Current SSD Programs

NOTE: The items shown in blue below are linked to the respective pages. Click on the title.

CONTENTS

Page

Section A. Strategic Weapon Systems

Section B. Tactical Systems

Section D. Ballistic Missile Defense

Section E. Civilian Programs

SECTION A. STRATEGIC WEAPON SYSTEMS

A1. Fleet Ballistic Missile Test and Evaluation

John P. Gibson

INTRODUCTION

The Navy's Fleet Ballistic Missile (FBM) Strategic Weapon System (SWS) is recognized today as the principal component of the U.S. nuclear strategic deterrent. The submarine-launched ballistic missile (SLBM) on its nuclear-powered submarine platform provides a mobile, long-patrol duration, covert, and invulnerable strategic deterrent force (Fig. A1-1). The FBM system has evolved through several generations of missiles and weapons systems deployed on five classes of submarines. The original fleet of 41 SSBNs has

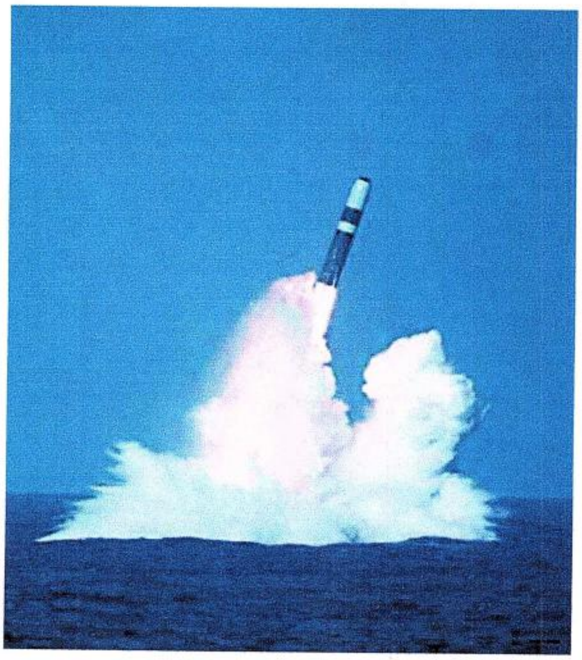


Figure A1-1. A Trident II (D5) SLBM broaches and ignites during a Demonstration and Shakedown Operation off the coast of Florida.

been replaced with 18 newer Ohio Class Trident SSBNs. Table A1-1 provides the specifications of the six generations of FBMs beginning in 1960. Each generation has been succeeded quickly by a more advanced version, resulting in a mixture of deployed systems for the Navy to sustain and manage. The increasingly complex features of each new FBM/SWS have provided unique challenges in the design and implementation of the test and evaluation program required to validate its capabilities and produce the high level of credibility essential to its national deterrent mission.

APL has assisted the Navy Strategic Systems Programs (SSP), since the inception of the FBM Program, in defining and conducting a continuing test and evaluation effort for each generation of SWS. APL has also assisted the U.K. Royal Navy with evaluations of their FBM fleet, starting with the U.K. Polaris program in the mid-1960s, and continuing through the current deployment of the U.K. Trident SSBN fleet. The results of the U.S. SWS evaluations are provided to the Navy technical and Fleet Commands, which then present them to the U.S. Strategic Command (USSTRATCOM) for strategic targeting requirements. This article describes the current, ongoing efforts of the Strategic Systems Department (SSD) in support of this national priority program.

not
jat
[illegible]

TEST PROGRAMS

Because of the vital importance of the FBM Program to the national nuclear deterrent force, and the requirement for annual performance estimates, an ongoing test and evaluation approach has been established to monitor these systems throughout their deployed life. APL has assisted SSP in structuring a comprehensive test program and is the principal agent for the continuing evaluation of the FBM weapon system for the Navy. The three primary test programs of the FBM SWS evaluation are (1) Demonstration and Shakedown Operations (DASOs)—testing that is conducted before strategic deployment, (2) patrol—recurring tests conducted during each strategic deterrent patrol, and (3) Commander-in-Chief (CINC) Evaluation Tests (CETs) or Follow-on CETs (FCETs)—

Table A1-1. Specifications of six generations of SLBMs.

	SLBM					
Feature	Polaris A1	Polaris A2	Polaris A3	Poseidon C3	Trident C4	Trident D5
Year deployed	1960	1962	1964	1971	1979	1990
Length (ft)	28.5	31.0	32.3	34.0	34.0	44.6
Diameter (in.)	54	54	54	74	74	83
Weight (lb)	28,000	32,500	35,700	64,000	73,000	130,000
Range (nmi)	1200	1500	2500	2500	4000	4000
Payload	1 RB	1 RB	3 RBs	MIRVs	MIRVs	MIRVs
Guidance	Inertial	Inertial	Inertial	Inertial+ stellar	Inertial+ stellar	Inertial+ stellar
Propulsion stages	2	2	2	2 + bus	3 + bus	3 + bus

Note: RB = reentry body; MIRVs = multiple independent reentry vehicles.

end-to-end weapon system tests, including missile flights, conducted with randomly selected SSBNs periodically throughout the life of the system.

The effective implementation of a comprehensive test and evaluation program is highly dependent on involvement during the earliest design and development phases of the system. To ensure availability of the required test data for the deployed system, identification and integration of necessary instrumentation, testing concepts, and special test procedures must be accomplished during the weapon system design.

SSD has played an important role in defining evaluation and data requirements for each generation of the FBM SWS. Novel system test concepts have been devised, and sensors and instrumentation have been conceived, built, and utilized in this continuing evaluation effort. Recent APL-developed innovations include the introduction of electronic log-keeping devices and a versatile, onboard ship-control training capability (see the article by Biegel et al., this issue). Recurring test and evaluation tasks include the design of individual flight test mission trajectories consistent with current FBM employment concepts, production and maintenance of test procedures unique to each test program and SSBN class, and specialized training sessions for SSBN crews.

The contributions of SSD to the continuing FBM SWS test and evaluation effort are summarized in Fig. A1-2 and discussed in other articles in this issue. To perform these tasks, SSD has maintained a permanent field office at Cape Canaveral, Florida, since 1959, provides an on-site representative to the staff of the Commanders of the Atlantic and Pacific Submarine Forces for required liaison with the operational forces, and maintains a dedicated data processing and analysis facility. The following sections describe the basic components of the FBM SWS test and evaluation program.

Demonstration and Shakedown Operations

DASO exercises are conducted by each U.S. and U.K. SSBN prior to strategic deployment after either new construction or a shipyard overhaul period. This is the first time that the new or upgraded weapon system undergoes full, comprehensive system tests and culminates in a test of the entire system, including missile launch. Interspersed throughout the dockside and at-sea operations are a series of activities and simulated countdowns (many with inserted casualties) to provide realistic training. Objectives of this program are to (1) certify the readiness of the SSBN weapon

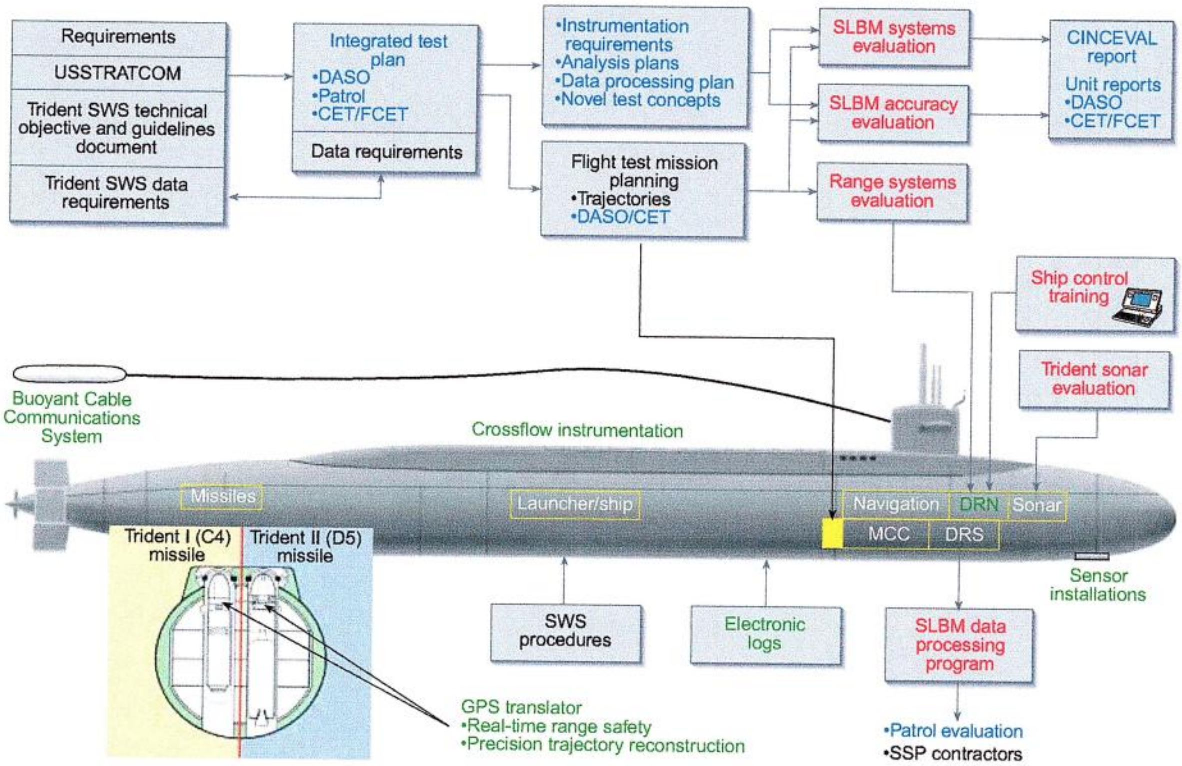


Figure A1-2. SSD contributes to the continuing Navy Fleet Ballistic Missile Strategic Weapon System test and evaluation programs. (SSD-developed hardware, green; documents/reports, blue; programs, red. MCC = master control console, DRS = Data Recording System, DRN = DASO Reference Navigator, CINCEVAL = Commander-in-Chief Evaluation.)

system and its crew for strategic deployment, (2) evaluate the technical performance of the weapon system while in an operational environment, identifying material and procedural deficiencies, and (3) provide data used to derive current reliability and accuracy measures of performance for the deploying weapon system.

