FORMULATION OF A METHODOLOGY FOR THE PROBABILISTIC ASSESSMENT OF SYSTEM EFFECTIVENESS

Danielle S. Soban†, Dimitri N. Mavris‡
Aerospace Systems Design Laboratory
School of Aerospace Engineering
Georgia Institute of Technology, Atlanta, GA 30332-0150
Atlanta, GA

Abstract

There exists a need for an integrated and efficient framework that can rapidly assess system effectiveness for today's complex systems. The analysis of air superiority in the theater environment is considered here, and the steps taken in the formulation of a cohesive methodology are presented. "System" and "system effectiveness" are clearly defined. The formulation of the new framework is based on existing probabilistic methodologies that define the aircraft as the system. Extrapolation of these methods to the theater level is proposed, redefining the system as the total warfighting environment, in which the aircraft becomes a system component. In the first part of the formulation, presented here, key issues are identified that must be addressed, and initial solutions to these issues are proposed, but have not yet been implemented. Current probabilistic methods were used to analyze an example scenario and operational situation, and this test case was used to identify the issues mentioned above. Results of this test case are provided, and indicate the feasibility of using these methods at the theater level.

Introduction

Assessing the success and effectiveness of today's complex systems becomes an increasingly challenging problem. System demands, including increased performance, lower system life cycle costs, longer operating capacities, and improved productivity and efficiency, must be balanced against limited resources, scant or unknown data, the identification and resolution of conflicts, and resource allocation. These tradeoffs point to the need for an integrated and systematic framework that can assess system effectiveness by identifying potential problem areas and aid in resource allocation and decision-making processes. Specifically, a framework needs to be established that aids in analyzing air superiority in the theater environment.

† Ph.D. Candidate, Student Member AIAA

‡ Boeing Chair in Advanced Aerospace Systems Analysis, Senior Member AIAA

Definitions of Concepts

In order to formulate such a framework, it is important to understand and clearly define the concepts of both "system" and "system effectiveness". There is general agreement across fields and disciplines as to what constitutes a system. The following definition is representative of this agreement, and is an acceptable definition for the developing framework:

> *A* **system** *may be considered as constituting a nucleus of elements combined in such a manner as to accomplish a function in response to an identified need…A system must have a* **functional** *purpose, may include a mix of products and processes, and may be contained within some form of hierarchy…*[1]

However, the definitions of system effectiveness vary widely and are often application dependent. Some examples that illustrate the diversity of these definitions include:

> "The overall capability of a system to accomplish its intended mission"[2]

> "The probability that the system can successfully meet an operational demand within a given time when operated under specified conditions"[3]

> "A measure of the degree to which an item can be expected to achieve a set of specific mission requirements, and which may be expressed as a function of availability, dependability and capability"[4]

The authors of an annotated bibliography on system effectiveness models in 1980 concluded "A wide range of definitions, and measures of system effectiveness are used without strong guiding logic."[5]

A new definition for system effectiveness, therefore, must be justified by identifying key elements crucial to a useful and informative definition. First, the term "effectiveness" implies that some sort of *quantification* needs to occur. This quantification must necessarily be the result of some sort of systematic *analysis* of variables and metrics that represent the system performing its function. In addition, in order to perform the quantification, an *intended or expected effect* needs to be identified in order to properly model the results of the system performance. Combined, these concepts result in the following definition put forth for use in formulating the framework for the probabilistic assessment of system effectiveness:

> *System effectiveness is a quantification, represented by system level metrics, of the intended or expected effect of a system achieved through functional analysis.*

The Aircraft as the System

The Aerospace Systems Design Laboratory (ASDL) at the Georgia Institute of Technology has formulated probabilistic methodologies that treat individual aerospace concepts as the system[6,7,8,9]. Metamodels based on regression methods have been created relating vehicle design variables (geometry, engine specifications, drag polars, etc.) to vehicle responses (takeoff gross weight, thrust-to-weight ratio, etc.). Further advances in the methodology added economic variables, requirements and mission constraints, as well as allowing analysis of the effect of new technologies. In this way, the aircraft itself was treated as the system, and aircraft Measures of Performance (MOPs) were the metrics used to assess the "goodness" of the system (Figure 1). System effectiveness, therefore, was a function only of that aircraft's design variables and parameters.

While this method of analysis resulted in the design of a vehicle that was optimized to its own mission and performance requirements, the vehicle remained independent of its role in the larger system (in this case, the warfighting environment). In other words, the aircraft was never placed in its correct context and evaluated as system fulfilling its intended function.

