ASSESSING SYSTEM RELIABILITY WITH LIMITED FLIGHT TESTING

Joseph T. Buontempo

The Problem

Expensive flight tests often cannot be conducted a sufficient number of times to yield estimates of system reliability with low uncertainty.

One of the challenges for DoD in developing and fielding the Ground-Based Midcourse Defense (GMD) system is testing the performance and reliability of the Ground-Based Interceptors (GBIs). (The GMD system is described below.) Because GBI flight tests are expensive (on the order of $250 million each), the Missile Defense Agency (MDA) can conduct only a limited number of them—typically one flight test per year. Such infrequent testing would yield considerable uncertainty in estimates of GBI reliability—and therefore in assessments of whether the interceptors are meeting requirements. In the past, similar challenges have arisen in assessments of the reliability of nuclear power plants, nuclear weapons, and some other weapon systems due to limited system-level testing. These situations helped drive the development of the Bayesian methodology for estimating reliability.

IDA examined the use of the Bayesian methodology as a way to estimate GBI reliability given limited flight testing and to reduce the uncertainty and risks associated with these estimates.

IDA examined the use of the Bayesian methodology as a way to estimate GBI reliability given limited flight testing and to reduce the uncertainty and risks associated with these estimates. A Bayesian approach quantifies a starting state of knowledge using a probability distribution (called the "prior distribution"), uses data as they become available to modify the state of knowledge (using the formalism of Bayes' Theorem), and summarizes the resulting state of knowledge with a refined probability distribution (the "posterior distribution").

In determining the starting state of knowledge for estimating GBI reliability, the GBIs can be divided into components (or subsystems), and the reliability of each component can be modeled

The GMD system is intended to engage limited intermediate- and long-range ballistic missile threats in the midcourse phase of flight to protect the United States. The GMD system, which employs GBIs, is supported by multiple sensors that detect and track the ballistic missile threats. The GBI is a three-stage, solid-fuel rocket carrying a 230-pound Exo-atmospheric Kill Vehicle (EKV) toward the target's predicted location in space. Once released from the booster, the EKV uses data received in-flight from ground-based radars and its own on-board sensors to attempt to close with and, using the kinetic energy from a direct hit, destroy the target outside Earth's atmosphere.

with its own probability distribution. For illustration, one can make the initial simplifying assumptions that the starting state of knowledge of each component is identical; the initial reliability of each component is unknown; and the components have independent reliabilities. The product of the component reliabilities, therefore, is the system reliability. Given these assumptions, the prior distribution can be produced for GBI reliability, and as new data become available, the reliability distribution for each component—and thus the GBI system—can be updated using Bayes' Theorem (see the callout box for details).

In general, the Bayesian methodology is flexible enough to capture other (more useful) starting states of knowledge and to incorporate many different kinds of data. Regarding the former, different prior distributions could be specified for each component, for example, rather than assuming that all components start with the same uncertainty about reliability. Regarding the incorporation of different kinds of data, appropriate models could allow the use of data obtained from modeling and simulation, ground tests, bench tests, and correlated failures. Through these methods, an initial prior distribution can be constructed that would enable more certain reliability estimates given a limited number of system-level tests (see Figure 1).

Thus, the use of the Bayesian methodology and other sources of data in addition to infrequent flight tests can reduce uncertainty in the estimates of GBI reliability and help identify and fix failure modes more rapidly. This analytic process provides a pathway to help reduce risk and enable assessments of whether the

The reliability of each GBI component can be modeled with a two-parameter (α, β) distribution called a beta distribution. The beta distribution gives the probability density of a value x (here reliability) on the interval [0, 1]:

$$Prob\ (x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)},$$

where $\alpha > 0$, $\beta > 0$, and $B(\alpha, \beta)$ is the beta function (which serves as a normalizing constant to ensure that the total area under the density curve equals 1).

The choice of parameter values for the GBI beta prior distribution depends on the assumptions regarding the starting state of knowledge (see main text) and the number of GBI components (*n*). If the initial reliability of each component is assumed to be unknown, the initial values of the two parameters would be chosen to fit a uniform distribution between 0 and 1 (roughly, all values between 0 and 1 are equally likely). The actual parameter values, as derived by Redd and Reese [1], following Goodman [2], are:

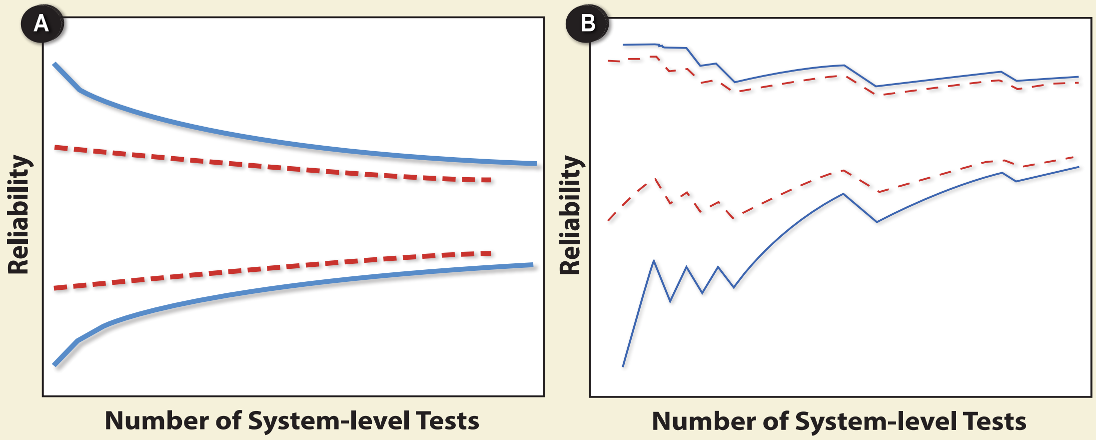
$$\alpha = \frac{(2/3)^{1/n}-1}{1-(4/3)^{1/n}}, \beta = \alpha \cdot (2^{1/n}-1).$$

This starting state of knowledge is updated with new data using Bayes' Theorem, which shows that the posterior distribution is equal to the prior times a factor that is dependent on the data. If we make the simplifying assumption that each trial is successful with the same probability, this implies that the number of successful trials for a particular component can be described by a binomial distribution. Using the beta distribution to describe the prior distribution and a binomial distribution to describe the data leads to a posterior distribution that is also a beta distribution.

Sources:

1 Redd, T. and S. Reese, Brigham Young University, personal communication.

2 Goodman, L., "The Variance of the Product of K Random Variables," *Journal of the American Statistical Association*, Vol. 57 No. 297, March 1962.



A provides a notional depiction of the reduction in uncertainty. The solid lines represent estimates based only on data from the system-level tests, and the dashed lines represent estimates that also include other suitable sources of data using the Bayesian methodology. B shows a similar set of results obtained for a simple problem* and contains noise and asymmetric behavior (the lower bound increases more rapidly than the upper bound decreases due to the nature of the problem). In these illustrations, the uncertainty estimates for a limited number of system-level tests are smaller when one incorporates a starting state of knowledge (based, for example, on component-level testing and/or modeling and simulation) prior to the onset of system-level testing. Over a sufficient number of system-level tests, the solid and dashed lines approach similar estimates of uncertainty. However, in practice, conducting a sufficient number of tests might not be possible, and thus the inclusion of other suitable sources of data can provide smaller uncertainty for a limited number of system-level tests.

* For B, results were randomly generated for 30 "tests" (observation of 1 for "success" and 0 for "fail") with probability of success equal to 0.75. For the estimates using the Bayesian methodology, the starting prior had 90 percent probability that the probability of success was between 0.5 and 0.95. The solid lines are the upper and lower classical 90 percent confidence interval. The dashed lines are the upper and lower bound of a Bayesian credible interval. These intervals, confidence and credible, correspond to our "uncertainty estimates." *[Acknowledgment: IDA Research Staff Member Alyson Wilson]*

Figure 1. Reduction in Uncertainty of Reliability Estimates with Increasing Number of System-level Tests

GBIs are meeting reliability requirements. In addition, projections of future reliability can also be made by employing additional assumptions involving, for example, the degradation of GBI reliability due to age, the occurrence of unforeseen failure modes, and the effectiveness of repairs.

IDA researchers are also employing Bayesian methods to estimate reliability for other military systems. For example, since there are many similarities among the vehicles in the Stryker family, we used Bayesian hierarchical models to estimate the reliability of the family based on data obtained for the different vehicles and across both developmental and operational tests. These models let the data determine the appropriate weighting of information across vehicle variants and test phases. The combination of information improves estimation and reduces uncertainty.

IDA is also examining the use of Bayesian methods for improving the estimation of reliability for the Ground Combat Vehicle. These methods are particularly well suited to integrating heterogeneous information sources. We are starting by considering information obtained from modeling and simulation to develop prior distributions that summarize what we anticipate seeing in subsequent testing. This will help with test design and potentially reduce uncertainty in the reliability estimates.

For many military systems, conducting a sufficient amount of system-level testing in operational environments to provide reliability estimates with low uncertainty can be quite expensive. The use of Bayesian methods can allow researchers to design test strategies involving a limited number of tests and to incorporate other sources of data—such as modeling and simulation, ground tests, and bench tests—that could help reduce uncertainty and risk.

Dr. Buontempo is an Assistant Director in IDA's System Evaluation Division. He holds a doctorate in physical chemistry from the University of Chicago.