Applying Advanced Statistical Analyses to Helicopter Missile Targeting Systems

Howard C. Keese and Steven A. Rabinowitz

THE PROBLEM

Advanced analytical methods often extract the essential information from test results that may not have been readily apparent by direct observation alone. When testing the effectiveness of naval helicopters to defend a carrier group from surface attacks, the use of sophisticated statistical methods can provide operators with a greater understanding of both system capabilities and limitations during real-world employment.

DEFENDING THE CARRIER STRIKE GROUP

The United States Navy's Carrier Strike Groups are critical components of our national defense infrastructure. They are also prominent targets for potential U.S. enemies, and are subject to multidimensional threats from air, surface, and subsurface attacks. Consequently, the Navy dedicates significant resources to protecting the aircraft carrier and other high value units at sea.

The end of the Cold War led to a shift in the strategic paradigm for the Navy. It could no longer focus on a single, monolithic threat. In the 21st century, the Navy must be able to adapt to a wide array of disparate regional threats and operating environments. Instead of being primarily concerned with blue-water, open ocean combat, the Navy now also must be prepared to operate in the littorals – in close proximity to the shoreline, which, in turn, exposes U.S. ships to a multitude of new threats. Also important is the radically different mindset of some adversaries. Instead of possessing at least a passing concern with "living to fight another day," some enemies now attack with a suicidal determination. A driven enemy with no regard for personal survival poses a different challenge. One such asymmetric threat is the small boat suicide attack. The grave nature of this threat was dramatically illustrated by the October 2000 attack on USS *Cole*. Although this suicide bombing occurred in port, the Navy is equally concerned about possible small boat attacks at sea involving small arms, missiles, and torpedoes. To counter this threat, the fleet employs a layered defense, with fixed-wing aircraft providing longer range standoff engagements and the ships defending themselves close in. Between these ranges, embarked helicopters provide another defensive layer.

IDA analysts identified and helped construct the test design for the Multispectral Targeting System [MTS] for the operational testing conducted in 2014...[and] identified the dominant factors expected to affect system performance.

NAVAL HELICOPTERS

The Navy deploys two medium-lift, tactical, rotary-wing aircraft aboard carriers and surface combatants: the MH-60R and the MH-60S multi-mission helicopters. Both of these Sikorsky aircraft are derived from the Army's UH-60 Blackhawk, but the construction of each is uniquely tailored to operate in the maritime environment in support of Navy missions. With both radar and sonar sensors, the MH-60R is optimized for antisubmarine warfare. The MH-60S fills combat search and rescue and airborne logistics roles. Both aircraft contribute to surface warfare, providing strike capabilities against small surface targets. Recently, the Navy has been testing various weapons systems aboard these aircraft, including 50mm machine guns, 2.75-inch rockets, and guided missiles.

MTS AND HELLFIRE WEAPON SYSTEM

Both helicopters were designed to employ laser-guided AGM-114 Hellfire missiles (Figure 1). While the Hellfire missile originally was designed for anti-armor land warfare by the Army, the Hellfire's size, range, and lethality have proven useful for a variety of warfare areas. The original plan was to use the MH-60R/S's Hellfire missiles against enemy surface combatant ships, but the missiles can also be employed against small boat targets. Each MH-60R/S can carry eight missiles.

Recently the Navy upgraded the Forward-Looking Infrared (FLIR) system on both aircraft. The new system is known as the AN/AAS-44C(V) Multispectral Targeting System (MTS) imaging system. The MTS uses advanced electro-optic technologies to support navigation, search, and surveillance activities. It can also detect, track, and range surface threats, and its laser designator can illuminate targets to guide Hellfire missiles. The system also has a Day TV capability. When combined with the FLIR camera, the MTS provides imaging from the visible through far-infrared spectrum under all lighting conditions. The MTS features an Automatic Video Tracking (AVT) software algorithm designed to maintain a consistent track on the target and keep the laser designator

Figure 1. MH-60R (left) and MH-60S (right) Helicopters Employing AGM-114 Hellfire Missiles

beam accurately positioned on the target, allowing the Hellfire missile's laser seeker to guide the missile all the way to target impact. Determining the MTS's ability to enable accurate Hellfire employments, therefore, was the focus of the operational test. The critical issue was the MTS's ability to establish a solid engagement-quality lock on the intended target, maintaining laser illumination on the aim point throughout the weapon's time of flight. Therefore, the testing focused on measuring MTS targeting effectiveness across a variety of operational conditions.