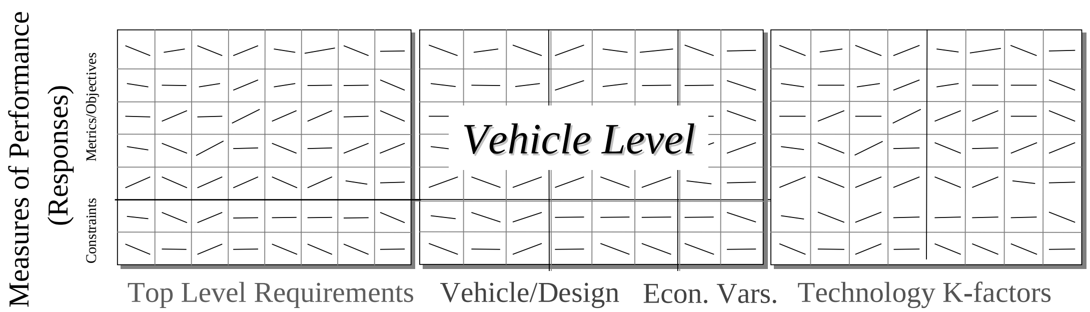

Figure 1- Analysis of the Aircraft as the System

System of Systems Approach

In order to place the aircraft in its correct context, the system must be redefined. No longer is the aircraft the system; rather, the theater becomes the system, and the aircraft is relegated to being a component of this new system. System effectiveness is thus considered at the theater level, and becomes a function of the vehicle-level components of the system, including the aircraft.

There is, however, a missing level in this formulation. The outputs of the vehicle level (performance parameters) do not usually map directly as the inputs to theater level modeling codes. Rather, the inputs at the theater level usually consist of probability of kill values or effectiveness values for component vs. component encounters. There must be an intermediary mapping that takes the outputs of the vehicle level as its inputs, and in turn generates outputs that serve as inputs to the theater level. This is illustrated in Figure 2. Theater level Measures of Effectiveness (MOEs) are functions of vectors of subsystem Measures of Performance, which in turn are functions of subsystem requirements, design and economic variables, and technology factors.

When the methodology is complete, there will exist a continuous mapping between vehicle level design parameters and theater level Measures of Effectiveness. Changes at the vehicle level can thus be propagated all the way to the theater level. Instead of optimizing an aircraft, for example, to its own pre-defined performance and mission constraints, the aircraft can now be optimized to fulfill theater level goals and objectives. In addition, as more system level components are treated as input variables, tradeoffs can be established not only at the individual component level, but *across* the components. In other words, the methodology will allow tradeoffs between, say, the effectiveness of a cruise missile compared to an aircraft carrying a specified weapons load. Tradeoffs could also be made between the number of system components needed: two of aircraft "A" could produce the same effectiveness of five of aircraft "B", but at less cost. Thus, the methodology becomes a key device for design decisions as well as resource allocation.

Finally, the completed methodology can be used to actually determine the mission and design requirements for the vehicles themselves that comprise the system. By using the Measures of Effectiveness at the theater level as a measure of goodness, tradeoffs can be made between vehicle design and mission requirements. These requirements, when optimized to maximize the overall effectiveness of the system, become the requirements to which the vehicles are then designed.

