INSTITUTE FOR DEFENSE ANALYSES

Applications of Modern Statistical Techniques to Operational Testing

Matthew Avery
Laura Freeman
Thomas Johnson
V. Bram Lillard

August 2016
Approved for public release.
IDA Document NS D-8119
Log: H 16-000955

INSTITUTE FOR DEFENSE ANALYSES
4850 Mark Center Drive
Alexandria, Virginia 22311-1882

IDA

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

Test data from the Department of Defense present numerous opportunities for statistical research on applied problems. Operational testing and evaluation is an integral part of the DoD acquisition process. Since conducting realistic tests with representative users in a full operational environment is costly in terms of time, money, and resources, it is imperative that we do it right. That means designing efficient tests, extracting all the information we can out of available test data, and making use of validated models and simulations. The field of statistics offers a variety of tools for addressing these problems and a framework for finding new solutions when existing techniques don't quite fit the unique challenges presented by operational testing. Over the past 5 years, the Institute for Defense Analyses has worked with the Director, Operational Test and Evaluation to make modern statistics the norm. This talk discusses some recent challenges in operational testing relevant to modern statistical approaches and the work that IDA has undertaken to address them.

INSTITUTE FOR DEFENSE ANALYSES

IDA Document NS D-8119

Applications of Modern Statistical Techniques to Operational Testing

Matthew Avery
Laura Freeman
Thomas Johnson
V. Bram Lillard

Executive Summary

A. Summary

Statistical data analysis techniques are a crucial part of operational test and evaluation. Operational testing is a critical part of the Department of Defense's acquisition program. Operational testing provides a realistic representation of scenarios in which systems will be used. Objective analysis of data generated from these tests is crucial for providing the warfighter with accurate information on the capabilities and limitations of these systems prior to their use in combat. Statistical approaches can ensure that the proper amount of data is collected and that the results from analysis are accurate and complete.

B. Experimental Design

Design of experiments (DOE) helps address many key questions when it comes to planning operational tests. Questions of how much data are required and what specific scenarios should be tested are critical when test resources are limited.

The Joint Strike Fighter (JSF) provides excellent examples of the benefits of DOE. The Mission Data File Scan Schedule defines the approach by which the JSF scans the battlefield for different types of emitters, some of which may represent threat radars. Optimizing this scan schedule requires maximizing the timeliness and accuracy of emitter detections while minimizing the incidence of non-detections. This is a many-dimensional problem that DOE is well-suited to address. Robust parameter design and optimal designs allow for an efficient test that maximizes coverage of the range of potential scan schedule settings while accurately estimating performance.

The JSF Initial Operational Test & Evaluation Combat Air Support test plan is an example of how optimal designs can be used to cover a complex design space efficiently. By using a modified D-optimal design, a seven-factor design was created that ensured estimability for all critical two-way interaction effects using only a fraction of the runs that would have been required by a full factorial design.

Improvements in test design have also shown benefits for penetration testing on helmets. Traditional approaches are optimized exclusively around estimating the velocity at which a projectile has a 50 percent chance of penetration. Alternatives allow for efficient estimation of the full curve, which is more useful to decision makers.

C. Data Analysis

Statistical techniques also play a crucial role in data analysis. The Q-53 counterfire radar is a complex system; thus the probability of detecting incoming projectiles, and the accuracy with which the points of origin of these projectiles can be estimated, depend on a variety of factors. Regression allows us to identify differences in performance across this factor space and quantify the impact each factor has on operations.

The Littoral Combat Ship (LCS) has four core mission systems upon which it relies during deployment. These systems work in series, and it is important to account for all of them when assessing overall ship reliability. Since each system functions differently (the Gun System is an on-demand system, while the communications system must be operating continuously, for example), estimating overall reliability can be challenging. Using Bayesian methods, this process is greatly simplified.

D. Conclusion

Statistical approaches improve operational test planning and analysis. These techniques are used across many disciplines, and allow analysts to provide the warfighter with the most complete, most accurate information possible.