APL participates in DASOs by providing a team of up to 14 professional staff members, supported by additional administrative and data processing personnel, at the DASO field site in Cape Canaveral. APL also provides field technical support to the SSP Weapons Evaluation Branch Team tasked to conduct and evaluate each DASO. The DASO simulates all phases of a typical strategic patrol to validate procedures and evolutions that will be conducted during deployment. In addition, SSP-approved special tests are scheduled and performed, as appropriate, to evaluate new capabilities and new equipment, or as diagnostic investigations.

SSD develops, conducts, and evaluates some of these special tests, such as the submarine crossflow evaluation depicted in Fig. A1-3. The crossflow instrumentation measures, records, and provides real-time display of the speed of water across the submerged SSBN missile deck during FBM launch operations or special at-sea tests. This instrumentation allows SSD to evaluate the dynamics of FBM underwater flight throughout the SSBN speed/depth launch envelope. During DASOs, SSD field-test teams evaluate all aspects of the performance for each weapon subsystem (navigation, fire-control, missile, launcher, and ship) and provide the Navy with a technical report documenting the results of that evaluation. This report is generated within 2 weeks of completing each DASO and is used to support the Navy's certification for deployment.

Patrol

Each SSBN and weapon system may be deployed for a decade or more. The continuous monitoring of each SSBN identifies hull-unique problems as well as Fleet trends that may evolve or change with time and affords a current, cumulative weapon system performance estimate, which is critical to maintaining the credibility of this strategic deterrent system. The objectives of patrol evaluations are to (1) provide FBM weapon system performance information in the actual patrol environment for use in deriving USSTRATCOM performance planning factors, (2) provide Navy Fleet

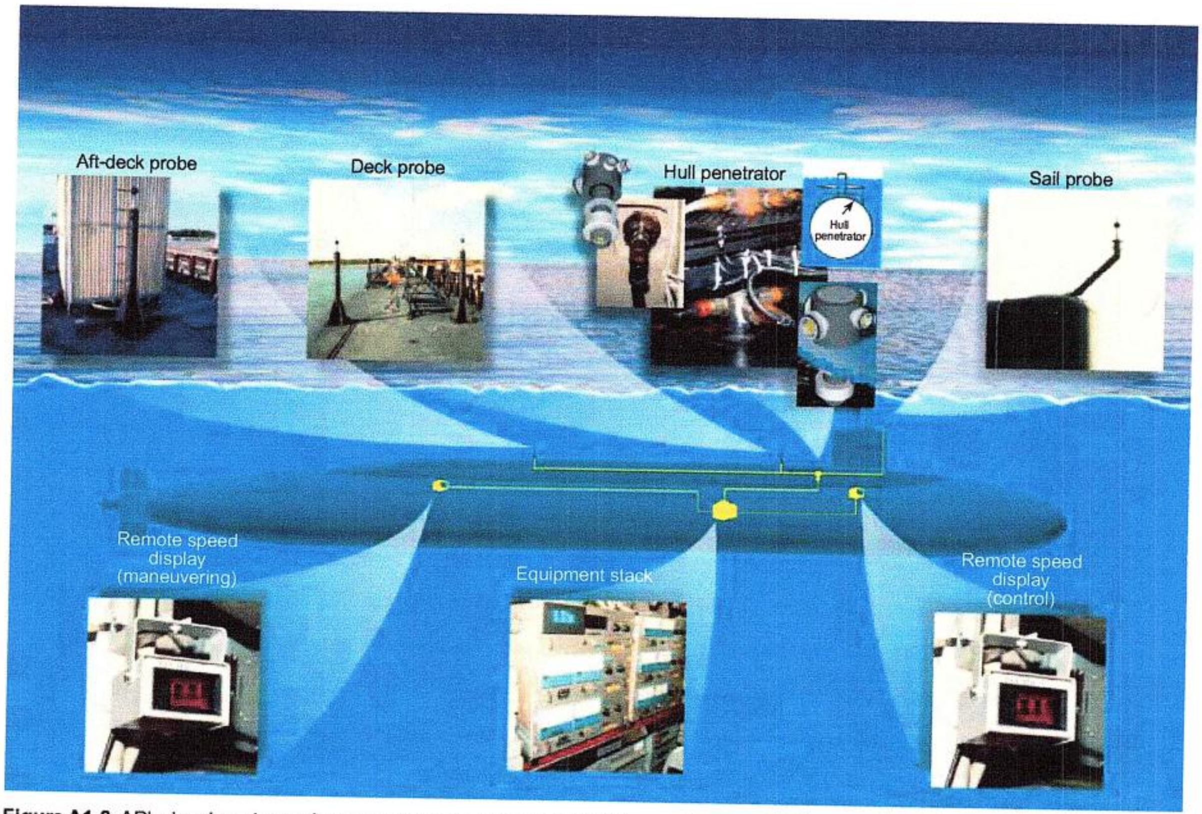


Figure A1-3. APL-developed crossflow instrumentation (see Fig. A1-2) is used during a DASO to measure, record, and display real-time water speed across the SSBN missile deck during an SLBM launch.

Operational Commanders and SSP with an independent system evaluation of each SSBN and its weapon system while on strategic deterrent patrol, and (3) provide individual SSBN crews with an analysis of the performance of their weapon system during patrol.

While on patrol, the submarine regularly conducts tests that activate the weapon system in a manner similar to an actual countdown to launch (but with appropriate safeguards in place). In addition, tests that monitor the health of each subsystem along with routine maintenance are conducted regularly. Data from all of these sources, both electronically recorded and manually logged, are sent to APL after each patrol. Engineers and analysts review raw and processed data to identify equipment problems, faults, and other abnormal conditions and to initialize simulations used in the patrol evaluations.

SSD has developed a set of electronic data-logging devices for SSBN crews. These electronic logs are replacing the traditional paper ones and will increase efficiency in documenting and evaluating patrol activities. Electronic data from the individual navigation, fire-control, launcher, and control and monitoring panel logs are transferred to the Electronic Weapons Log (EWL) base station at the end of each watchstander's shift. The EWL base station maintains the historical log record onboard the SSBN and allows the crew to use these files for analysis. EWL data are copied to compact disc and transferred to APL at the end of each patrol or upkeep cycle. After the patrol data package is reviewed, patrol and upkeep quicklook reports are produced to provide an overview of in-port and underway activities and problems that may need attention before the next patrol. A more detailed patrol summary report gives a synopsis of each subsystem's performance throughout the period and provides the data from which hull-unique or class performance trends are examined.

Certain patrols are evaluated randomly in greater depth, consistent with the need to obtain information to form credible annual estimates of weapon system performance. APL engineers meet with the crew to review the evaluation and confirm interpretation and understanding of the logged activities. After this review, a final patrol report is published and distributed to SSP and its contractors, the Operational Commanders, and the submarine crew. The data and evaluations from these patrols contribute to the APL annual weapon system performance estimates that are used by USSTRATCOM to prepare strategic targeting.

CET/FCET

The CET/FCET is the continuing operational test program conducted annually with randomly selected SSBNs. The results of these tests, which include the launch of multiple test-configured FBMs, provide the basis for the annual performance estimates of the FBM SWS. The objectives of this program are to (1) determine operationally representative weapon system performance characteristics for targeting purposes, (2) ensure that planning factors do not significantly change with time, (3) determine the adequacy of tactical procedures, and (4) provide diagnostic information that may lead to system improvements.

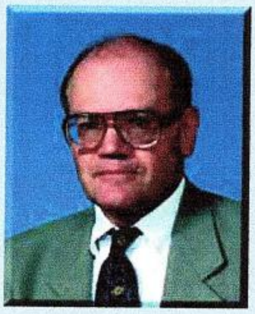
Without advance notice, a selected SSBN is recalled from patrol for a CET. Two or more missiles, selected randomly from the onboard complement of tactical missiles, are converted to a test configuration alongside the wharf at the normal refit site. Following this evolution, the submarine proceeds to a launch area and resumes operations as if on a strategic deterrent patrol. The USSTRATCOM transmits an exercise launch message at random via the strategic communications links. When the message is received, the submarine, using tactical procedures, launches the designated CET missiles at the tactical firing rate. Data obtained from instrumentation onboard the SSBN, from a launch area support ship, and from downrange support sites (on ship, aircraft, and land) provide the information necessary for SSD engineers to assess total weapon system performance in this near-tactical end-to-end test.

An APL representative meets the submarine when it returns from the exercise to review the operation with the crew, inspect the condition of the weapon system, and provide a quicklook report to the Navy on the overall operation. A team of APL engineers conducts a subsequent detailed evaluation of all the data from the exercise, and a report covering the entire test operation, including missile performance during flight and reentry, is provided to the Navy.

The evaluations that APL performs for these three major test programs (DASO, patrol, and CET/FCET) provide the complete data set necessary to derive current weapon system planning estimates that are prepared annually for the CINC evaluation reports. These annual reports are prepared for the cognizant CINC of each SWS in accordance with evaluation guidance specified by USSTRATCOM. For the FBM Program, APL has prepared them for the CINCs of the Atlantic and Pacific Fleets since 1966. The CINC evaluation reports are forwarded by those commands to USSTRATCOM for use in the annual strategic targeting laydown. They provide estimates of weapon system prelaunch, in-flight, and reentry reliability, accuracy, reaction time, launch interval, and missile performance capabilities. A detailed discussion of the validity of the test program and the sources of demonstrated performance data used in developing the estimates is also provided, as required by USSTRATCOM.

These APL-generated reports provide an annual assessment of the complete FBM SWS for the nation's strategic planning processes. Future generations of the FBM SWS will most likely be developed in an entirely different manner than earlier generations, i.e., their development will be more evolutionary and involve less whole-system replacement. Furthermore, the complexity and cost of these advanced strategic weapons will undoubtedly limit the number of full-scale tests that can be conducted. The ability to test and demonstrate system capability to potential adversaries will nevertheless remain a crucial element in the credibility of these strategic deterrent systems. Therefore, the Laboratory must continue to develop and implement improvements to its test and evaluation approaches. The challenge will be to execute this continuing, nationally important task efficiently and cost-effectively while maintaining the credibility that has been the hallmark of APL's contributions to the FBM Program.