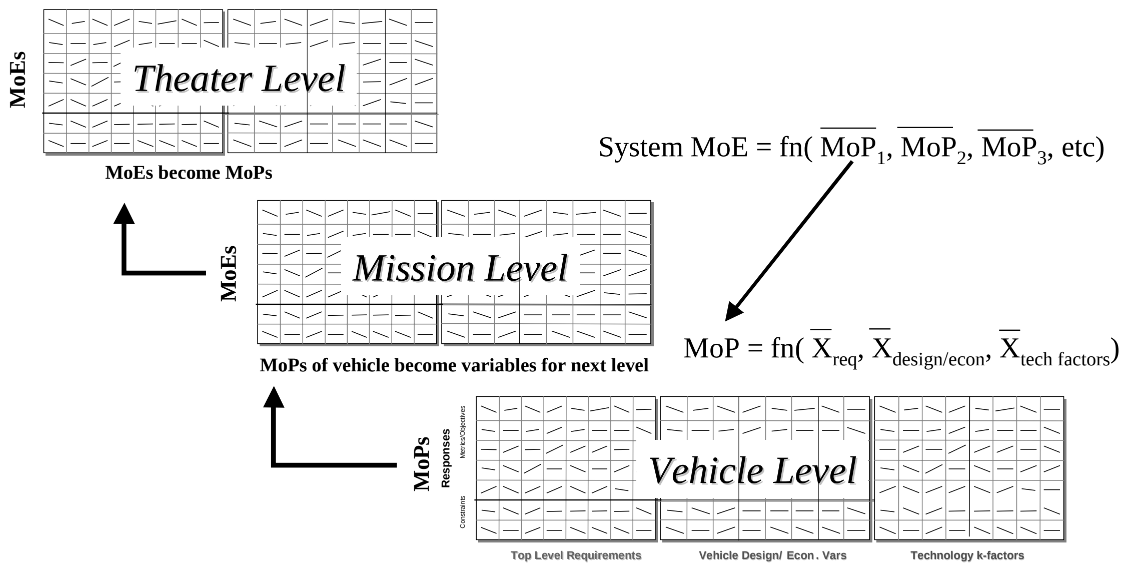

Figure 2- Mapping and Sensitivity Cascades Relating Vehicle Characteristics to Theater Level MoEs

Current Methodology

The new methodology will be proposed as an extrapolation of current existing ASDL probabilistic methodologies from the vehicle system level to the theater system level. This section will outline the key concepts and components currently in use, and is a summary (with some verbatim parts) of sections presented in Reference 10.

Design of Experiments and Response Surface Methodology

The cornerstone of the ASDL probabilistic methodologies is Response Surface Methodology (RSM) combined with Design of Experiments (DOE). RSM is an efficient, multivariable approach to system modeling that defines clear cause-and-effect relationships between design variables and system responses, and is based on a statistical approach to building and rapidly assessing empirical metamodels[11,12].

The RSM methodology, employing a DOE strategy, creates metamodels of a particular synthesis code by selecting a subset of all possible combinations of variables to run which will guarantee orthogonality (i.e. the independence of the various design variables). Using regression techniques, the subset of inputs are related to selected outputs to create an equation that represents the relationship between inputs and outputs of the synthesis code. This technique allows the maximum amount of information to be gained from the fewest number of experiment executions, and thus provides trade study results in a more cost-effective manner.

The first step in the creation of the metamodel is to select an appropriate Design of Experiments. This DOE is expressed as a table of experimental cases, specifying the values of the variables to be used for each individual execution of the synthesis code. These values are usually normalized to a low, high, or midpoint value of the variable (represented by a –1, 1, and 0 to aid in the statistical analysis). An example DOE table is shown in Table 1. Typically, the response is first modeled using a second order quadratic equation of the form:

$$R = b_o + \sum_{i=1}^{k} b_i x_i + \sum_{i=1}^{k} b_{ii} x_i^2 + \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} b_{ij} x_i x_j$$

R is the desired response term
b_0 is the intercept term
bi are regression coefficients for the first order terms
bii are coefficients for the pure quadratic terms
bij are the coefficients for the cross-product terms
x_i ,x_j are the independent variables
k is the total number of variables considered

This equation is called a Response Surface Equation (RSE). Other forms of the equation may be used (for example, during a screening test, a first order linear regression is appropriate). If the non-linearities of the problem are not sufficiently captured using this form of the equation, then transformations of the variables and/or the responses need to be found which improve the fidelity/accuracy of the model.

A Response Surface Equation (RSE) is created by executing multiple runs of the synthesis code, with each execution using as its inputs the values of the variables determined by the DOE table. The resulting responses of interest for each run are then collected from the output and added to the table (the blank columns in Table 1). A statistical analysis package (in this case, JMP[13]) provides the ability to take this data and perform a regression analysis to create these polynomial representations (Analysis of Variance or ANOVA) to determine these sensitivities, relative importance, fidelity, etc. JMP also aids in providing the experimental setup, as well as facilitating visualization of the results. There is one Response Surface Equation created for each response, and this equation is a function of all input variables. The resulting RSEs, thus, are in actuality metamodels of the synthesis code used in their creation. The equations represent a quick, accurate way of determining a response for given values of variables (as long as these values are within the range of variables for which the RSE is defined).