Applications of Modern Statistical Techniques to Operational Testing

Dr. Matthew Avery

Dr. Laura Freeman

Dr. Tom Johnson

Dr. V. Bram Lillard

2 August 2016

Operational testing provides assessments of systems in realistic combat conditions

Testing should be realistic and objective and put representative warfighters in operational scenarios

Provide objective information on capabilities and limitations before a system is used in combat

Ensure testing is adequate to support reporting objectives

Statistical techniques should play an important role in operational test planning and analysis

Test Planning: Design of Experiments (DOE)

» Adequate coverage of the operational envelope

» How much testing is enough

» Analytical basis for assessing test adequacy

Data Analysis: Statistical Analysis Methods

– Maximize information gained from test data

– Incorporate all relevant information in analyses

– Ensure conclusions are objective and robust

» Regression Analysis

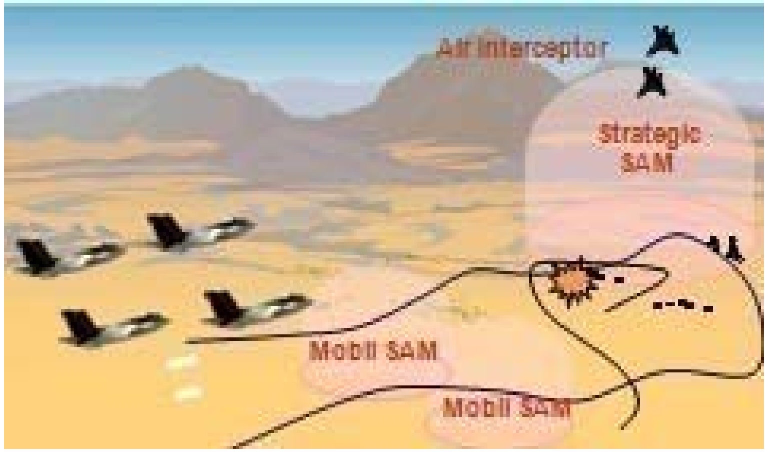
» Hypothesis Testing

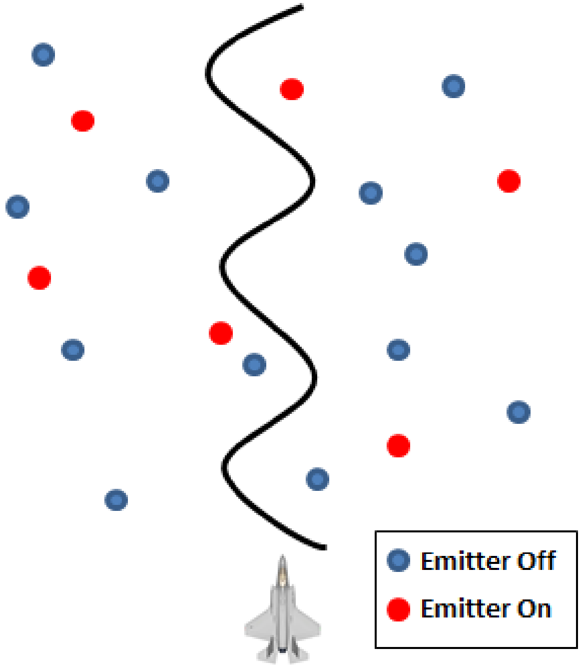
» Confidence Intervals

Efficiently cover the design space to optimize the F-35 Joint Strike Fighter Mission Data File Scan Schedule