THE AUTHOR



JOHN P. GIBSON is a member of APL's Principal Professional Staff. He received a B.S. degree in aerospace engineering from the Pennsylvania State University in 1964, and an M.S. degree in aerospace engineering from Drexel University in 1967. Since joining APL in 1967, he has worked in the Strategic Systems Department. Mr. Gibson has served as Program Manager for the Trident I (C^4) Strategic Weapon System Evaluation and as Program Area Manager for the Strategic and Tactical Systems Programs. His e-mail address is john.gibson@jhuapl.edu.

A2. The Fleet Ballistic Missile Accuracy Evaluation Program

Dean R. Coleman and Lee S. Simkins

BACKGROUND

In the early 1970s the Navy was asked to respond to a DoD request to produce a development plan for a future highly accurate FBM Strategic Weapon System (SWS). The Trident I SWS, then in development, as well as its predecessors (Polaris and Poseidon), were designed to meet accuracy goals that were well within the existing state-of-the-art. The observed system accuracy for each generation of FBM met those goals but was not thoroughly explainable. As a result, insight was lacking into the technical limitations on the incremental improvement in accuracy that might ultimately be achieved in an advanced system (Trident II). In order to plan a set of design options with confident, quantifiable accuracy improvements, the Navy needed an improved technology base. In 1975, Strategic Systems Programs (SSP) initiated an Improved Accuracy Program to gain the understanding and tools necessary to validate the accuracy of the design options as well as the instrumentation needed to evaluate a new high-accuracy system.

SSD played a leading role in the Improved Accuracy Program over its 8-year course, fulfilling a system evaluation task for the Navy in helping to achieve the accuracy technology base for Trident II. Advanced instrumentation, data processing, and error estimation techniques were developed by SSD together with other members of the Navy/contractor team and were used to gain insight into the sources of inaccuracy during flight tests of the Trident I weapon system, which provided the springboard for Trident II development concepts.

An SSD system-level accuracy model validation effort, in conjunction with subsystem-level investigations by hardware contractors, led to high-fidelity analytical accuracy models that were used in Trident II trade-off studies. SSD long emphasized to the Navy the importance of accuracy instrumentation, in particular, to enable errors to be sufficiently visible so that test results could be extrapolated to untested, tactical conditions.

APL was asked to determine the instrumentation and evaluation concepts that would be needed for Trident II to ensure a high-confidence accuracy evaluation capability. Through a joint effort between APL's Space Department and SSD, the Accuracy Evaluation System for Trident II was defined by early 1982. A satellite-based instrumentation system known as SATRACK had been conceived by APL in the early 1970s and proven in Trident I applications. It would become the backbone of SSD's evaluation capability for the advanced Fleet Ballistic Missile Strategic Weapon System.

TRIDENT II AND ADVANCED SYSTEMS

The stringent Trident II accuracy performance objectives motivated the development of demanding performance evaluation criteria and objectives. The Navy's desire to understand the system's performance with high confidence was translated into several specific accuracy evaluation objectives. These had significant implications with respect to analysis methodology, instrumentation, and modeling and simulation.

The Accuracy Evaluation System study outlined the process for attacking the accuracy evaluation problem. First, the evaluation objectives required that system performance be estimated. It would no longer be sufficient to use model validation approaches wherein test data were used to validate or invalidate contractor-supplied performance models. Without a methodology that provided direct estimates of parameter values, knowing that a model was to some degree invalid begged the question: If the current model is invalid, then what is the better model? Thus, model parameter estimation was established as the fundamental approach, and the method of "maximum likelihood" was adopted as the preferred methodology for identifying accuracy parameters from test data.

The requirement to estimate performance did not end there, however. Quantified confidence was also necessary. There had to be a procedure by which the uncertainty with which we observe performance as well as the finitude of test programs was translated into specified confidence (or uncertainty) in the accuracy parameters being estimated. Information theory provided the basis for developing algorithms that could quantify the confidence with which accuracy would be estimated. Next, performance was required to be known, and not just at the system level. The accuracy evaluation system had to be able to isolate faults and estimate performance of the subsystems or the various phases of the weapon system. This required that instrumentation and measurements be made not only at termination (e.g., reentry vehicle impact or airburst) but also during tactical patrol and at every phase of a full system test (prelaunch, powered flight, reentry body deployment, free-fall, and reentry). Figure A2-1 depicts the current Trident II flight test instrumentation suite.

Since the number of allowable tests used for the determination of estimates was specified at fairly low-to-modest levels (about 10 to 20 tests), the instrumentation had to be of sufficient quality to provide the high-confidence estimate; thus, a high-level goal was established to maximize information from the expensive and limited flight test samples. In addition, the evaluation objectives required that we be able to extrapolate to untested conditions, that is, to predict tactical performance, with high quantified confidence, from test data.

The need to predict tactical accuracy from test data had a profound impact on how the modeling was performed. Accuracy contributions had to be modeled at a fundamental level, independent of the test environment. For instance, inertial guidance errors would be characterized and modeled in detail at the hardware component level, i.e., complete mathematical descriptions (including cross-coupling and higher-order terms) of the input/output characteristics of the individual gyro and accelerometer hardware, component misalignments, etc. The structure of these detailed error models was derived from physics, first principles, or contractor component and subsystem tests. However, the values of the parameters would be derived from demonstrated operational test data.

In some cases, it was impractical or unnecessary to require modeling at such a level or to restrict data sources to flight tests. Additional sources of "demonstrated" data to supplement the flight testing were devised. For example, a novel approach for gathering representative navigation data, called the Navigation Accuracy Test, was developed by SSD to be conducted periodically during strategic deterrent patrols. Procedures and instrumentation were developed so that navigation contributions to system inaccuracy could be ascertained from simulated system countdowns during tactical alert periods; a missile did not need to be launched in order to understand a ship's navigation performance. This approach would provide significant insight and more data samples than would have been available if the evaluation were limited solely to the missile flight test program.

Data from each accuracy test were analyzed using some variant of a Kalman filter. Within the filters are the detailed models of both the system and instrumentation for each subsystem. Figure A2-2 depicts notationally how this analysis is accomplished. Given a particular test or scenario (say, a flight test) measurement data are collected on the various subsystems. Using rigorous methods, these data are combined with prior information generally developed and maintained by contractors responsible for various parts of the system under test. This prior information is necessary for single test processing, given the incomplete observability of error sources.

The outputs of the filter provide a basis for understanding particular realizations of system and subsystem

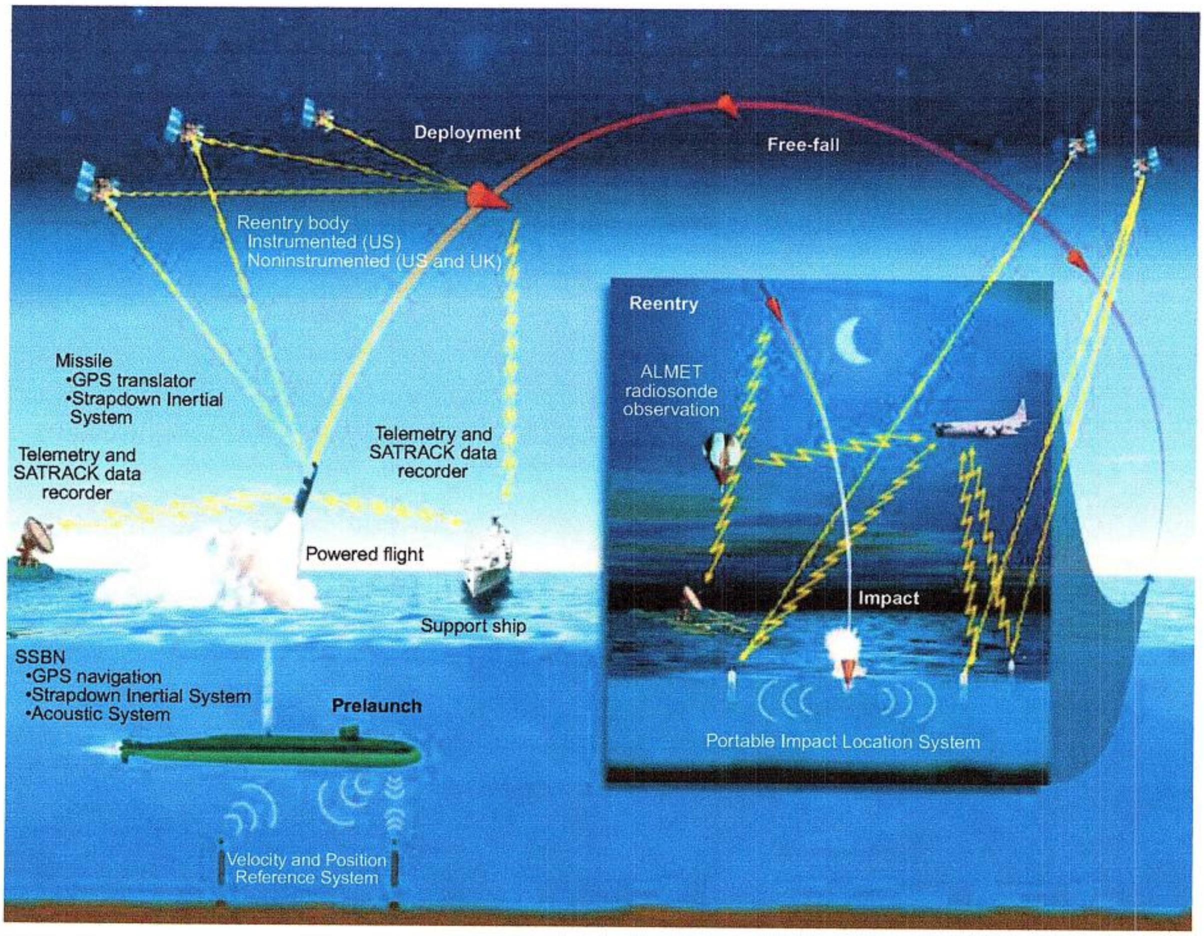


Figure A2-1. Current Trident II (D5) accuracy instrumentation suite. Measurements are made at every phase of a full-system test. (GPS = Global Positioning System, ALMET = Air-Launchable Meteorological System.)