Table 1- Example Design of Experiments Table

CASE	Wing Area	Sweep	Engine Scale Factor	Response 1 (R_1)	…Response n (R_n)
1	-1	-1	-1		
2	-1	-1	1		
3	-1	1	-1		
4	-1	1	1		
5	1	-1	-1		
6	1	-1	-1		
7	1	1	-1		
8	1	1	1		
9	-1	0	0		
10	1	0	0		
11	0	-1	0		
12	0	1	0		
13	0	0	-1		
14	0	0	1		
15	0	0	0		

The Response Surface Methodology is comprised of two basic steps, facilitated by the program JMP. The first is referred to as the effect screening. It creates a linear model which is used to determine the sensitivity of a response to various inputs and to screen out, using a Pareto analysis, those variables that do not contribute significantly to the variability of the response. The second step is called surface fitting, and yields a polynomial representation that gives the response as a function of the most important input parameters. These steps are illustrated in Figure 3.

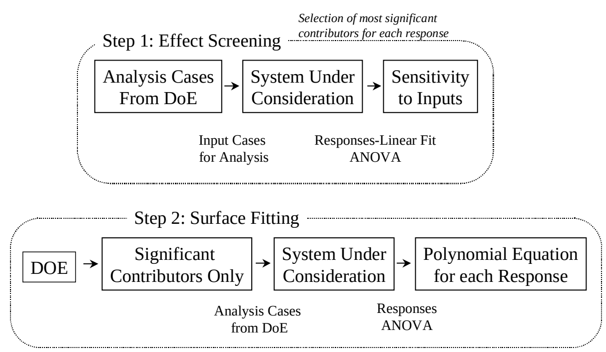

Figure 3- Basic Steps of Response Surface Methodology

The benefit of RSM is that it provides an almost instantaneous evaluation time. The equations are portable and can be run in a spreadsheet, a computer code, or even by hand. Within the variable ranges given, the results can be highly accurate. Caution should be exercised as to the ranges of applicability of these equations since they do not, as with all polynomials, extrapolate well. If variable values are needed outside the range of the RSEs generated, a new DOE experiment should be created and executed. In addition, the equations are continuous, and thus cannot account for discontinuities or higher order effects.

Prediction Profiles

Once the RSEs are created, JMP can then be used to create prediction profiles. These profiles allow the designer to see graphically how the responses vary with respect to changes in each of the variables. Figure 4 shows a sample prediction profile. The lines in Figure 4 denote the sensitivity of the response with respect to each variable. In essence, they are the partial derivatives of the response with respect to the variable with all other variables set at a given value. A flat or barely sloped line indicates that the variable does not have much impact on that response.

When using the prediction profile tool while in JMP (as opposed to a hard copy printout of the graph), the program allows the designer to change the value of the variables by using a click and drag technique. Using the RSEs, the graph is then updated in real time to show the new values of the responses. In this way the designer can manipulate the equations to gain insight into the problem and also to seek optimal configurations.

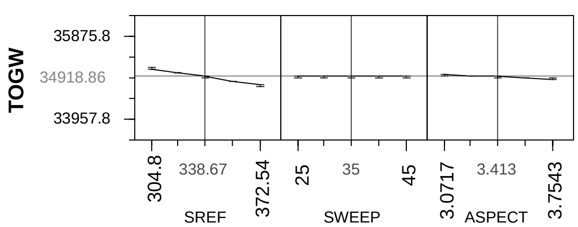

Figure 4- Example of a Prediction Profile

Probabilistic Analysis

K_factors

Once the RSEs have been created, they can be manipulated in a multitude of ways. One common use for the RSEs is to explore the effects of varying the inputs probabilistically. For example, this technique could be used to model a new technology concept. A new technology concept is characterized by ambiguity and uncertainty with regards to its performance, cost, etc. In order to introduce these uncertainties into the model, variability must be added to each input variable. This variability may be modeled by creating a multiplier of a disciplinarian metric and putting a probability distribution around it. By using the RSEs in a Monte Carlo environment, the effect of this variability can be quantified. For example, Figure 5 shows a shape distribution for the multiplier, or K_factor, associated with wing weight. This particular shape distribution would be appropriate for a technology that is expected to give a 7.5% decrease in wing weight, yet recognizes, through the use of a skewed distribution, that there is some chance of achieving either a greater or lesser change in wing weight. Other distribution shapes that may be used include a uniform distribution, used for when each value is as likely as another value, or a normal distribution which is used when there is an equal uncertainty around a particular value. The K_factor concept is not limited to modeling technologies, but can be used whenever the impact of the variability of an input factor is desired.