Maximize Accuracy and Timeliness

Minimize Non-detections





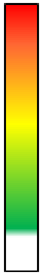
Complex operational space

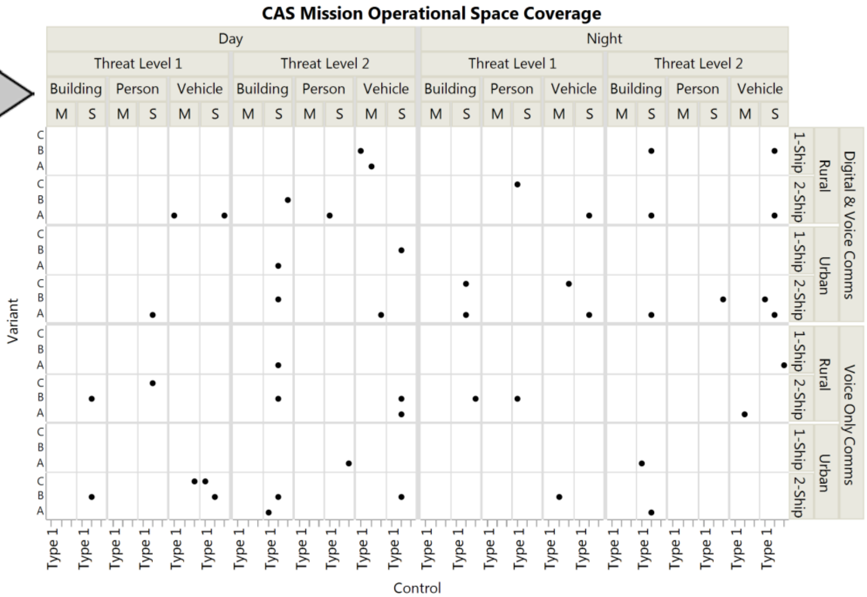
Robust parameter design for the scan schedule settings

Optimal designs to span the complex emitter/ scenario space

Planning the Initial Operational Test & Evaluation (IOT&E) Planning for the F-35 Joint Strike Fighter

	Core Mission Areas				
	Air-Surface	OCA / Strike / SEAD	DCA	CAS	CSAR/ FAC[A]
JSF Variant	✓	✓	✓	✓	✓
Day/Night	✓	✓	✓	✓	✓
Air Threat					
Ground Threat					
Mission Specific Factors					



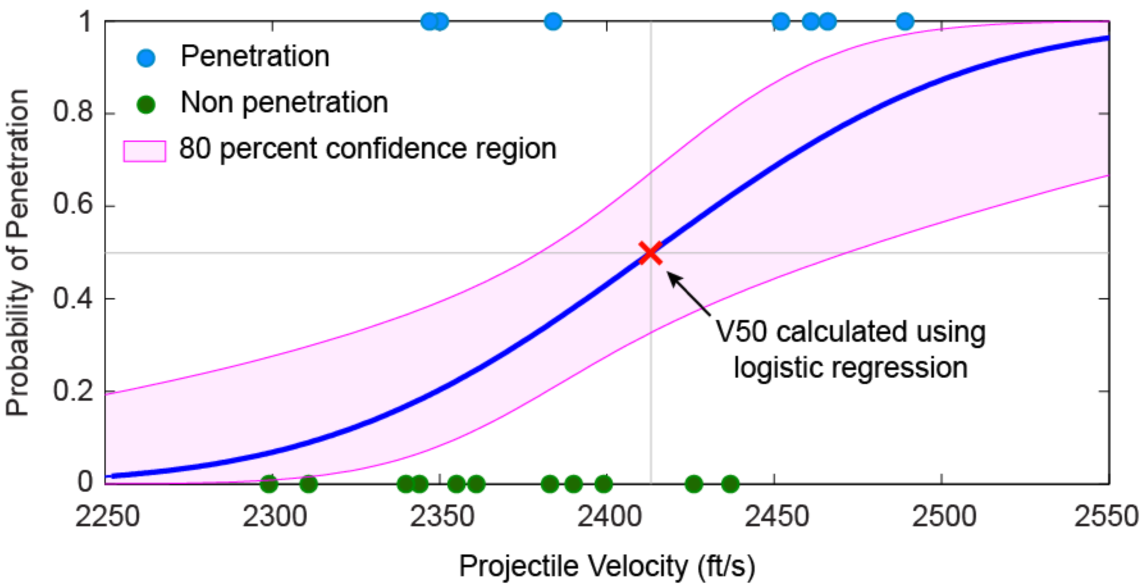


Defensible, efficient methodology for covering the operational space and comparison testing