TEST DESCRIPTION

In support of the Director, Operational Test and Evaluation, IDA researchers have been employing increasingly sophisticated statistical methods to plan and analyze field tests of critical defense systems. Using the Design of Experiments (DOE) methodology to develop the test plan, analysts identify specific factors that are expected to affect system performance in an operational setting. In the case of a weapon system, these might include the individual and relative motions of the launch platform and target, environmental factors, and weapon-specific data such as firing mode. The ultimate goal is to rigorously characterize weapon performance across the entire operational envelope as a function of those factors, singly and in combinations, rather than simply rolling up the data into an aggregate result. Consequently, instead of reporting out a single overall hit percentage, analysts can demonstrate how particular circumstances and their combined effects may increase or decrease a system's overall effectiveness. Design of Experiments techniques can generate an optimal run plan that provides statistically significant coverage of the various factors without requiring explicit testing of every possible combination. This allows testers to make the most efficient use of limited resources. During the planning phase, IDA analysts work with Service test personnel to determine how to control test scenarios in a manner that provides sufficient data to support factor analysis while preserving operational realism.

Working with the Navy and DOT&E, IDA analysts identified and helped construct the test design for the MTS for the operational testing conducted in 2014. Using both engineering judgment and tactical experience, IDA researchers identified the dominant factors expected to affect system performance. These included target size (large/small), the target's speed (fast/slow), and whether it is maneuvering (yes/no), all of which can be controlled by the run plan. Additionally, target aspect, lighting conditions (day/night), Hellfire targeting mode (target lock before launch/target lock after launch), and airframe (MH-60R/S) were considered in the analysis. Calculating every possible combination of the two levels for each of the seven factors (known as a full-factorial DOE) generated 128 total configurations (2^7), which established the basic data collection requirement for the test.

Although the Navy has not identified specific performance thresholds for the MTS, the helicopter requirements documents specify that the aircraft must be able to fire air-to-ground missiles capable of disabling or destroying a small boat at a standoff range that is beyond the threat of small arms fire and man-portable anti-aircraft missiles that might be carried on the threat boat. Testing of the MH-60R/S with MTS was conducted in the Chesapeake Bay area from August 2012 through January 2013, in two parts. The first phase was a simulated fire period where the helicopters acquired and tracked targets but did not launch any actual missiles. Instead, they carried a Captive Air Training Missile (CATM), which is a specially built Hellfire missile body without a rocket motor. The CATM replicates all Hellfire missile activity up to the point of missile launch, allowing testers to examine the crew's use of the MTS to track and lase the target. During the first phase of testing the targets were emulated by two different types of small fast manned boats: the 26-foot High-Speed Maneuvering Surface Target (HSMST) and the 50-foot Fast Attack Craft Trainer (FACT) (shown in Figure 2).

During the second phase of testing, the helicopters launched five actual Hellfire missiles to demonstrate end-to-end functionality against towed surface targets.

ANALYSIS

In general, data collected from testing are rich with information despite the fact that the results are a simple series of successes and failures (1's and 0's). Several methods exist to analyze such data. One traditional analysis method is to simply tally the successes, divide by the number of trials, and determine the overall success rate. Although this lends itself to simplicity in reporting ("overall MTS is successful X percent of the time"), it may be misleading because it is dependent on the specific allocation of test conditions conducted and might not be representative of a global average of system performance across a variety of future combat conditions. Furthermore, it hides important information about system performance.

Figure 2. The High-Speed Maneuvering Surface Target (HSMST) (top) and the Fast Attack Craft Trainer (FACT) (bottom)

Another methodology groups the 128 data points according to general categories of test conditions. Thus, based on technical experience, analysts might split the data according to threat, calculating the probability of a successful engagement using 64 observations for each of the two target sizes. Next, they might decide to divide the data sample further by target speed, producing four separate results with sample sizes of 32 each, or simply report the average success rate for all fast speed targets and all slow speed targets. This approach may continue for additional factors. Figure 3 plots the data as grouped by three different factors (to avoid revealing classified details the results are normalized and the names of the specific factors are aliased).

Clearly, with each division of the data, we learn more detailed information about system performance, but it comes at the price of statistical confidence. For a binomial (yes/no) response, the confidence intervals, sometimes known as error bars, for small sample sizes can be very large and therefore not particularly informative. In addition, examining the individual point estimate calculations for selected subsets of the data could mask important performance limitations that may exist.