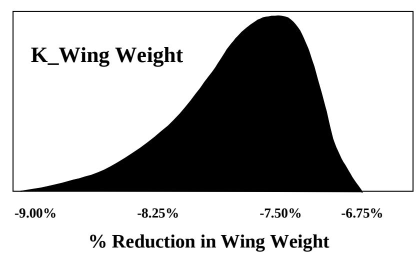

Figure 5- Notional Shape Function for a Wing Weight Reduction K_factor

Monte Carlo Simulation

After determining shape distributions for all of the variables, a Monte Carlo simulation, utilizing the Crystal Ball [14] software, is conducted. Variable values are chosen randomly based on the distribution given. The responses are then calculated through the use of the RSEs. The results are probability distributions that indicate the likeliness of achieving a certain result. Figure 5 shows examples of the two ways that the probabilistic results can be presented. The first is the probability density function (PDF), which depicts the frequency that a certain value is observed in the simulation. The second is the integral of the PDF, called the cumulative distribution function (CDF), which shows the probability or confidence of achieving a certain value. By examining the CDF in Figure 5, the designer can see that there is about a 10% chance of achieving a takeoff gross weight of 33,475 pounds or less, but a 100% chance of achieving a takeoff gross weight of less than 33,850 pounds (find 33,475 on the horizontal axis, follow it up to where it hits the curve, and read the corresponding probability from the vertical axis).

The designer can interpret information from the probability distributions in a number of ways. If the distribution has quite a bit of variability, but some or most of it fulfills the requirement being examined, this would suggest that the assumptions, including any technology infusions, are viable. It would be beneficial, therefore, to invest more resources into the technologies or options that the assumptions represent. This addition of resources could have the effect of narrowing the uncertainty associated with the technologies or options. On the other hand, if the distribution indicates that the probability of meeting the requirement is low, then it might be more provident to examine other options before investing money into a technology or decision that might not be sufficient to solve the problem. This kind of system-level investigation can also show how much the detrimental effects of certain decisions are penalizing the system. This information, shared with the disciplinary experts that engage in the development of the technologies or assumptions, could be investigated to see how resources need to be allocated towards reducing the penalties, as opposed to improving benefits.

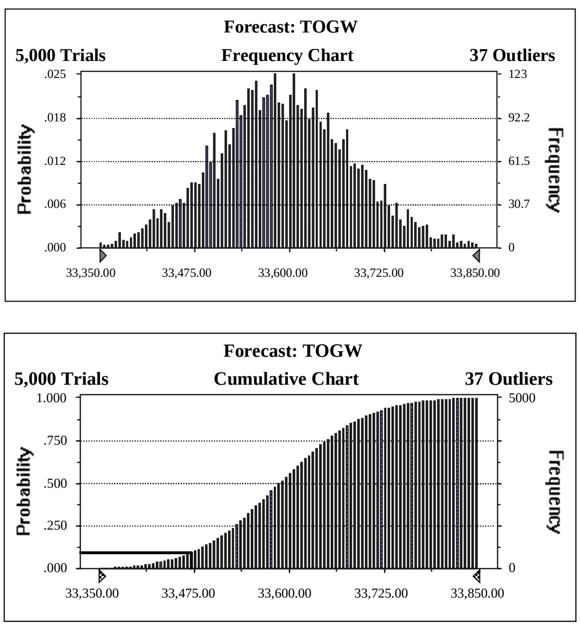

Figure 6- Examples of a Probability Density Function and a Cumulative Probability Function

Project Approach

A project approach to aid in the formulation of the methodology was developed that consisted of the following six steps:

- Select a representative system of system code
- Apply current methods
- Identify issues of concern
- Propose and implement solutions to issues
- Determine new methodology based on solutions to issues
- Apply new methodology to representative example

The first step was to identify an appropriate system of system analysis code for use in developing the methodology. Although this code would be the test bed for the methodology, it was assumed that the new methodology would be applicable to any similar systems code. Next, the existing current probabilistic methods would be applied in order to identify any specific issues of concern that might arise from the extrapolation of the method from the vehicle level to the theater level. Once these issues were identified, solutions to these issues would be proposed. Based on the incorporation of these solutions, a new methodology would be developed. The final step would be to test and justify the new methodology by applying it to a representative test case.

Selecting a Representative System of System Code

A code needed to be selected that had the ability to model theater level interactions. In addition, this code needed input and output variables that were relative, provided insight, and were easily manipulated. Finally, in order to be compatible with the existing probabilistic methodology, the code needed to have the ability to run multiple cases quickly and efficiently. Many theater level codes exist, but all have different emphasis and capabilities, and there are certain codes favored by different organizations (Figure 7).