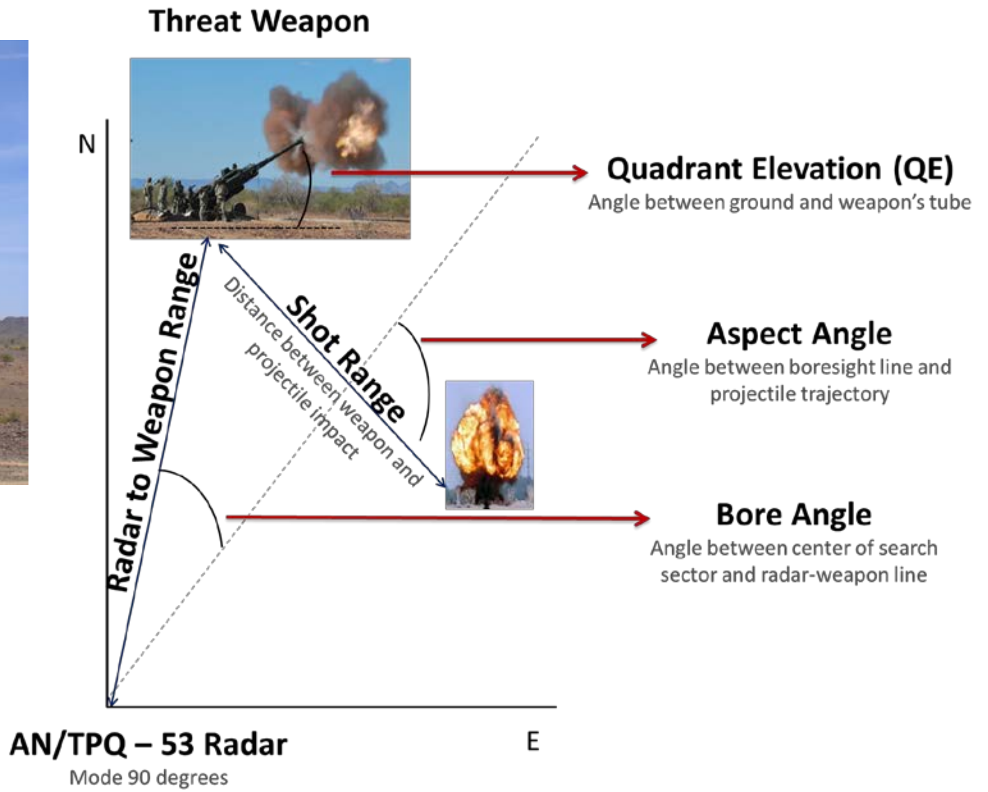
Design tests with the goal of estimating the full probability of penetration curve

