[13]

Activity	Description	Techniques, Methods or Examples
Animation (Visualization)	The model's operational behavior is displayed graphically as the model moves through time	▪ Example: Movements of parts through a factory during a simulation run are shown graphically.
Comparison to Other Models	Various results (e.g., outputs) of the simulation model are compared to results of other valid models	▪ Ex: Simple cases of a simulation model are compared with known results of analytic models; simulation model is compared with other simulation models that have been validated.
Degenerate Tests	The model's degeneracy behavior is tested by appropriate selection of values of input and internal parameters	▪ Ex: Tracking whether the average number in the queue of a single server continues to increase over time when the arrival rate is larger than the service rate.
Event Validity	"Occurrences" in the simulation model are compared with those of the real system to determine similarity	▪ Ex: Number of fires in a fire department simulation compared to historical fires.
Extreme Condition Tests	Check model structure and outputs for plausibility for extreme and unlikely combinations of factors	▪ Ex: If in-process inventories are zero, production should be zero.
Face Validity	Asking individuals knowledgeable (SMEs) about the system whether the model and/or its behavior are reasonable	▪ Ex: Assure the logic in the conceptual model correct and the model's input-output relationships reasonable.
Functional Decomposition	Validate the whole based on testing the parts, see DA PAM 5-11, VV&A of Army M&S.	
Historical Data Validation	If historical data exist (or if data are collected on a system for building or testing a model), use part of the data to build the model and the remaining data to determine (test) whether the model behaves as the system does	▪ This testing is conducted by driving the simulation model with either samples from distributions or traces.

continued

[13] Sargent, "Validation and Verification of Simulation Models"; *Department of the Army Pamphlet 5-11, Verification, Validation, and Accreditation of Army Models and Simulations, 30 September 1999.*

Activity	Description	Techniques, Methods or Examples
Historical Methods	The three historical methods of validation are *rationalism*, *empiricism*, and *positive economics*.	▪ Rationalism assumes that everyone knows whether the underlying assumptions of a model are true. Logic deductions are used from these assumptions to develop the correct (valid) model. ▪ Empiricism requires every assumption and outcome to be empirically validated (reliance on data from experiments or observation). ▪ Positive economics requires only that the model be able to predict the future and is not concerned with a model's assumptions or structure (causal relationships or mechanisms).
Internal Validity	Make several replication (runs) of a stochastic model to determine the amount of (internal) stochastic variability in the model	▪ A large amount of variability (lack of consistency) may cause the model's results to be questionable and, if typical of the problem entity, may cause questions about the policy or system being investigated.
Multistage Validation	Naylor and Finger (1967) proposed combining the three historical methods of rationalism, empiricism, and positive economics into a multistage process of validation. This validation method consists of: • Developing the model's assumptions on theory, observations, and general knowledge • Validating the model's assumptions where possible by empirically testing them • Comparing (test) input-output relationships of the model to the real system	
Operational Graphics	Graphically show how values of various performance measures change, e.g., the number in queue and percentage of servers busy, as the model runs through time	▪ Ex: Visually display the dynamic behavior of performance indicators as the simulation model runs through time to ensure they are correct.
Parameter Variability - Sensitivity Analysis	Change input and internal parameter values to determine the effect upon the model's behavior or output and compare to the real system	▪ Parameters that are sensitive, i.e., cause significant changes in model behavior or output, require reasonable accuracy prior to model use (which may require iterations in model development). ▪ Example of Regression testing: Run the model across a wide parametric range for the key variables. Examine the boundary conditions, see DA PAM 5-11, VV&A of Army M&S.
Predictive Validation	Use the model to predict (forecast) system behavior, and then compare the actual system's behavior to the model's forecast to determine if they are the same	▪ System data may come from an operational system or be obtained by conducting experiments or field tests on the system

continued

Activity	Description	Techniques, Methods or Examples
Traces	Follow the behavior of different types of specific entities through the model to determine if the model's logic is correct and if the necessary accuracy is obtained	
Turing Tests	Ask individuals who are knowledgeable about the operations of the system being modeled if they can discriminate between system and model outputs	

THIS PAGE LEFT INTENTIONALLY BLANK

REPORT DOCUMENTATION PAGE	Form Approved OMB No. 0704-0188

1. REPORT DATE (DD-MM-YY)	2. REPORT TYPE	3. DATES COVERED (From – To)
January 2009	Study (Final)	July 2008—December 2008

4. TITLE AND SUBTITLE	5a. CONTRACT NO.
Mobility Capabilities Study 2005: Use Case Study	DASW01-04-C-0003
	5b. GRANT NO.
	5c. PROGRAM ELEMENT NO(S).

6. AUTHOR(S)	5d. PROJECT NO.
Dr. Jack Jackson	
	5e. TASK NO. BA-8-2911
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	8. PERFORMING ORGANIZATION REPORT NO.
Institute for Defense Analyses 4850 Mark Center Drive, Alexandria, VA 22311-1882	IDA Document D-3710 Log no. H 09-000001

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR'S / MONITOR'S ACRONYM(S)
Director of Study and Analytical Support Division Office of the Director Program Analysis and Evaluation 1800 Defense Pentagon Washington, DC 20301-1800	OSD(PA&E) and MSCO
	11. SPONSOR'S / MONITOR'S REPORT NO(S).

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited. 3 February 2011.

13. SUPPLEMENTARY NOTES

Unclassified

14. ABSTRACT

This use case study is a written assessment of the Department of Defense's (DoD's) Mobility Capabilities Study 2005 (MCS 2005), based on the General Accountability Office's (GAO's) evaluation. The GAO identified several deficiencies that undermined the credibility of the analysis and found that DoD practice did not match DoD policy. In response, the DoD drafted new guidelines for the Verification, Validation, and Accreditation (VV&A) of models and simulations. This document is an application of those guidelines to evaluate the VV&A practices in MCS 2005 after the fact.

15. SUBJECT TERMS

Mobility Capabilities Study 2005; modeling and simulation; models and simulations; M&S; verification, validation and accreditation; VV&A

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NO. OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE	UU	47	Mr. James Bexfield
U	U	U			19b. TELEPHONE NUMBER (Include Area Code) (703) 695-7945

THIS PAGE LEFT INTENTIONALLY BLANK

The Institute for Defense Analyses is a non-profit corporation that administers three federally funded research and development centers to provide objective analyses of national security issues, particularly those requiring scientific and technical expertise, and conduct related research on other national challenges.