USN	USA	USMC	USAF	OSD
ITEM GCAM	TAC WAR VIC	COMBAT IV	THUNDER	TAC WAR ITEM

Figure 7- Common Theater Level Codes

The code selected for the study was ITEM (Integrated Theater Engagement Model) developed by SAIC, Inc[14]. ITEM is an interactive, animated computer simulation of military operations in theater-level campaigns. It has fully integrated air, land, and naval (including amphibious) warfare modules and contains a strong emphasis on visualization. The inputs and output are GUI-driven (Graphic User Interface), and an example of this interface is shown in Figure 8. ITEM is fully object-oriented in design and execution and contains a hierarchical structure of its database.

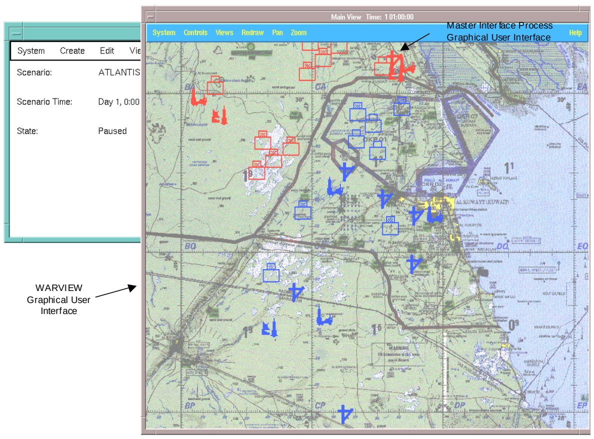

Figure 8- The ITEM Graphical Interface and Environment

Apply Current Methods

The current probabilistic methods as discussed in the previous section were applied using the ITEM code as the analysis code. An outline of this method is presented in Figure 9. During this step, several key issues were identified that were a function of both the analysis code itself as well as conducting the analysis at the new system level.

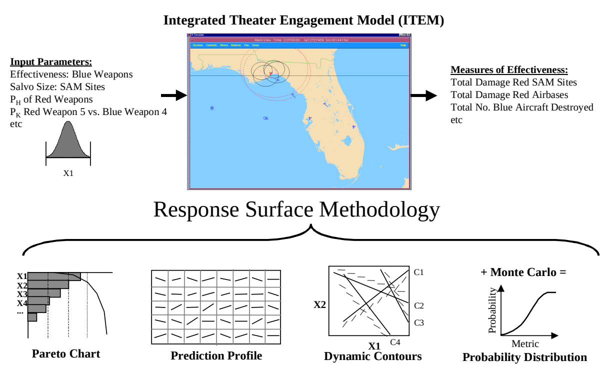

Figure 9- Applying Current Methodology Using ITEM

Identify Issues of Concern and Propose Solutions

It was expected that applying the traditional methodology to the theater level would uncover some problems or issues that would not be a factor at the vehicle level. Investigation and preliminary implementation of the current methodology identified four key issues of concern.

Human in the Loop

A major difference between vehicle level and theater level analysis codes is how the user interacts with the code. In a traditional vehicle sizing code, the user will supply a set of inputs and the code will iterate on a sizing scheme to converge the vehicle according to the laws of physics and empirical relationships. In ITEM and other similar theater codes, however, the user becomes an integral part of the analysis process. This means that the user periodically evaluates the effect of his/her decisions and can then change the parameters (either from that point or change initial input parameters and rerun the simulation) to provide improved results. ITEM was specifically designed to incorporate the use of human judgement to make strategic decisions based on the state of the forces at any given time. Figure 10 shows a typical analysis scheme for using a theater level code.

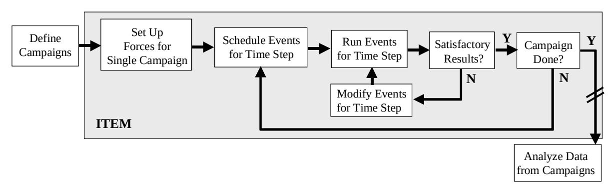

Figure 10- Flowchart for Decision-Making for ITEM[15]

The alternative to having the human in the loop is to use some sort of embedded rules (expert systems) to make decisions. There are some theater level codes that do this. The key drawback to this is that the rules have an inherent lack of flexibility to simulate real operational plans. In addition, these rules lack transparency in assessing cause and effect relationships. An example of this drawback in illustrated in the following example. Say that an embedded rule system is used to model the decisions made in a particular scenario. The results are summarized as follows: "The analysis shows that there is an 85% probability that this scenario (with its inputs) results in the loss of two aircraft carriers in the first four days of the event." What is wrong with this statement? In the real world, losing two aircraft carriers is so completely unacceptable that, after the loss of the first carrier, **the decisions (inputs) would be changed** in order to ensure that a second carrier would not be lost. With embedded rules, unrealistic results such as these could be modeled and decisions based upon these results.