Ballistic limit tests often only quantify a V50 point estimate

Tests should characterize the entire response curve and precision



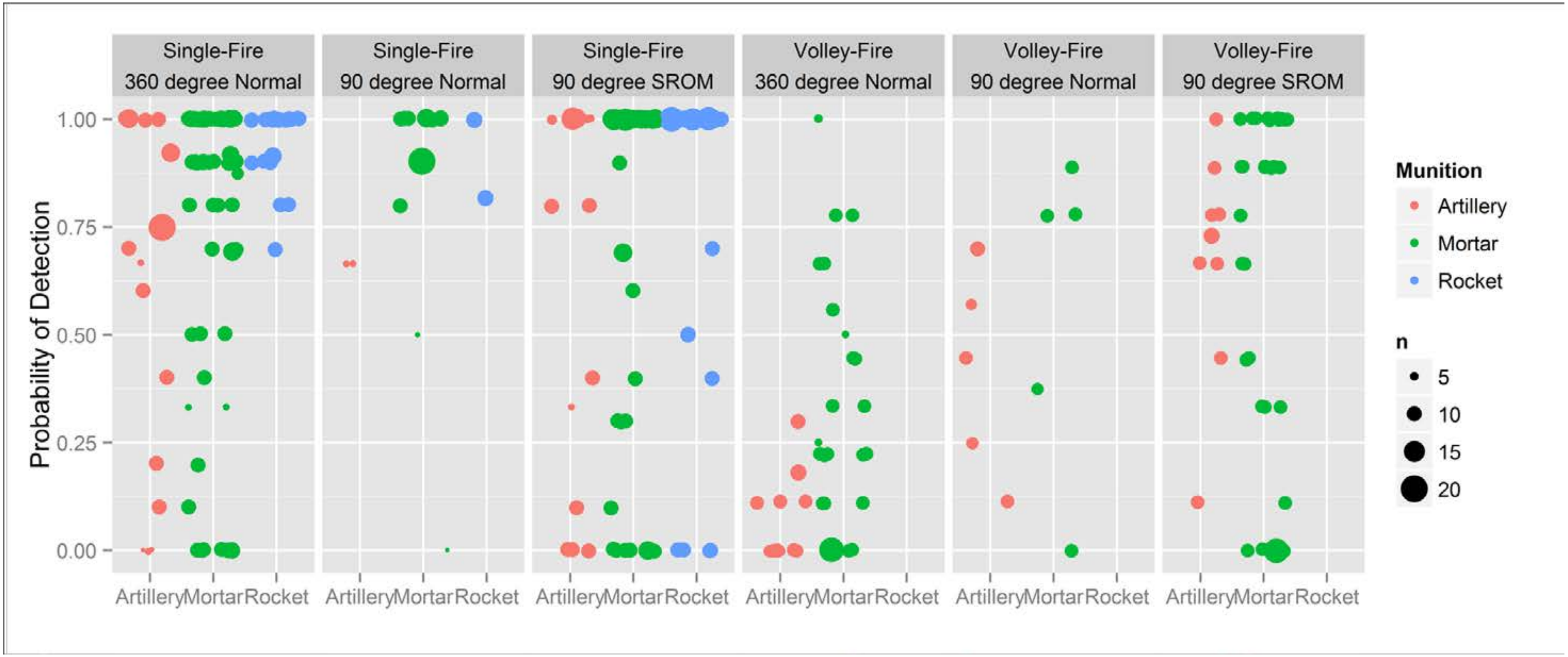
Performance of the Q-53 Counterfire Radar depends on many factors



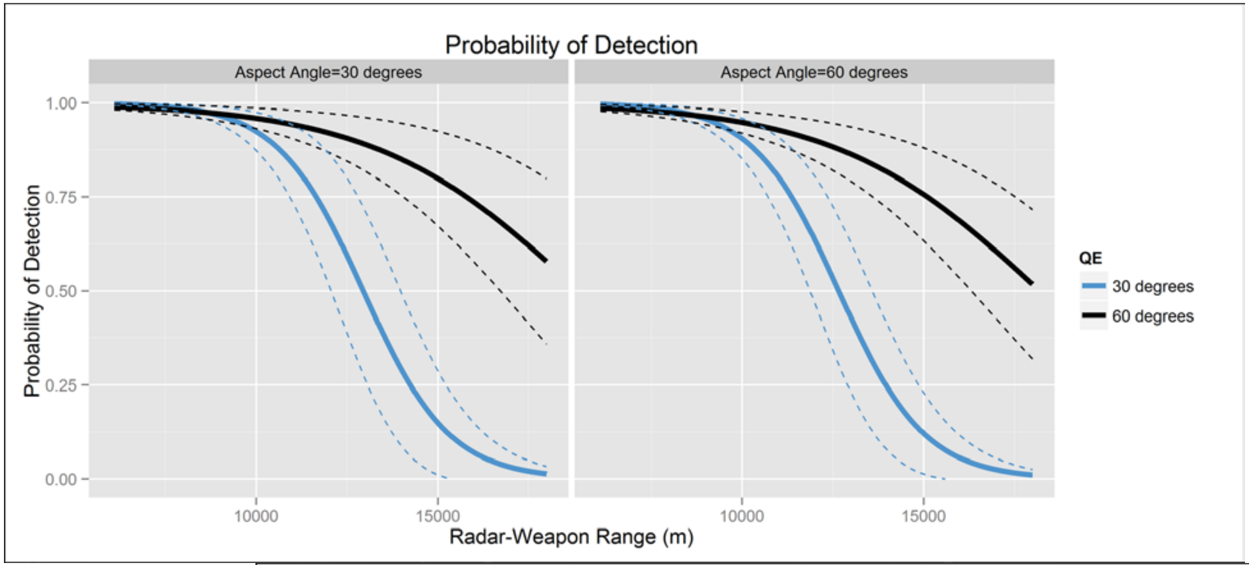
Detects incoming threat projectiles to warn friendly troops