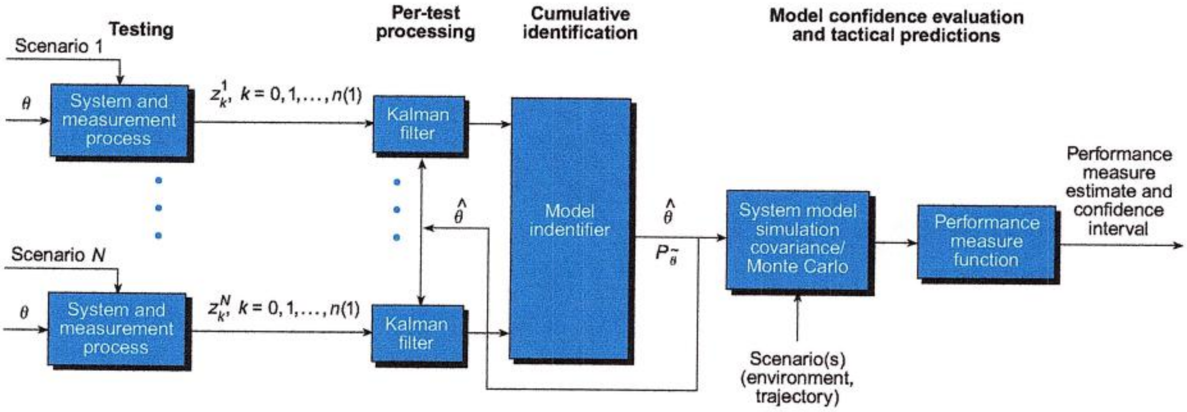


Figure A2-2. Strategic Weapon System accuracy evaluation concept. θ = model parameter, $\hat{\theta}$ = estimate of parameters derived from tests, $P_{\tilde{\theta}}$ = estimation error covariance matrix, z_k^j = measurement k from test j.

behavior. Analysis results provide insight into the sources and causes of inaccuracy (Fig. A2-3). The results of multiple tests (the outputs of the Kalman filters) serve as input to the cumulative parameter estimation process; however, all prior information relative to the error models is removed so that the estimated accuracy is derived solely from the test data.

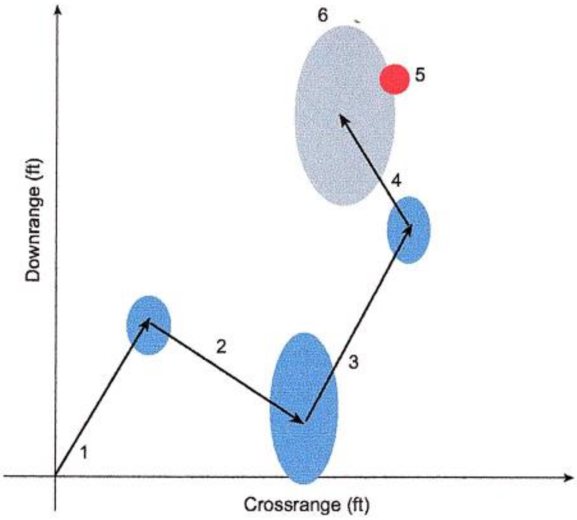


Figure A2-3. Reconstruction of sources of missile impact miss distance error. 1 = initial conditions, 2 = guidance, 3 = stellar residual, 4 = deployment, 5 = reentry body measured impact, 6 = total uncertainty.

This process solves the highly nonlinear equations for the means, variances, and Markov parameters that characterize the overall system accuracy performance. In addition, uncertainties in the parameter estimates are calculated so that we have a quantitative measure of our confidence in the solution. The ultimate desired product is a performance prediction for the system under tactical, not test, conditions. Here we rely on models of the tactical gravity and weather environment developed from data and instrumentation. These models, along with deterministic simulations of the system, are then used to "propagate" the fundamental model parameter estimates and uncertainties to the domain of interest—system accuracy at the target.

TECHNOLOGY ADVANCEMENTS

The development, maintenance, and evolution of the Trident II Accuracy Evaluation System provided considerable technical challenges in terms of methodology, numerical methods, mathematical modeling, algorithms, software, and instrumentation. Noteworthy developments include constrained numerical optimization algorithms; efficient gradient approximation techniques; large-scale, efficient, and numerically stable filtering algorithms; high-fidelity models and simulations of inertial guidance, navigation systems, and gravity; the use of the Global Positioning System (GPS) for precision tracking; development of GPS translator concepts and hardware; advancements in GPS signal tracking and receiver technology; modeling and development of precision acoustic reference systems; and target pattern optimization. Many of these developments have been extended to other weapon systems and programs, including the Air Force's Peacekeeper ICBM, the Army/Ballistic Missile Defense Organization (BMDO) Exo-atmospheric Reentry Interceptor System, and the ongoing test and evaluation of the BMDO exo-atmospheric kill vehicle.

PRINCIPAL ACHIEVEMENTS

The Trident II Accuracy Evaluation Program has contributed to the success of the SWS in several important ways.

Instrumentation Requirements and Test Planning

While in the early development phase, models and simulations of accuracy evaluation processes supported rigorous quantitative trade-off studies designed to support management decisions about instrumentation and test program requirements.

Accuracy Understanding

Analysis has provided unprecedented understanding of and confidence in system performance. The analytical accuracy model has been refined to where current performance is faithfully predicted and is known to be a fraction of the original objective. Biases have been isolated and estimated. System use is enhanced by virtue of our understanding system performance as a function of the tactical operational and environmental parameters. Anomalous test performance has been more easily detected, and causative factors have been isolated.

System Improvements

Improved models and understanding of accuracy provide improved system performance by way of embedded system software. The calculation of system gains used when processing guidance stellar sightings or reentry body fuze information relies on an accurate characterization of system performance. The calibration of reentry body release parameters has been improved by knowledge gained from onboard inertial instrumentation. Accuracy enhancement potential through modified operational scenarios has been demonstrated to be viable.

Accommodation of Testing Cutbacks

Proper instrumentation and a rigorous analytical approach required less testing to achieve the desired initial confidence. In addition, follow-on testing of the deployed system was reduced without significant risk as a result of near-optimal use of the limited flight test assets.

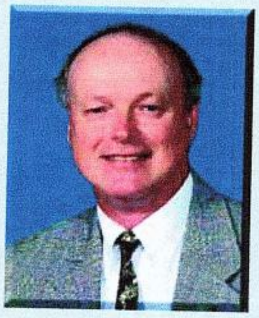
FUTURE DIRECTIONS

In the last several years, there has been considerable interest in a GPS/Inertial Navigation System as both instrumentation and as a candidate tactical missile or reentry body guidance system. Several special tests of missiles and reentry bodies have been conducted with various combinations of inertial systems (space-stable and strapdown), GPS receivers, and GPS translators, as well as various RF/antenna designs. Technologies have been developed to enhance and extend signal-tracking capabilities further, including during periods around onset of plasma blackout and recovery following blackout. Interest in achieving even greater accuracy has been facilitated by the detailed understanding of Trident II performance. Special tests have demonstrated that accuracy can be achieved to support potential new and extremely demanding tactical strike scenarios. Sophisticated tools for exploring optimal target patterning have been developed to support these studies.

Future FBM systems may look very different from present systems. Current modeling and simulation efforts are drawing upon Trident II experience to predict and trade off system design options. Techniques for properly merging ground test (e.g., centrifuge test) data with flight data are being developed in response to the changing test and evaluation environment, where there is much emphasis on affordability and cost reduction. Technology and hardware that support precision intercept system evaluation have been demonstrated, extracting from Trident II technology and extending it through independent research and development projects.

The success of the Trident II system and the Accuracy Evaluation Program is due, in large measure, to SSP leadership. SSP's desire to mitigate risks in the development and maintenance of a high-accuracy strategic deterrent created a vision for an evaluation approach developed as an integrated part of the system. Instrumentation, analytic methods, and modeling and simulation were exploited to optimize the procurement and use of limited and expensive flight test assets. The program has been, and continues to be, successful in meeting its objectives.

THE AUTHORS



DEAN R. COLEMAN is manager of the SLBM Accuracy Evaluation Program within APL's Strategic Systems Department. He received his B.S. in mathematics from the University of Maryland in 1970 and M.S. degrees from The Johns Hopkins University in 1972 and 1985 in numerical science and technical management, respectively. Mr. Coleman has over 30 years' experience in test and evaluation of the Navy's Fleet Ballistic Missile weapon systems. His work has included complex system modeling and simulation as well as evaluation methodology development. Since 1980, he has focused on system accuracy evaluation, where he coordinated APL's SLBM Improved Accuracy Program Analysis Plan and led the Laboratory's efforts to apply this methodology initially to evaluate Trident I and subsequently to evaluate the high-accuracy Trident II weapon system. His e-mail address is dean.coleman@jhuapl.edu.

LEE S. SIMKINS is a member of APL's Principal Professional Staff. He received a B.S. in aerospace engineering from the University of Michigan in 1972 and an M.S. in mechanical engineering from the University of Maryland in 1977. Since joining the Laboratory in 1972, he has worked in the Strategic Systems Department. Mr. Simkins is the Supervisor of the System Development, Test, and Evaluation Branch. For most of his career at APL, he has contributed to the development, implementation, and use of system identification and estimation methodologies that allow the evaluation of submarine navigation, missile guidance, and other military systems. Those efforts have made extensive use of the Global Positioning System as well as other instrumentation consisting of telemetry, inertial systems, and external reference information. His e-mail address is lee.simkins@jhuapl.edu.