The question now becomes: how do we apply a probabilistic methodology if we need to maintain the information provided by having human in the loop? There are several possible approaches. The first of these is the simplest. Acknowledge the problem, but schedule the events and run the cases anyway. In other words, ignore the issue and continue. The unrealistic solutions and decisions are accepted and identified, while still gaining insight into the overall problem. In addition, care can be taken to try and eliminate unrealistic decisions through careful scheduling. If it is decided that it is important to include the issue, the next logical step is schedule events, create nodes of key decisions, and examine all possible results. The concept of decision trees would be useful here, and can aid in identifying the impact of key decisions.

The final decision of the authors was that the issue was important and crucial enough to address, and critical to the formulation of the new methodology. In addition to using decision trees to identify and model key decision points, the authors will explore the concept of Time-Dependant Response Surface Equations (TDRSEs). In the current methodology, there is a direct input-output relationship between the design variables and the response metrics of interest. A TDRSE would try and model an input variable that changes during the course of the analysis. Instead of the response being a function of a set of variables, the response would be a function of a *vector* of variables. Each vector would represent the set of decisions that could be made at each decision node. Another advantage of this formulation is that probability distributions could be applied to each possible path at each node. In this way, the human decision maker can be modeled. For example, a decision node may have two identified paths. A "practical" decision might have, say, an 80% chance of occurring while a more "aggressive" decision would only be chosen 20% of the time. In this way, one could model the personalities and decision-making abilities of several different types of decision makers, but with the ability to assess them in one modeled environment.

Batch Mode vs. GUI

The current methodology relies on the ability to run the analysis code multiple times with changing inputs. Typically, hundreds of cases need to be run for each scenario. Currently, there exists a batch mode in ITEM, but this batch mode takes as its input a binary file. At this time, the most efficient way found to running multiple cases was to write a script that would create a separate text input file for each case. The user must then load each of these files by hand into the graphical interface and convert the text file into a binary file and rename it. A separate script is then used to run the binary files and parse the output. While the middle step is rather tedious, it is not prohibitive to the method. Ideally, a batch mode with text files would be created. Realize, however, that this issue is strictly a function of the ITEM code. If the methodology were to be applied to other theater level codes, an efficient way of running the multiple cases must be created in order for the method to supply useful and timely data.

Hardwired Measures of Effectiveness

Currently in ITEM, the user can specify which responses are desired as selected from a set of hardwired Measures of Effectiveness (Figure 11). While this set is rather complete and certainly adequate for the formulation phase of the methodology development, the exploration of less traditional metrics would necessitate changes to the source code or close collaboration with SAIC.

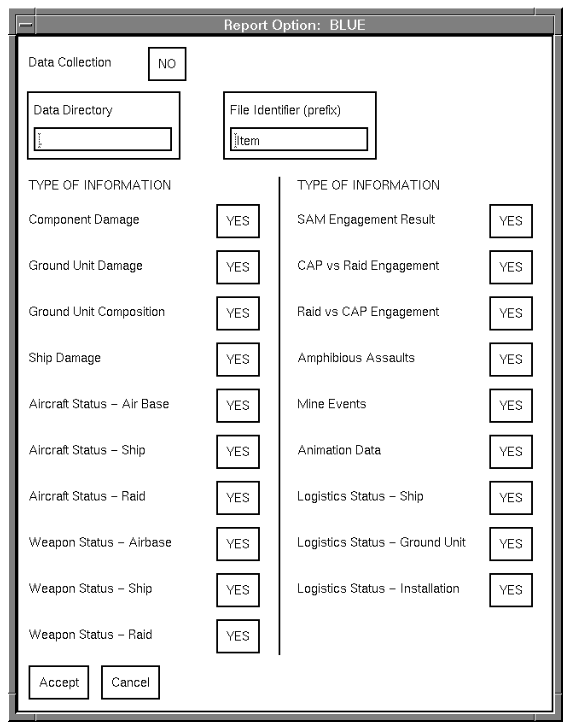

Figure 11- Measures of Effectiveness Available in ITEM

System Differentiation

The final issue touches on the "system of systems" approach discussed earlier and involves how ITEM (and other theater level codes) define their inputs and components. A weapon is defined primarily by its effectiveness value against a target. As shown in Figure 12, weapons can then be grouped into Standard Conventional Loads (SCL) and attached to aircraft. At this point, the aircraft is merely the weapons delivery system, and flies in a straight line path to the target. It is important to note that no system design attributes are defined or used at this level of analysis. In other words, the variables that define "aircraftness" are essentially missing from the theater level. This highlights the disconnect between the analysis levels and illustrates the necessity of having a continuous flow of information from the vehicle level to the theater level, as shown in Figure 13.

