DTIC FILE COPY

(2)

DIRECTOR
OPERATIONAL TEST & EVALUATION

AD-A210 796

89 8 02 015

DIRECTOR, OPERATIONAL TEST AND EVALUATION

POLICY

FOR THE APPLICATION OF

MODELING AND SIMULATION

IN SUPPORT OF OPERATIONAL TEST AND EVALUATION

24 JANUARY 1989

DISTRIBUTION "A
PER: DPS SAUNDERS of OSD (OT&E)
WASHINGTON, DC 20301

INTRODUCTION

Purpose:

This document provides focus and guidance in the critical area of modeling and simulation. As weapon systems become more complex and expensive to test, reliance on modeling and simulation will increase. Models and simulations are different from, and must, by definition, be used in support of operational test and evaluation (OT&E); and when models and simulations are used for such support, special care is necessary to ensure the results have validity. This policy establishes a framework for finding ways for modeling and simulation to effective y complement actual field testing.

Background:

A primary definition of operational test and evaluation (OT&E) is found in Title 10 U.S.C. Section 138. That section of public law established the Director of Operational Test and Evaluation (DOT&E) in the office of the Secretary of Defense (OSD) and defined OT&E as:

> "the field testing, under realistic combat conditions, of any item of (or key component of) weapons, equipment. or munitions for the purpose of determining the effectiveness and suitability of the weapons, equipment, or munitions for use in combat by typical military users; and the evaluation of the results of such tests."

The wording of that section of Title 10 precludes modeling and simulation (M/S) from being included as OT&E. However, M/S can have an important role in OT&E, particularly when a comprehensive evaluation of many modern weapon systems based on only field testing is not possible. Constraints against testing, such as cost, security, safety, ability to portray threat capabilities, limitations on use of the full electromagnetic spectrum, test instrumentation, treaty constraints, available test time, number and maturity of test articles, test maneuver space, and representative terrain and weather may combine to preclude a complete evaluation of a system through field testing alone. In some technically complex systems, particularly where relatively small quantities are to be procured, insights into expected operational capabilities are needed before the complete system is available for testing. This is particularly valid in concurrent programs where all components of a weapon system (i.e., combat system(s), support elements, representative manpower) are not available but decisions must be made concerning a commitment to some degree of production. It is, therefore, also imperative that avenues be explored to provide a better understanding when full-scale testing is not possible.

Operational test and evaluation has been described as a total process, which, to be effective, must use analytical studies, systems analysis, component testing, and eventual testing of the actual weapon system. Models and simulations are tools which can potentially augment and/or complement actual field tests and provide decision makers necessary information to assess the progress of a system toward fulfilling the operational needs. As an adjunct to actual field testing, those tools can provide valid, credible, and timely operational effectiveness and suitability insights which would not otherwise be available. Consequently, it is appropriate to utilize them as an aid to OT&E planning, as an evaluation tool prior to the availability of a complete system, and to augment, extend, or enhance actual field test results, provided appropriate discussion is included in the reports. While it is true that M/S has limitations, it is also true that operational test and evaluation (OT&E) has limitations in that the number of relevant combat environments that can be addressed in OT&E is very limited. Together, models and simulations offer the ability to expand the operational assessment, while field testing offers a means to more effectively calibrate and validate the models.

The following definitions of modeling and simulation are found in DoD 5000.3-M-1, "Test and Evaluation Master Plan (TEMP) Guidelines":

> <u>MODEL</u>: A representation of an actual or conceptual system that involves mathematics, logical expressions, or computer simulations that can be used to predict how the system might perform or survive under various conditions or in a range of hostile environments.
>
> <u>SIMULATION</u>: A method for implementing a model. The process of conducting experiments with a model for the purpose of better understanding the behavior of the system modeled under selected conditions or of evaluating strategies for the operation of the system under selected conditions or limits imposed by the development of operational criteria.

This document provides guidance on decisions to use modeling and simulation (M/S) as part of the OT&E process, and M/S application, development, and reporting considerations.