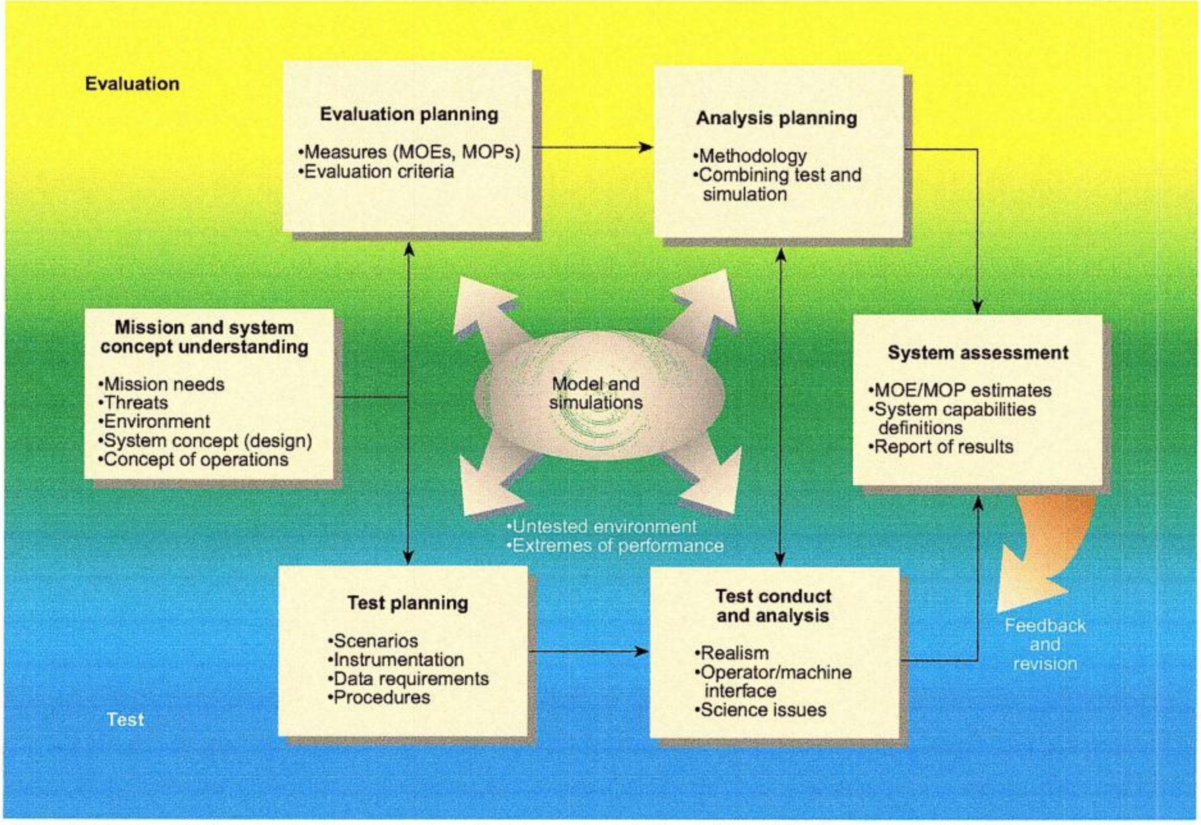


Figure 3. Systematic deployed system assessment is characterized by testing that is both representative of tactical situations and responsive to science issues. Testing is driven by evaluation programs that produce useful measures of performance and effectiveness (MOPs and MOEs, respectively).

predictions of system effectiveness in untested environments and situations.

ESSENTIAL FEATURES OF FULL-SYSTEM EVALUATION IN THE OPERATIONAL ENVIRONMENT

Effective system-level testing in the operational environment involves four key features.

1. An integrated mixture of testing (equipment and subsystem surveillance, representative free play with complete systems, environmental measurements, and component-level engineering data collection) supported by instrumentation, an evaluation process, and a mission-based assessment
2. Validated models and simulations to leverage test results into assessments of system performance in untested or infrequently tested areas
3. Instrumentation that is built into the system design
4. A fundamental understanding of the system, the environment in which it operates, and the interactions among components, subsystems, and the environment

These features strengthen the view that postdeployment testing should support more than pass/fail decisions. We can reach an appropriate level of understanding of system performance that allows us to extrapolate into untested areas or respond to changing plans and emerging threats. The T&E process is a continuum of increasing understanding of the system, in which detailed analysis of system success is as important as fault isolation after system failure.

To illustrate the advantages of this systems engineering approach to the testing and evaluation of deployed systems, we describe the T&E programs of two Navy systems—the Fleet Ballistic Missile/Strategic Weapon System and the passive sonar suite of SSBNs (nuclear-powered fleet ballistic missile submarines).

The Submarine-Launched Ballistic Missile (SLBM) Strategic Weapon System

The Joint Chiefs of Staff requires the Navy to provide credible, quantitative information for strategic target planning. The Navy developed, with APL providing primary support, a T&E program that continues throughout the life cycle of the weapon system.

Instrumentation is integrated in all subsystems. Annual flight test programs are planned, and instrumented missiles are flown from test ranges with special impact areas to evaluate in-flight performance and accuracy. This test program includes significant hallmarks.

- To supplement system-level testing, a comprehensive test program is conducted at the subsystem level to bolster statistical confidence in the performance of critical elements.
- Modeling and simulation is a major contributor to test design, instrumentation specification, and extrapolation of test results to untested environments.
- Instrumentation is designed into the system; all aspects of the weapon system contain built-in instrumentation that is deployed with the strategic assets. This approach provides an early understanding of system performance. Supported by detailed system understanding, the need for follow-on testing is reduced.
- The evaluation requirements drive the test program—test objectives, test design, and test sizing are specified with respect to integrated performance measures.

Instrumentation for the newest SLBM weapon system (Trident II) records on board real-time data, flight telemetry/Global Positioning System data, and impact area sonobuoy detections. Figure 4 shows a schematic of the instrumentation support for a Trident II missile launch. Data from a test flight can be used to identify contributions to miss distance down to the subsystem level.

The Trident II Accuracy Model is the analytical tool that combines subsystem and system data from flight tests to validate fundamental system accuracy parameters and provide estimates of overall system accuracy. During system development, developers and evaluators collaborated under the Navy's Strategic Systems Programs to model Trident accuracy to the component level and to develop a propagation model to combine the elements. This model is now used routinely in system accuracy assessments and is the preeminent tool for identifying the accuracy of system design or employment changes. Our reliance on this model continues to increase as the numbers of missile test ranges as well as trajectory options available to the Navy for direct systems accuracy tests have decreased. Figure 5 shows how the Accuracy Model is used to "systems engineer" the results of flight tests by decomposing missile reentry body impact results from a test flight to component accuracy contributors.

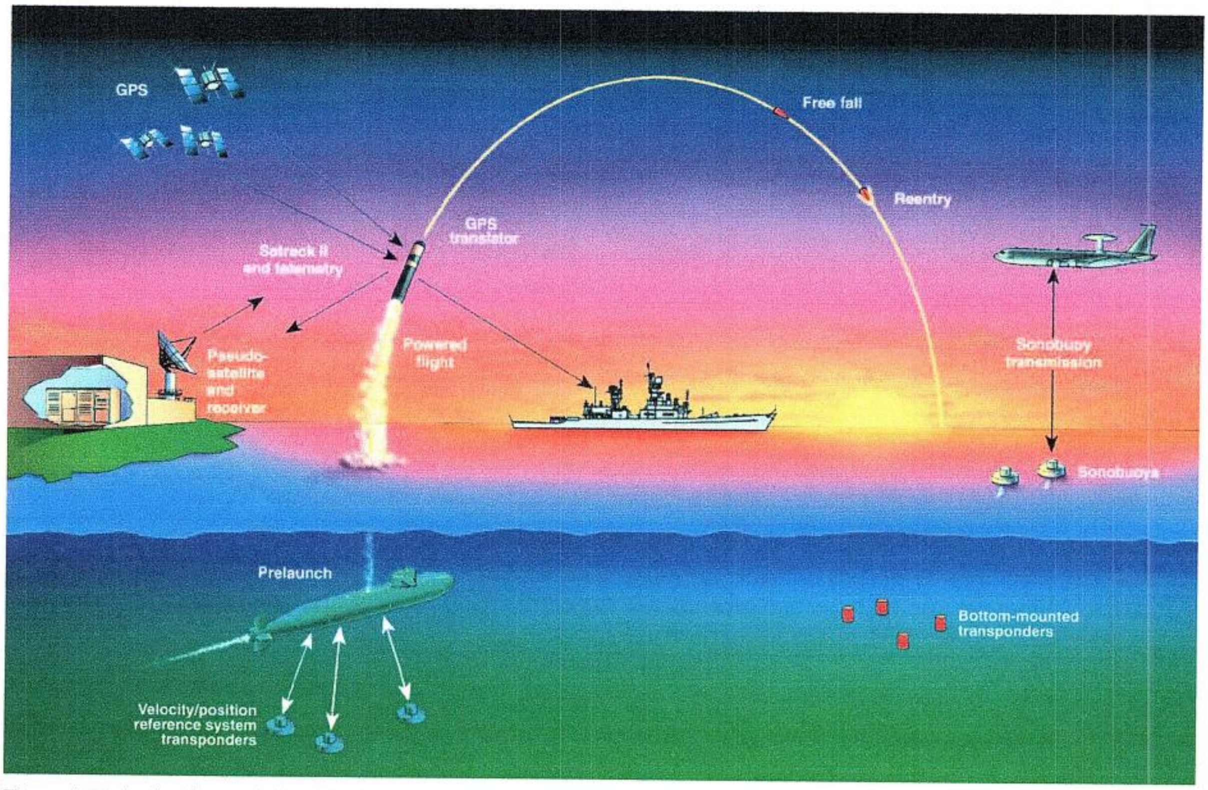


Figure 4. Evaluating the capability of the Strategic Weapon System requires instrumentation covering launch, boost, free fall, and reentry, complemented by sophisticated processing and evaluation methods. Onboard real-time recording includes data on navigation, missile launch-tube environment, missile/guidance initialization, major interfaces, system events, and voice communications. (GPS = Global Positioning System; white lines and arrows denote acoustics, black represent telemetry, and yellow show trajectory.)

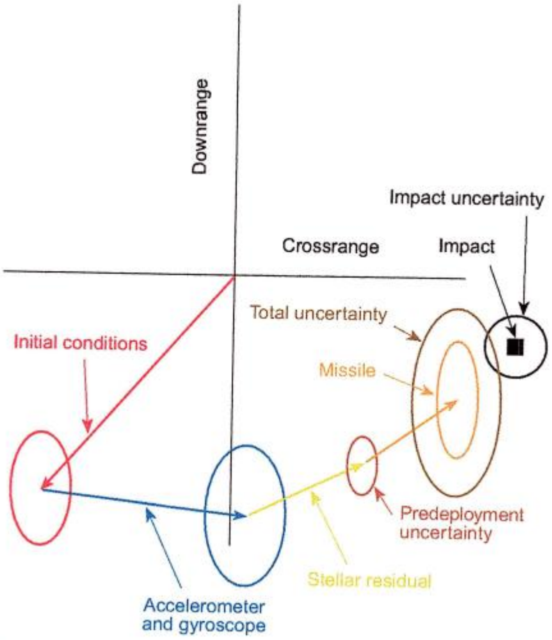


Figure 5. Identification of sources of miss distance, even for successful tests, allows identification of compensating errors that could adversely affect other scenarios (total uncertainty includes modeled reentry aerodynamics uncertainty).