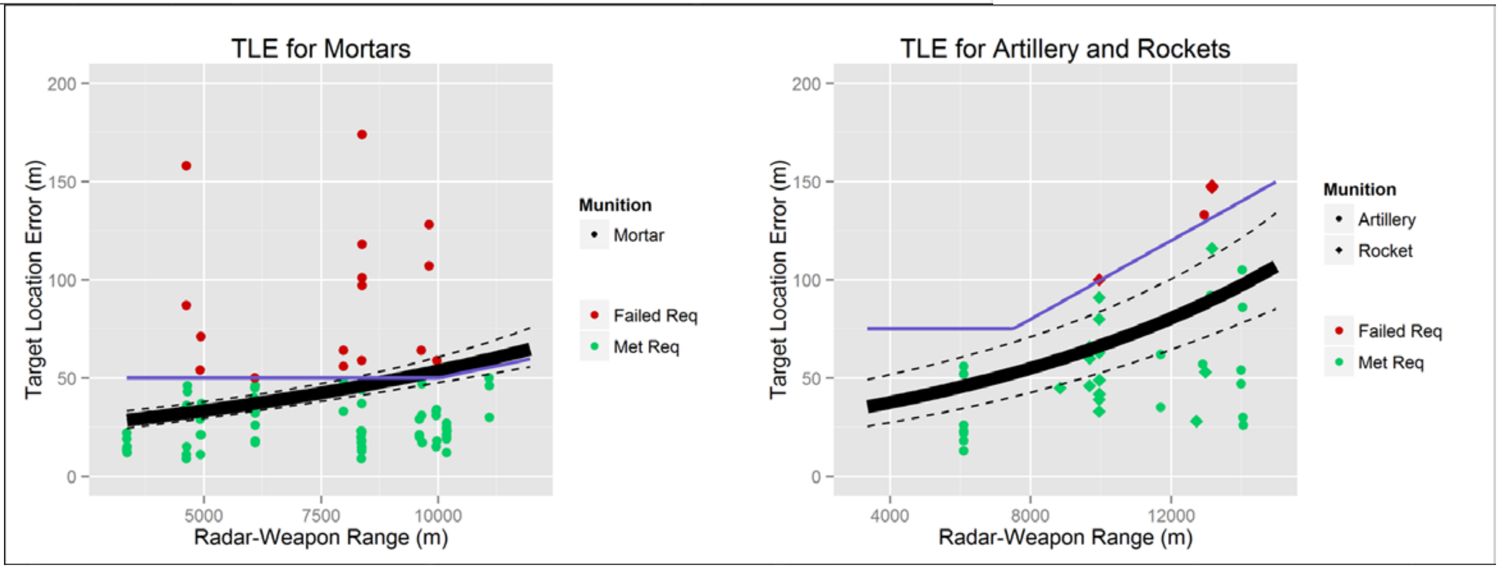
Locates origin of projectiles for counterfire

Data from operational testing is rarely distributed optimally for inference



Modern analytical techniques allow us to characterize system performance across a complex battle space