A more rigorous approach uses logistic regression analysis. Logistic regression is similar to the more common linear regression, which predicts performance for a given input values using a linear relationship. Logistic regression employs the same techniques for finding the best fit to the data but is constructed using a more complex function to handle the binary nature of the data and predict the probabilistic outcomes. The logistic regression analysis can be extended to any number of factors (regressors)

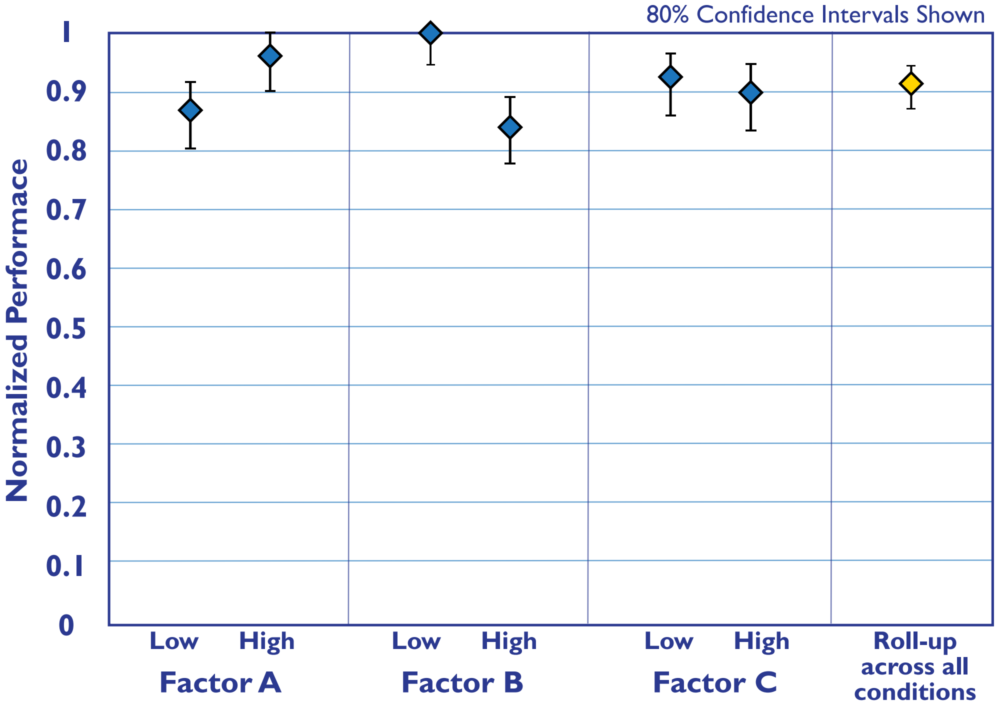

Figure 3. Binomial Point Estimates (Data Grouped by a Single Common Condition)

in order to produce a response surface. The most general form of the MTS logistic regression model is:

$$\log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 Speed + \beta_2 Size \cdots + \beta_k\, Manuever + \beta_{12}\, Speed * Size + \cdots$$

where p is the probability of success, and the β_i's are linear coefficients linking the factors varied in the test to the probability of success. The analysis estimates each β_i. If β_i is not zero (more technically, is statistically significantly different from zero), then the factor or condition is important in explaining the probability of interest. This form of the equation is used because it shows that the factors and conditions impact the "log-odds" of the probability (the left-hand side of the equation) in an additive, linear way. We can rewrite this expression as:

$$p = \frac{\exp(\beta_0 + \beta_1 Speed + \beta_2 Size \cdots + \beta_k Manuever + \beta_{12} Speed{*}Size + \cdots)}{1 + \exp(\beta_0 + \beta_1 Speed + \beta_2 Size \cdots + \beta_k Manuever + \beta_{12} Speed{*}Size + \cdots)}$$

which gives a direct expression for the probability of interest.

In the case of the MTS analysis, IDA researchers utilized 128 data points to construct a regression model that includes all possible interactions between the factors. This technique readily identifies the combinations of factors that result in significant degradation of system performance that would not be easy to isolate through the manual data parsing method discussed above. IDA analysts were able to iteratively build and evaluate different regression models based on various combinations of factors and model terms. The ideal model is the simplest one that includes the most significant factors and their interactions while accurately predicting the system performance based on the data collected. In other words, the statistical analysis is formed and molded by the data alone.