The attention to fundamental knowledge of the phenomenology of system performance is central to a successful T&E program. For the Navy's Strategic Weapon System Program, this knowledge is gained through an integrated test plan that combines the results of an assortment of tests to provide statistical significance to system assessments that complement the solid understanding of the underlying physics of system performance. These assessments range from end-to-end system tests (including missile firing) to a range of system and subsystem tests that collect data on various components of the weapons system and evaluate important interfaces. Most of these tests serve a dual role, accommodating the requirements of the integrated test program while providing training for systems operators.

The data provided for validation of models and simulations constitute an important aspect of the integrated test program. Modeling and simulation has often been heralded as the primary complement to live testing for estimating a system's performance. During the acquisition process, for example, it is cheaper to explore design alternatives by computer simulation than by building and testing prototypes. Without an adequate body of test data to support model validation and verification, however, there can be little credibility in predictions based solely on these evaluation tools. For that reason, modeling and simulation plays the central role in coupling tests of live systems with evaluations of the performance capabilities of those systems.

In the Strategic Weapon System Program, many of the tests intended to support estimations of systems capability are augmented by test phases whose sole purpose is to validate simulations or to provide data for constructing credible models. For example, consider the automatic depth control system whose purpose is to maintain the submarine at a nearly constant depth during a ripple launch. The ability to launch 24 missiles without ship control delay cannot be demonstrated, so a ship control simulation provides the basis for establishing this capability. Tests to support these simulations include "closed-loop" launches in which the depth control system operates automatically to transfer ballast water into and out of the submarine. In addition, some launches are conducted in "open-loop" mode with sea valves closed. These tests provide data to estimate launch forces on the submarine to be used as inputs to depth control simulations. Although the open-loop launches are not representative of actual system application, they provide the best data for estimating the complex forces acting on the submarine as the missile is ejected and seawater floods back into the vacated launch tube (Fig. 6). A physics-based submarine ship control simulation is used to estimate these forces.

The Trident Passive Sonar Suite

The Strategic Weapon System exemplifies a system for which a comprehensive life-cycle and valuation

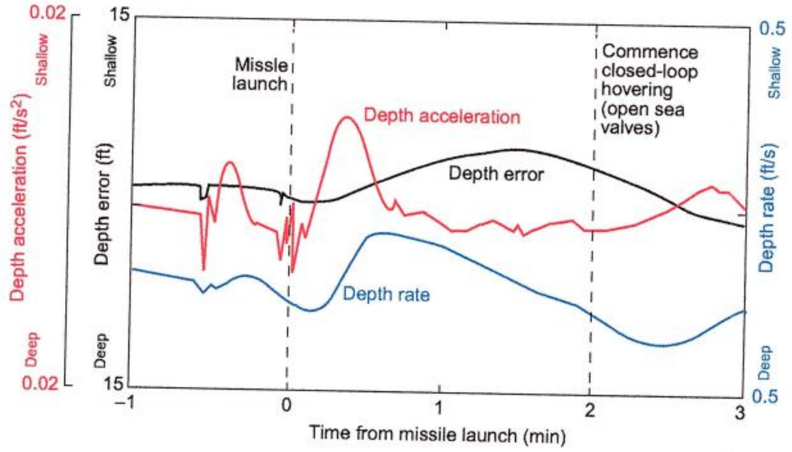


Figure 6. To validate depth control system models, the missile is launched from a submarine with the depth control computer operating but sea valves closed. The computer output, shown here, provides data for deriving launch forces on the submarine and for validating the complete depth-keeping simulation.

program was envisioned during the design phase. The test program for the SSBN passive sonar suite was established after the system had been deployed. Nevertheless, it is an ambitious program involving many test types and a sophisticated instrumentation suite supported by elaborate postprocessing capability and evaluation methods.

The Navy and APL have fitted selected submarine sonar systems with instrumentation to record raw and processed hydrophone data for postpatrol processing and analysis. The recording system interfaces with the tactical sonar suite and records up to 1300 sonar acoustic inputs (6.25 MB/s) continuously during an entire deployment. These data are then processed at the Laboratory on state-of-the-art sonar processors (Fig. 7), which provide flexible beamforming as well as user-selectable displays.

We use a mix of tests and integrate the results into a complete assessment of sonar capability. These tests include an evaluation of sonar performance against selected contacts of opportunity chosen from those available during routine patrol operations. We also use data from security exercises, during which submarines are ordered to operate together in an ASW scenario in which one plays the role of the pursuer and the other the role of the evader. The test program is rounded out with sonar demonstration and shakedown operation exercises in which the exercise geometries may be controlled to provide data for special investigations. Some parts of the controlled exercises are usually reserved for collecting data for model validation (e.g., to establish a baseline for modeling acoustic propagation).

The test program for deployed Trident sonars typifies the way in which developing a thorough understanding of system capability supports planning and decision making when faced with a changing threat. Twenty years ago, when the program was conceived, the threat to U.S. strategic assets consisted of Soviet SSNs (nuclear-powered attack submarines) that were generally noisy and inefficient compared with our submarines. Since then, the successors to those Soviet SSNs are the submarines of the modern Russian Navy. Despite the collapse of the Soviet Union, these new submarines are highly capable, with quieting levels that rival some U.S. submarines. Our understanding of the capability of U.S. strategic assets to tackle a threat as capable as these Russian submarines comes from detailed system knowledge, stemming from a test program that concentrates on quantifying performance factors throughout the system's life cycle rather than establishing whether it has passed preestablished test criteria.

CONCLUSION

The role of T&E during the acquisition cycle is well established. We have identified a class of defense systems that require a credible understanding of performance and a continuing test program to support planning, system modifications, and responses to changing threats. Because we are interested in more than test thresholds, the evaluation of test data concentrates on understanding system performance at the subsystem or even component level. Accurate assessments of the system's interaction with its environment are also necessary. Finally, the test program should support model development and validation.

A full-scope integrated test program is not appropriate for all systems. For some critical systems, a continuing evaluation of system capability ultimately saves the cost of planning military campaigns in the face of unnecessary uncertainty, or developing follow-on or replacement systems without knowing how current systems might perform in new or modified roles. Increasingly, however, other developed systems require planning to support system modifications, life extensions, and responses to changing threats. The core components of this form of T&E provide another potential tool in the systems engineering process of developing, fielding, and modifying these systems.

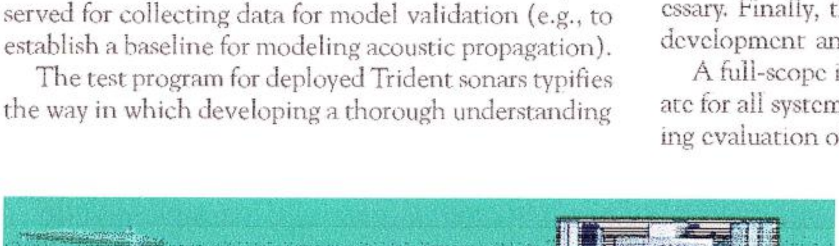


Figure 7. To provide an independent assessment of the capability of Trident sonars, hydrophone outputs are recorded throughout a patrol. These data are analyzed to establish baseline capabilities to compare with demonstrated performance (top right, acoustic recorder; bottom left, Signal Processing Laboratory).

ACKNOWLEDGMENTS: Kevin Currier supplied the information supporting the illustration of model validation for an open-loop missile launch from a submarine. Lee Simkins and James Huneycutt, Jr., supplied the information illustrating the use of models and instrumentation to identify sources of miss distance.

Mission - Profile - PR (new 1/8/96) - Space Dept - *Technical Digests*

The Johns Hopkins University Applied Physics Laboratory

Welcome to The Johns Hopkins University Applied Physics Laboratory (APL) Home Page for the World Wide Web.

The Applied Physics Laboratory is a not-for-profit research and development division of The Johns Hopkins University dedicated to solving urgent problems of national and global interest. While focused primarily on U.S. Navy and defense tasks, the Laboratory also conducts major programs for civil federal agencies, primarily in space science for NASA, and continues to make contributions in civilian sectors, where national needs are in concert with our capabilities. At any given time, APL's diverse organization is working on approximately 200 active tasks for many different sponsors.

With more than 140 experimental laboratories and facilities, APL is recognized for its expertise in research, analysis, simulation and modeling, design, prototype development, and test and evaluation.

Mission Statement

A Profile of APL

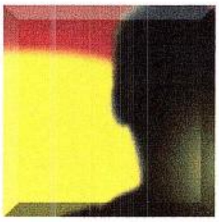
Press Releases (new release dated 1/8/96)

Space Department

The Johns Hopkins APL Technical Digests

APL Contact Information

e-mail: closjs@jhuapl.edu. Updated 17 January 1996.



The Systems Analysis, Test, and Evaluation of Strategic Systems

Larry J. Levy

The systems analysis, test, and evaluation of strategic systems is a "preeminent technical leadership role" for APL. This activity encompasses the planning, design, development, operation, and performance assessment of the Trident II Weapon System, uniquely providing confidence-based performance assessments over untested trajectories. Physics-based scenario-independent statistical models are cumulatively fit to operational tests by maximum likelihood estimation techniques for maximum extraction of model information. The estimated model propagated into the performance factor domain provides the performance factor estimates and computable estimation error statistics for confidence interval estimation. The flight test restricted environment of the present aging weapon system, new global strike missions, and the ballistic missile defense system will present new technical challenges to providing confidence-based evaluations.

INTRODUCTION

The U.S. national leadership recognized that not understanding how well our strategic deterrent (offensive) systems would perform (i.e., with quantified confidence) would be unacceptable, setting specific quantitative guidelines for testing and evaluating of these systems in classified requirements as early as 1966. The need was not just knowing how well the systems would perform but also how confident we could be in our prediction. Quantified confidence is knowing the system's performance to within a quantified uncertainty (confidence interval). It is statistically knowing what you do not know about the system's performance. Building a weapon system with a good performance estimate (e.g., high reliability) but with a large confidence interval (high uncertainty) about that estimate could be dangerous. This was recognized early on in Submarine Launched Ballistic Missile (SLBM) development and testing, requiring a top-down systems engineering approach to define the test programs. Being able to extrapolate test results to the tactical domain became more important as the accuracy requirement became more stringent. Traditional testing by "shoot and score" could not satisfy the top-level evaluation requirements cost-effectively.