MODELING AND SIMULATION APPLICATIONS IN OT&E

Modeling and simulation, as analytical tools, have been used in military applications for many years (cost and operational effectiveness analyses, wargames, etc.), but their use in the operational test and evaluation process has been limited. While not replacements for actual weapon system tests (to include hardware, software, and operators) that are feasible within reasonable resource constraints, there are areas within the OT&E process for which these tools are appropriate. Such areas include test planning, test data, analysis and evaluation (to augment, extend, or enhance test results), system simulation, development of system tactics and techniques of employment, and early operational assessments of expected capabilities (with the incumbent risk of inexact modeling of both the physical reality and system interactions).

While actual field testing is the preferred primary data source for operational evaluations, it may be only a partial replication of the expected wartime environment. It is normally not possible to complete hardware testing in all operational environments, with all force interactions instrumented to provide sufficient data, to determine, with certitude, system combat performance. Therefore, most field tests are only partial representations of the total operational environment. Other representations of the "real world" provide flexibility, precision, and scope not found in field tests.

Given that a decision has been made to consider using M/S results to support the OT&E of a specific system, certain elements of caution should be exercised. For example, if used to augment, extend, or enhance field test results, a definitive statement should be prepared, delineating which evaluation issues, or parts thereof, will be addressed by the simulation effort. It is important that both the implicit and explicit assumptions used in a M/S be carefully documented; and that the risks of inexact modeling (both of the physical reality and system interactions) be well understood.

In planning any operational test, it is important to understand which elements of system performance are the "drivers" in assessing whether or not the system meets the user's requirements. Models may be utilized to assist in the identification of those "drivers" which should be verified by field tests.

Possessing validated intelligence documentation on the threat environment is also important in planning for an operational test. By using the current DIA-validated System Threat Assessment Report (STAR) for a particular weapon system and focusing on the Critical Intelligence Parameters listed therein, planners will be assured of developing a model and/or simulation that correctly addresses a weapon system's threat "drivers."

When models are used to complement field testing, or provide early insights into system capability, it is important that the questions to be addressed are clearly defined and related to the critical operational issues. In selecting a model to address a specific question, the selected model should be no more complex or detailed than needed to address the question.

Decisions concerning the use of M/S to support the OT&E process should be made early in the material acquisition cycle so as to support timely development of the new M/S. Ideally, the user, developer and tester would agree on the model(s) or simulation(s) needed to provide operationally-oriented assessments for a system under consideration not later than Milestone I. A plan should be developed at this time to transition the generic modeling capability normally available at Milestone I, to a mature, high fidelity model at Milestone III. It will not be necessary, in all cases, to develop new models or new simulations. Use or adaptation of existing M/S may be more appropriate and cost effective. Whatever the source, M/S should be consistently updated, verified, and validated with test data, field or bench measurements, and analyses to enhance predictions of real world capabilities. Only then can a model or simulation be accredited for use in support of the OT process.

The extent of M/S use in support of OT&E will be a function of the acquisition phase, weapon system, cost, and availability of the supporting resources, most notably the availability of input data. Use of a "balanced" mix of M/S and actual field testing should be addressed during the earliest stage of the acquisition cycle, and any such plans must be briefly described in the Test and Evaluation Master Plan (TEMP), if M/S results are to be used to augment, extend, or enhance field test results. As a general rule, any plans or requirements for the use of M/S in support of operational test and evaluation not addressed, <u>in advance</u>, in the TEMP will not be approved by the Director, Operational Test and Evaluation (DOT&E). Exceptions to this rule will require strong justification. In addition, any M/S plans or requirements not addressed in the current TEMP shall be considered with and included in a TEMP update as soon as practical. Some examples of applications of M/S in support of OT&E are summarized in Appendix A.

MODEL DEVELOPMENT PROCESS LEADING TO CREDIBILITY

The credibility of M/S results is a fragile commodity. Credibility, as applied to the M/S processes and results, is a combined impression of the inputs, processes, outputs, conclusions, the persons or agencies involved, and the strength of the evidence presented (see Appendix B). To be of use to decision makers, M/S results must be credible, and the process for planning, executing, and reporting on the development and use of a M/S should be very similar to that for a field test, with the added requirement to provide an audit trail (traceability, end-to-end) to allow an assessment of its credibility. It is imperative that any anticipated use or development of M/S to support the operational test and evaluation process be documented in the TEMP by reference to the verification and validation (V&V) plan. If changes to the M/S are planned after the TEMP is approved, then the annual updates to the TEMP should address the revised verification plan.