The Littoral Combat Ship relies an numerous systems working in series

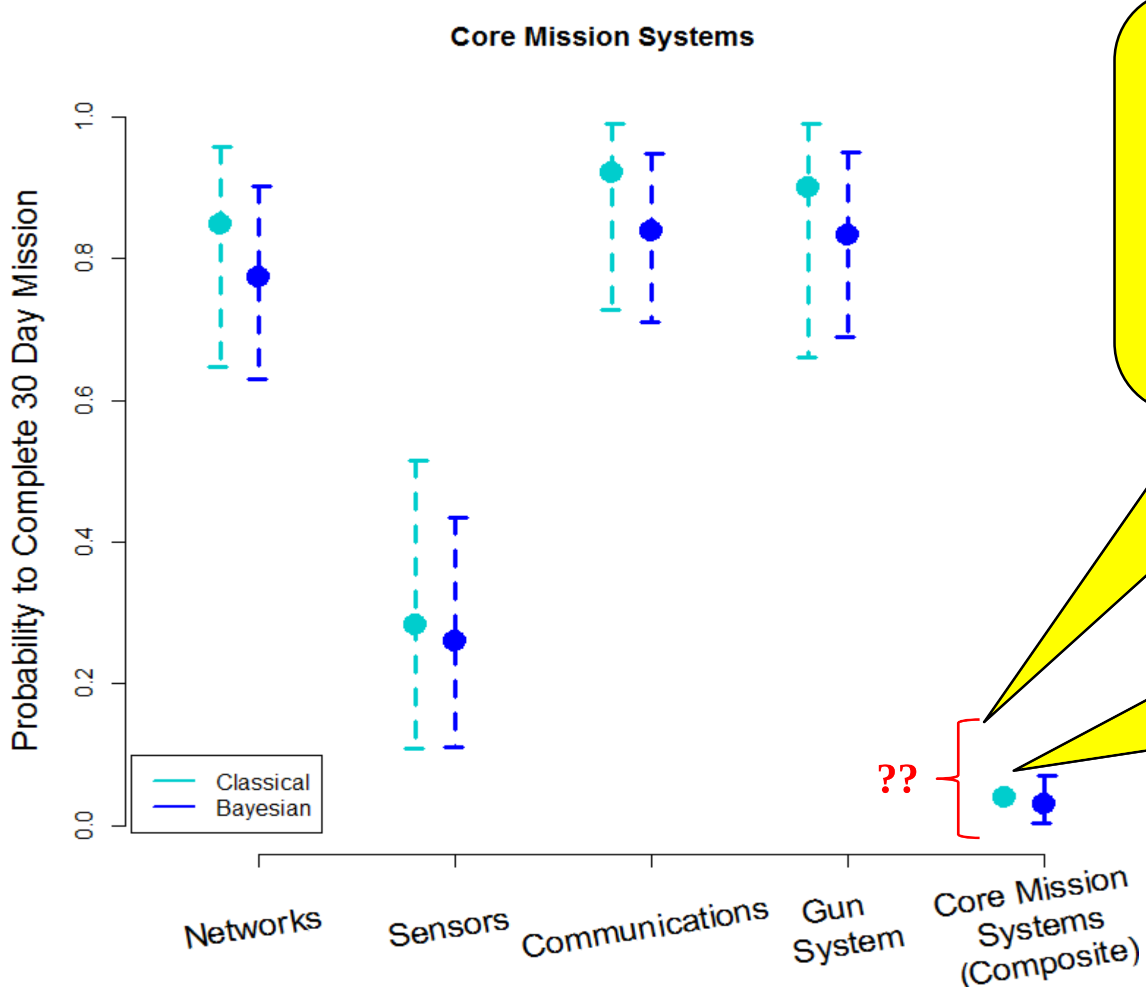
<u>Threshold Reliability</u>
Probability = 0.80
(30-day mission no failures)

Core Mission Systems

- Computing Environment (Networks)
- Sensors
- Communications Systems
- Gun System

Continuous Use   On Demand

Statistical methods allow us to generate point estimates and quantify uncertainty



No universally acceptable classical approach for combining data, estimating uncertainty

Impossible to estimate with classical methods if zero failures observed

Better Science, Better Testing, Better Information

Make Operational Testing More Scientific

- Best practices used across many disciplines
- Conclusions more defensible
- More certain of results
- Test as much as necessary but not more

Provide the best information possible

- Informing the Warfighter
- DOT&E Annual Report
- Beyond Low Rate Initial Production Report
- Congressional Testimony