In this article, the approach to SLBM test and evaluation (T&E) is presented from three perspectives: (1) the top-down systems engineering approach that produced the T&E system, (2) a description of the T&E system as applied to Trident II, and (3) a discussion of future technical challenges that could be addressed.

SYSTEMS ENGINEERING APPROACH TO TEST AND EVALUATION

APL's systems engineering approach to T&E is shown in Fig. 1. This was extrapolated from experience with previous weapon systems T&E and especially that of Trident II. The approach is discussed generically here to illustrate its use for other weapon systems as well. The left side of Fig. 1 illustrates the planning steps required to properly design an overall test program to provide adequate evaluation capability at certain milestones in the test program. The right side describes the execution steps in the T&E process. This process can be rather elaborate, as it was for Trident, or simpler, as for nonstrategic systems, depending on the system type, stage in the acquisition process, and APL's role.

The key starting point in the systems engineering approach is specifying the top-level performance evaluation requirements (not how well the weapon system should perform, but how well we should know its performance, i.e., confidence). A few test successes do not guarantee that the system will meet its objectives; it only shows that success is possible. If there are no top-level measures of effectiveness (MOEs) evaluation requirements in terms of confidence, then one can be developed. This would be an iterative process involving developer, evaluator, and user.

The next step is to determine a complete set of lower-level measures of performance (MOPs) with associated confidence requirements over a reference set of scenarios needed to achieve the required MOE and confidence bound. Testable MOPs (or ones that are extrapolated from tests) are sampled from distributions commensurate with assumed confidence bounds, and scenario simulations are used to calculate the resulting MOEs (and confidence bounds). This process is iterated until an optimized set of MOPs (and confidence bounds) is achieved. A possible optimization strategy might be to "balance" the contributions of each MOP confidence contribution to MOE confidence. Other strategies might reflect the difficulty (e.g., cost) in achieving certain MOP confidence such as reliability. Many trade-offs could be evaluated.

A test program and analysis methodology are then designed to meet each MOP confidence requirement by hypothesizing various feasible tests (system, subsystem, component), test sizes, instrumentation quality, and evaluation methodologies. Appropriate simulation models (covariance or Monte Carlo) are used to evaluate each hypothesized set until an optimized set is obtained. The results of this phase might require going back to the previous phase to revise the required MOP confidence bounds.

Such a process provides trade-offs while quantifying the implications of decisions to test more (or less), to instrument different functions or systems, or to change the quality of the instruments. As defense spending and costs associated with system development and T&E come under increasing scrutiny, it becomes even more important to be able to quantify the relative benefits of test size and instrumentation quality. Quantifying the confidence with which we will know system performance provides a metric by which we can assess the value of our test programs, instrumentation, and analysis approaches.

To execute the steps of the T&E process (right side of Fig. 1), tests are conducted by traditional testers and evaluators, but with the evaluation outputs complying with the system evaluator's requirements. Test types include system, component, or subsystem tests; monitoring of an in-place system as it awaits operational use; and subsystems assessment "in-the-loop" of a simulation. Detection/isolation of faults on each test is conducted by traditional tester/evaluators, but again with results validated by the system evaluator. Isolated faults are fixed by the developer and removed from the database and models.

The system evaluator calculates a cumulative update of the MOP models, confidence intervals, and estimated distributions. Physics-based models to fit data (system identification) from diverse tests are used where possilbe to gain maximum information from each test. If the model can be broken down into a set of parameters that are independent of scenario, then statistical leverage can be gained by accumulating across all relevant but disparate tests.[1] The associated uncertainty (confidence bound) in the model estimates is calculated from the known observability, instrumentation quality, and number of tests. Prior information and tests from development testing can also be used initially until an adequate number of post-deployment tests can be accumulated. Periodic reassessment of the test program's adequacy to estimate the MOPs and associated confidences may require feedback to the planning stages to reassess the confidence requirements.

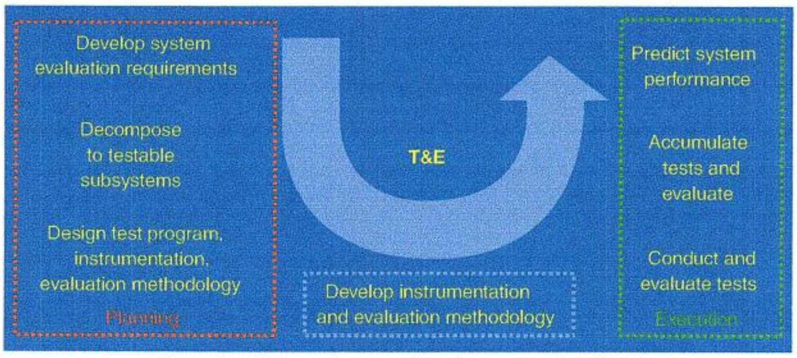


Figure 1. APL's systems engineering approach to T&E.

Next, the system evaluator predicts the MOE and confidence bounds for the required reference set of scenarios using the force-level simulations to flow up the MOPs (and confidences bounds) to MOEs (and confidence bounds). Model fault isolation follows to determine which MOP is out of specification and its resultant contribution to the MOE. Periodic reassessment of the test program adequacy for current MOE requirements must be done.

Finally, the system evaluator conducts force-level evaluations with the latest estimated models by using force-level simulations to flow up the estimated MOPs (and confidences bounds) to MOEs (and confidence bounds) to evaluate the adequacy of the systems for many different campaigns. This allows trade-offs to be made for optimal planning of the force-level deployment such as in ballistic missile defense.[2] The evaluator can also develop and update a functionalized performance prediction model to be used in the real-time employment of the weapon system against an operational threat.

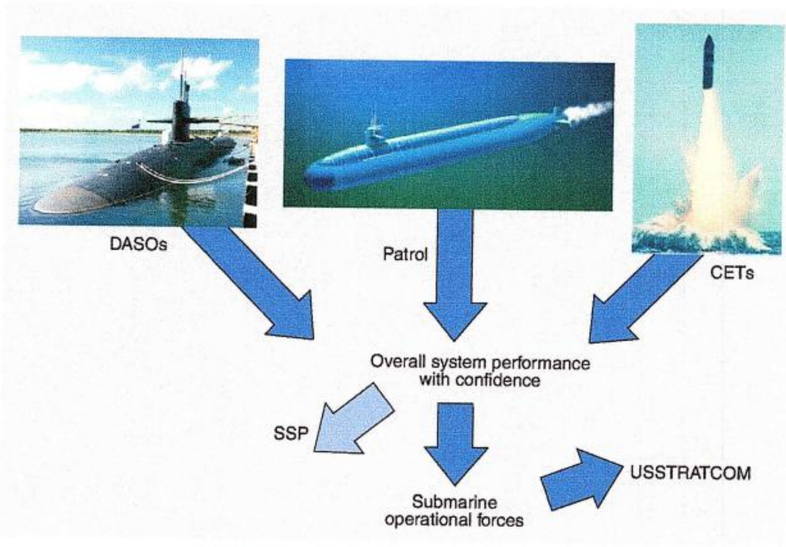


Figure 2. Strategic deterrence systems T&E.

STRATEGIC DETERRENCE TEST AND EVALUATION

Because of the national importance of our strategic deterrent systems, APL instituted a T&E program of the highest caliber that began in the late 1950s for the Navy's Fleet Ballistic Missile Strategic Weapon System, sponsored by Strategic Systems Programs (SSP). The SLBM on its nuclear-powered submarine platform provides a mobile, long-patrol duration, covert, and invulnerable strategic deterrent force. Figure 2 depicts the three major types of system testing of the SLBM: (1) demonstration and shakedown operations (DASOs), i.e., flight testing that is conducted before deployment after either new submarine construction or a shipyard overhaul period; (2) patrol, i.e., recurring nonflight tests conducted during each strategic deterrent patrol; and (3) Commander-in-Chief (CINC) evaluation tests (CETs) or follow-on CETs (FCETs), i.e., end-to-end weapon system tests, including missile flights, conducted with randomly selected missiles periodically throughout the life of the system. The results of the evaluations are provided directly to the Fleet Commands, which then present them to the U.S. Strategic Command (USSTRATCOM) for strategic targeting requirements. In this way APL's T&E is considered "independent" of the developer, SSP.

The scope of these ongoing evaluations encompasses about 220 staff years per year and is the largest concentration of T&E expertise at the Laboratory. SLBM T&E was developed using the full scope of the systems engineering approach described previously. The major S&T innovations—SATRACK, the Accuracy Evaluation System (ACES), and Trident II accuracy—are detailed next.

SATRACK, developed in the late 1970s, uses GPS satellites to precisely track Trident missiles from DASO and CET tests. As illustrated in Fig. 3, the GPS satellite radiates to the test missile containing a GPS translator (instead of a receiver), which relays the raw

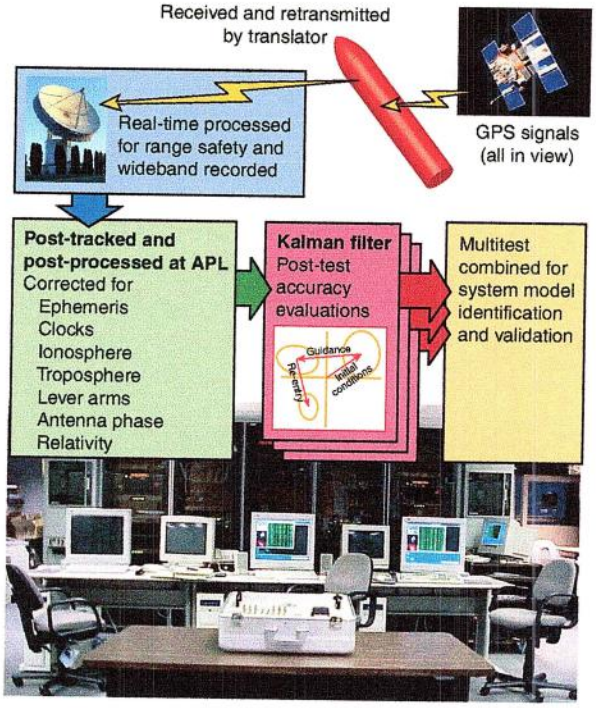


Figure 3. SATRACK for missile accuracy evaluation.