The following points should be considered throughout the M/S development, implementation, and review processes:

- <u>Acceptability of the M/S approach</u> by having decision makers involved early (and updated periodically) on a formal or informal basis.

- <u>Confidence in the model</u>, based on a sound, coherent, systematic process used in development; sound model management structure, including configuration management; model descriptions, including usage, strengths and weaknesses, past history, adequate documentation; thorough description of verification and validation efforts; threat representation and usage; and thorough descriptions of accreditation efforts.

- <u>Confidence in the M/S team</u>. M/S practitioners must be experienced with the simulation being used and with the system being simulated. That part of the team tasked with establishing the confidence in the model must be independent of the developers and users of the model or simulation.

- <u>Confidence in M/S methodology and use</u>, based on applicability or appropriateness to address requirements and issues under consideration, adequately described methodology and assumptions, certified and documented input data (including scenarios).

- <u>Confidence in M/S results</u>: M/S results should be consistent with actual field test results when the M/S input data used are representative of actual field test conditions. Further, M/S results should be consistent with other M/S results when their input data are comparable. Any M/S results that appear counter-intuitive must be fully investigated in order to determine whether or not the results are in error or if the results actually reflect some new insight not previously anticipated. M/S reported results must include a description of related OT&E field tests, apparent inconsistencies, and any available resolution of issues identified.

- <u>M/S Verification</u>: Verification is the process of determining whether a computer program or a simulation model performs as intended. A verification plan must be prepared for M/S planned for use in operational test and evaluation. This plan must be referenced in the weapon system TEMP. For new and modified simulation models, the verification plan must describe the verification process(es) and the documentation for reporting verification results. For existing simulation models, previous verification efforts which led to accreditation, if any, must be referenced.

- <u>M/S Validation</u>: Validation is the process which, as a minimum, addresses the following primary concerns: (1) the appropriateness of the model to adequately answer the questions or issues under study; (2) the degree of confidence in the conclusions that can be drawn from the M/S results; (3) the appropriateness of the threat data and threat tactics used in the model; and (4) model consistency. A M/S is appropriate if it addresses the critical issues and the supporting measures of effectiveness (MOEs), and if the M/S is a realistic representation of the weapon system and its operational environment. M/S appropriateness depends on the modeling techniques, assumptions and limitations, the input sources and quality, the ability to measure performance, and the design of the experiment. Confidence in M/S results can be enhanced by comparison to other data, e.g., actual test results, other models, or historical data. The sensitivity (driving and limiting factors) should be well understood and documented. Plans to recalibrate, reverify, and/or revalidate models and simulations based on actual test results should also be documented and implemented.

- <u>M/S Accreditation</u>: Accreditation is the process of certifying that a computer model has achieved an established standard such that it can be applied for a specific purpose.

GUIDELINES FOR THE DEVELOPMENT, APPLICATION, AND DOCUMENTATION OF M/S TO SUPPORT OPERATIONAL TEST AND EVALUATION

The following guidelines are applicable to the process of developing, applying, and documenting M/S to support OT&E:

1. Establish a process for M/S accreditation for use in the OT&E process which maintains independence between the model's development and its evaluation.

2. Provide centralized guidance and oversight to model developers and users for each accepted model <u>in a particular application</u>.

3. Provide for the maintenance and configuration control of accredited models and simulations throughout the weapon system's life cycle.

4. Develop a clearinghouse of reference models, including model versions, points of contact, etc.

5. Provide for dissemination of model software, data and documentation, when requested.

6. For categories of related models, establish user/developer groups, to exchange information on model requirements, voids, capabilities, limitations, and experiences.

7. Establish standard scenarios when appropriate.

8. Establish a process to ensure that threat representation and usage modeled or simulated are consistent with national and departmental intelligence estimates (DIA or Service intelligence validation of the threat data, and how the threat data are used, is required).

