test data is not straight forward, requiring advanced system modeling techniques.

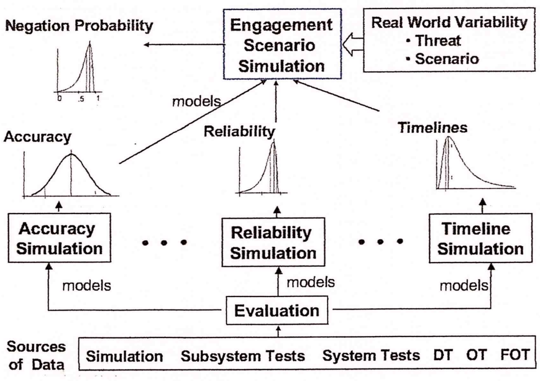

Figure 2 - BMD MOE Evaluation with Confidence

Importance of Confidence in a TMD Example

A related paper, Mitchell et.al. (2001), provides an example of the importance of confidence in TMD. Starting with assumed estimates (called expected values in the paper) and associated distributions for accuracy, reliability and timelines, the Extended Air Defense Simulation (EADSIM) was used in a Monte Carlo mode, as illustrated in Figure 2, to evaluate a fictitious theatre engagement scenario. It was found that the variations in the sampled distributions could sometimes cause the FoS to perform radically different than predicted by just the projection of the assumed estimates. In other words, the reasonably possible statistical variations in the associated lower level distributions caused significant tails (and even multiple modes) in the $\hat{P}_n$ distribution. So it is possible that one could have predicted a reasonably high $\hat{P}_n$ with the true value being significantly lower, a potentially disastrous result! Again, a properly constructed test program must be developed so as to achieve sufficiently close confidence bounds to the truth.

Confidence in Trident II Accuracy Prediction

Goals for Trident II accuracy evaluation were specified in IDA/WSEG (1966) and the evaluation requirements were specifically defined in US Strategic Command (1998). The requirements specified quantified confidence goals for top-level MOE estimates of reliability and accuracy for initial performance estimates and change detection with time. For brevity, only accuracy evaluation will be described.

The process followed very closely the steps outlined in the next section except it was applied to the MOE of target accuracy. An overview description is given in Simkins et.al. (1990). New evaluation methodology (a satellite missile tracking system and maximum likelihood system identification for modeling) was developed to minimize system tests with greater functionality. Thirty system tests were needed using the traditional ("shoot and score") evaluation approach with only ten tests needed with the new methodology for initial model estimation. Ten tests were needed using traditional evaluation to four tests using the new methodology for detection of model changes in follow-on testing. Only the new methodology enabled extrapolation to untested conditions. Individual guidance error models and launch area gravity models were corrected. Increased system understanding was obtained to accurately predict performance over long-range non-tested trajectories. The estimated Trident II performance was considerably different than was expected. This would not have been known or understood with the traditional approach. This has enabled test-based predictions of capability to support other (non-traditional) missions & requirements.

Conceptual Application to T&E of BMD

The systems engineering approach to test and evaluation of BMD with confidence is shown in Figure 3. This was extrapolated from experience with many previous weapons systems T&E and especially that of Trident II. The left side illustrates the planning steps required to properly design an overall test program to provide adequate prediction confidence at certain milestones in the test program.

The key starting point is specifying the top level Performance Evaluation Requirements (not how well the weapon system should perform, but how well should we know it) in terms of required specifications (e.g. negation probabilities for realistic overall force level scenarios). At present, there does not appear to be "official" evaluation requirements on how well we must know P_n as there is for Trident II accuracy and reliability. This will be a serious impediment to successful employment of the BMD system. A few test successes does not guarantee that the system will meet its objectives; it only shows that success is possible.