Accuracy Model is based
APL data.

navigation signals to land- and sea-based receiving stations for wideband recording. The recordings are tracked/corrected following the test at the APL SATRACK facility and processed in a large Kalman filter along with missile telemetry for estimation of individual guidance system errors. These estimates can then be propagated to the target point to explain the observed test miss.

Since Trident II was to have a more stringent accuracy requirement, the ACES study, conducted in 1980–1981, used the systems engineering approach to develop system evaluation requirements in terms of accuracy confidence. Instrumentation, test programs, and processing methodology were then determined to satisfy the confidence requirements, resulting in the instrumentation suite shown in Fig. 4. Flight testing then featured an improved SATRACK system for powered flight, inertial instrumentation for deployment and reentry, and improved underwater navigation instrumentation for the prelaunch phase. The major new addition from the ACES study was the cumulative model estimation with confidence, where the per-test results from each test were accumulated via a maximum likelihood method as shown in Fig. 5. Here, a physics-based model of the system, where the unknown parameters are fundamental errors (e.g., gyro drifts) common across all tests, is fit to all the data (even though the test scenarios are different) to estimate the underlying system model and the associated confidence. This results in an estimated model (vs. a validated model) capable of predicting accuracy performance to untested conditions with quantified confidence. The new accuracy modeling, coupled with the traditional reliability modeling, enabled Trident II performance to be predicted with quantified confidence. Starting with Trident I in the late 1970s, more than 180 flights have been processed by SATRACK, with about 100 being Trident II.

CHALLENGES FOR THE FUTURE

SLBM systems will require life extensions, and new missions are being considered such as global flexible response precision strike with low collateral damage. Budget constraints will limit traditional flight testing, requiring new reliability evaluation techniques and

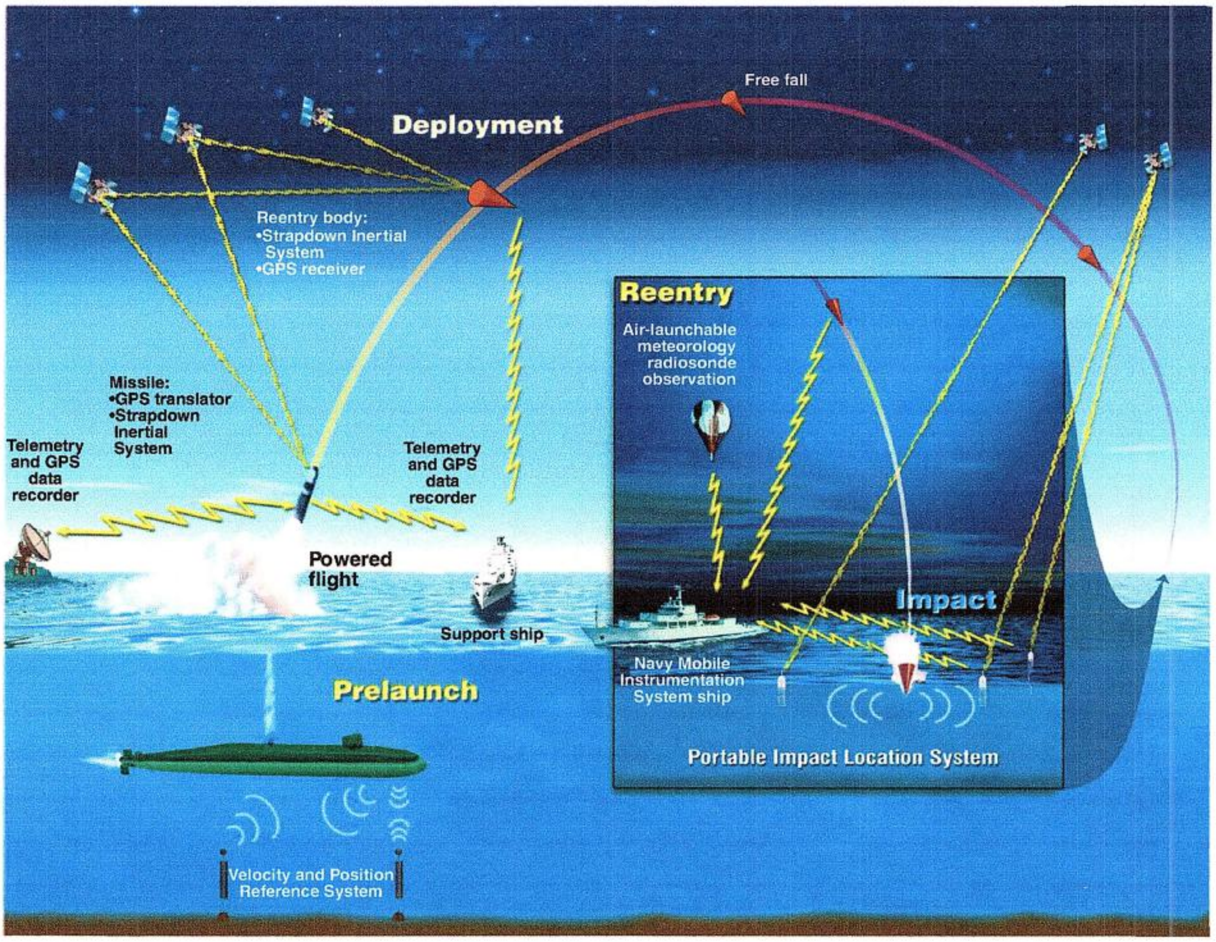


Figure 4. Trident II instrumentation.

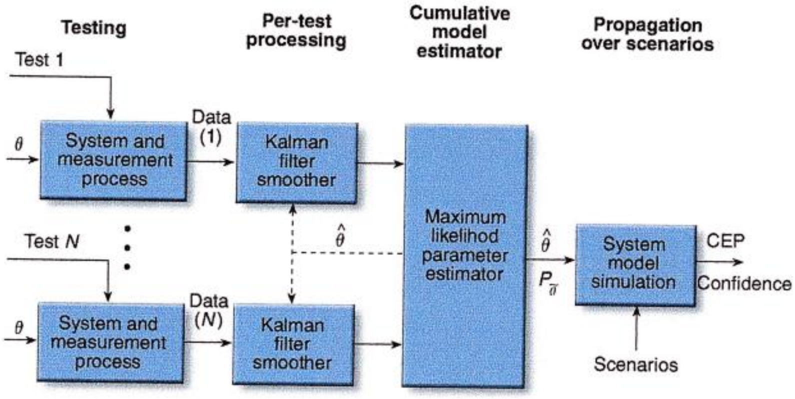


Figure 5. Model estimation for Trident II resulting in the credible performance prediction of a critical system to the government and military system. (θ = true model parameter vector, $\hat{\theta}$ = estimate of θ, $P_{\tilde{\theta}}$ = covariance of estimation error in θ.)

other new testing and instrumentation approaches. Additional test data will be needed to offset the lack of flight testing (there is no "free lunch"). Simulations per se provide no new information. Extensive subsystem ground tests with representative vibration/shock, thermal, and depressurization environments plus centrifuge and high-acceleration aircraft tests can nonsimultaneously replicate the missile environment. New processing methodologies, such as Bayesian Hierarchical Modeling,[3] can be used to appropriately combine ground and aircraft tests with traditional testing. All of these testing and processing methods must be able to provide quantifiable confidence to the performance predictions.

The importance of defending against ballistic missiles with strategic warheads (nuclear and chemical/biological) will require credibility (confidence) in ballistic missile defense performance on the same scale as for our Trident SLBM. This will require a paradigm shift from the traditional defensive systems T&E approach to provide quantified confidence in the performance assessments. The same systems engineering approach to T&E must flow down top-level force-on-force evaluation requirements into detailed subsystem evaluation requirements, followed by appropriate T&E of the subsystems and limited end-to-end tests. All types of testing providing usable performance information will be needed. High-fidelity force-on-force simulations will then propagate scenario-independent parameter estimates and confidences to top-level performance factors.[2] An independent system-of-systems evaluator will be needed to integrate all areas of subsystem T&E with the few available system-of-systems tests.

SUMMARY

Confidence-based performance evaluations of large-scale, complex systems of systems have been demonstrated for the Trident II weapon system, providing a unique approach to systems T&E. It uses detailed physics-based models fit to representative test data to extract maximum information from all relevant tests, providing quantifiable confidence in the model predictions on untested scenarios. Extension of this approach to new critical systems such as ballistic missile defense is possible in principle and necessary to ensure mission success.

REFERENCES

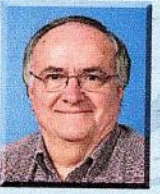
[1] Levy, L. J., and Porter, D. W., "Large-Scale System Performance Prediction with Confidence from Limited Field Testing Using Parameter Identification," *Johns Hopkins APL Tech. Dig.* 13(2), 300–308 (1992).

[2] Mitchell, B. L., and Spriesterbach, T. P., "Confidence-Based Performance Assessments for the BMDO Family of Systems," in *Proc. AIAA 2nd Biennial Nat. Forum on Weapon System Effectiveness* (2001); http://www.stormingmedia.us/72/7226/A722604.html.

[3] Telford, J. K., *Estimating System Reliability with Confidence When Only Limited System Testing Is Available*, SSD-POR-03-7405, JHU/APL, Laurel, MD (Sep 2003).

THE AUTHOR

Larry J. Levy is the Chief Scientist of APL's Strategic Systems Business Area. He received his Ph.D. in electrical engineering from Iowa State University in 1971, specializing in control theory. Dr. Levy has worked in the areas of navigation system development and system test and evaluation. He has more than 37 years of experience in Kalman filtering and system identification and is the author of numerous papers and presentations on these subjects. Dr. Levy was the co-developer of the GPS translator concept in SATRACK (a GPS-based missile tracking instrumentation system) and was instrumental in developing the end-to-end methodology for Trident II accuracy evaluation. He has developed multiple hypothesis fusion methods for multisensor, multitarget tracking and identification. He teaches graduate courses in Kalman filtering and system identification at the JHU Whiting School of Engineering as well as courses on these subjects for Navtech Seminars. His e-mail address is larry.levy@jhuapl.edu.



Larry J. Levy