9. Documentation supporting the development and use of M/S to support OT&E must address the following points:

   a. the need or rationale for the use of M/S, and how M/S results will address the mission requirements, critical operational issues, and test criteria by showing a clear audit trail (traceability);

   b. how field testing and M/S will overlap and complement each other;

   c. how M/S will support field test design and execution;

   d. what simulation resources are available for the particular application and which capabilities will have to be developed at what cost.

APPENDIX A

EXAMPLES OF APPLICATIONS OF M/S IN OT&E

1. To support pre-test planning.

2. To assist in the identification of critical issues to be addressed in a test.

3. To identify important test parameters earlier.

4. To grossly bound the problem and proposed solutions based on the intended environment, force structure, threat, tactics, strategy, and doctrine.

5. To identify oversights and flawed logic.

6. To determine sensitivity of a program to various input parameters.

7. To conduct non-destructive evaluations of high cost items which would, by their nature, be destroyed in actual hardware tests.

8. To provide better understanding when full-scale testing is not possible. To augment, extend, and enhance test results, in general.

9. To provide multiple "environments" for examining test questions.

10. To provide advantages of time compression, controlled expenditures, replicability, and reduction of variables under study.

11. To assess impact of known parameters of unavailable threat systems.

12. To accomplish human factors evaluations in part-task or limited fidelity "mock-ups."

13. To provide estimates of potential test outcomes.

14. To extrapolate, with caution, test results into other scenarios and levels of force aggregation if M/S and assumptions are applicable.

15. To address issues which cannot be physically tested.

16. To address "what if" questions during post-test analyses.

17. To develop and refine test scenarios and data matrices to obtain maximum data from limited test resources.

18. To develop new tactics for the employment of new weapon systems under test.

19. To provide overall system, scenario, or environment representation.

20. To represent the input, process, and output of non-available systems, subsystems, or components (friendly or threat).

21. To represent the whole integrated system when all components are not available.

22. To allow an assessment of test events that would otherwise be exposed to threat intelligence exploitation.

23. To act as a system driver or stimulator in order to stress a system beyond available field test scenarios.

24. To determine adequacy, effectiveness, and suitability of the planned operational and maintenance concepts.

25. To estimate mature system mission reliability, availability and logistics support frequency.

APPENDIX B

CREDIBILITY ISSUES TO ADDRESS

Credibility, as applied to modeling and simulation processes and results, is a combined impression of the inputs, processes, outputs, conclusions, the persons or agencies involved, and the strength of the evidence presented. This appendix contains questions that the developing, review, and user communities should ask and/or be prepared to answer when models and simulations are used to support operational test and evaluation.

Has the M/S gone through an approved process to establish its credibility?

Why was this model used in lieu of testing?

Was M/S discussed in the TEMP?

Were M/S results compared with combat, field test, and other models? If so, what were the results?

Did the simulation accurately reflect the system requirement and any available developmental test data?

What is the linkage between DT&E and OT&E and theater modeling?

What have the results been validated against? What is the availability and source of data?

What is the statistical confidence in the results?

How robust are the results on operational capability and supportability?

Who built the model? Certified the inputs? Certified the tactics/scenarios?

Who did the verification and validation? What implicit and explicit assumptions were made?

What sensitivity analyses have been performed?

Why was this particular model chosen? What was it designed to do? What are its strengths/weaknesses? Where has it been used before?

Do we always win in this model application? If so, why?

How far has the model been pushed to extremes and how has it performed? Has the M/S domain been established?

What field test results have been fed back into the model for validation?

Is there a documented audit trail? Will it provide traceability of critical decisions?

Is there adequate funding to support the M/S? By whom? Is the M/S cost-effective?

What elements of M/S should be confirmed by operational testing?

Were excursions made? If so, why and what were they?

What impact (if any) did excursions have on the evaluation?

What is the degree of independence of modelers with respect to the program office? If the M/S developer is associated with the program office, was an independent assessment of the M/S applicability made?

Has this model been used by the developer of the weapon system? What were the results?

Who is maintaining the model?

What is the source of threat data? Is it consistent with data used in other analyses? What is the source of threat (Red) tactics used in the scenario?

What variables of the operational environment are not represented?

Who is expected to use or operate the model?

Can one design and build the model faster or cheaper than the system it represents?

If multiple models are used, what are the linkages? Data structures?