Figure 4 shows the successful engagement predictions of the IDA regression model for MTS based on the data collected in the operational test. In order to mask the classified results, the specific factor names are aliased, and the probability is plotted on a normalized scale. The vertical bars on each performance estimate indicate the confidence intervals (error bars) for each combination of conditions. Note that while these bars are slightly larger than the ones shown in Figure 3, this is due to the finer binning. In fact, by using a regression model, the confidence intervals shown in this plot are smaller than they would be for calculating simple point estimates in each bin. This is because the regression model exploits information from across the data set, resulting in a more precise estimate of the statistical confidence. It is immediately obvious from the plot that the system's performance for most of the conditions is quite consistent. In other words, the MTS for some conditions is not likely to show much performance variation. However, for one particular combination of factors, circled in red, the probability of success is significantly lower. In this case, data exploration via regression modeling allowed IDA researchers to clearly identify a set of conditions that measurably degrade performance. Providing this information to the Navy might allow operators to make adjustments in their employment of

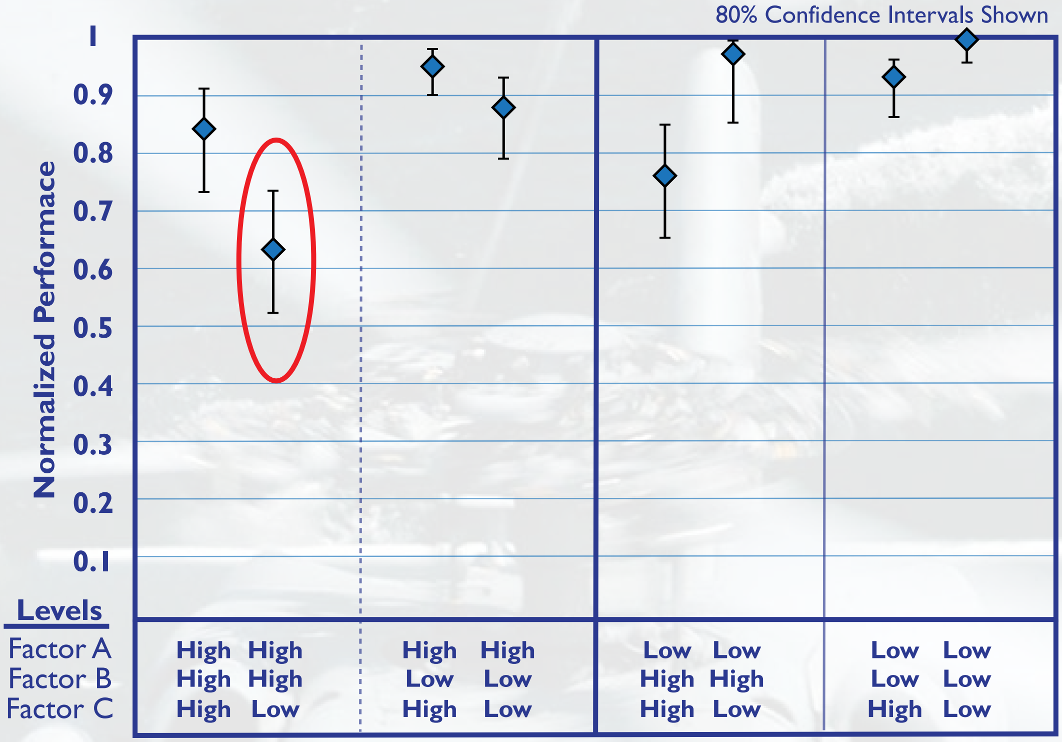

Figure 4. Logistic Regression Results

the MTS while helping the program manager and developers to focus resources on improving the system.

CONCLUSIONS

Properly evaluating system performance is critical to providing effective systems to our armed forces. IDA recently conducted an analysis of a new targeting system for Navy helicopters, applying rigorous statistical methods in order to discern key performance limitations. The resulting analysis provided the operational user with a more comprehensive understanding of their systems and highlighted key characteristics of operational performance that otherwise would not have been apparent. Armed with this knowledge, the Navy can develop the appropriate capabilities-based force structure and the most effective front line tactics, techniques, and procedures to counter the threat and safeguard our forces.

CAPT Keese (USN, ret.) is a Research Staff Member in IDA's Operational Evaluation Division. He holds a Master of Science in national resource strategy from the Industrial College of the Armed Forces, National Defense University and a Master of Science in information systems technology from the Naval Postgraduate School.

Dr. Rabinowitz is a Research Staff Member in IDA's Operational Evaluation Division. He holds a Doctor of Philosophy in physics from Columbia University.