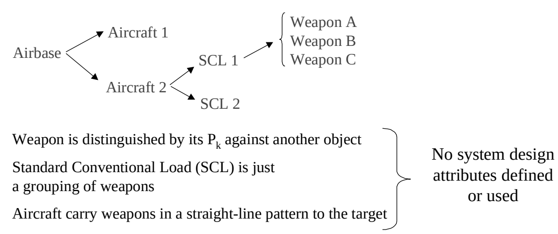

Figure 12- Example of System Differentiation and Component Breakdown

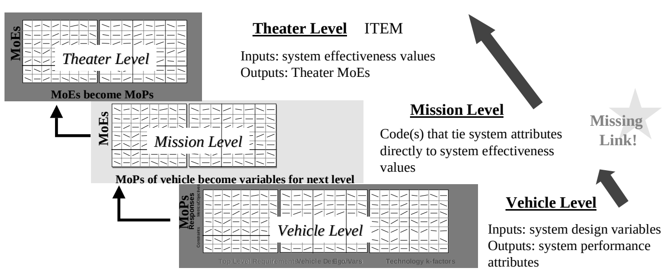

Figure 13- Ideal Continuous Flow of Information from Vehicle Level to Theater Level

Formulation of Methodology

To summarize the progress in the formulation of a new probabilistic methodology to assess system effectiveness: at this point, a representative system of systems code was selected, the current methodology was applied, issues of concern were identified, and proposed solutions to these issues proposed. The remaining steps are to implement these solutions and thereby determine a new methodology based on these solutions, and then apply this new methodology to a representative example. A mission level code that maps the output of the vehicle level to the input of the theater level needs to be selected and incorporated. The remaining steps are in progress, with results forthcoming.

Results and Discussion

The results in this section come from applying the current methodology techniques to the system level code, ITEM. This step was necessary in order to validate that the basic method could indeed be applied to a code that was not a vehicle level sizing code, as well as to help identify any issues raised by the application of the method.

The Scenario and Operational Situation

The scenario, operational situation, and inputs and outputs were provided by Johns Hopkins Applied Physics Laboratory (see Acknowledgements section). The test scenario chosen was a fictional conflict in which South Florida attacks North Florida with missiles and aircraft. An Air Superiority Operational Situation was constructed, and is summarized in Figure 14. Two blue aircraft were modeled, and are called Strike 1 and Strike 2. The variables that are available to model aircraft are shown in Figure 15, and the difference between the two aircraft specified. (It is again important to note how few variables are used to represent an aircraft, and the disconnect between these and traditional design variables becomes obvious.) The red SAM sites are shown in Figure 16, with their main goal to protect the red airbase.

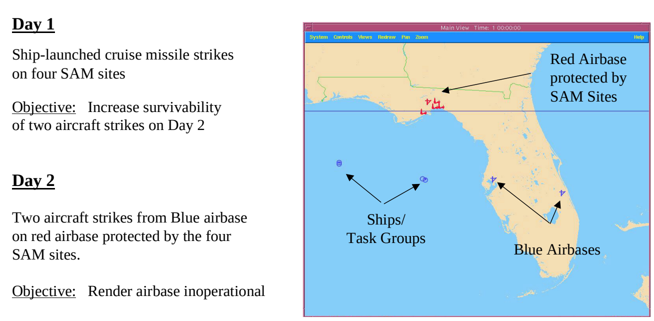

Figure 14- Florida Scenario with Air Superiority Operational Situation

Strike 1

Mission Duration	2.5 hrs
Turnaround Time	5 hrs
Speed	1600 kts
Range	4000 nm
Relative Detectability	0.4
Maneuverability Degradation	Variable
Shelterable	
Maximum Altitude	25000 ft
Standard Conventional Load	L1, L2

Strike 2

Mission Duration	3 hrs
Turnaround Time	6 hrs
Speed	1500 kts
Range	4500 nm
Relative Detectability	0.25
Maneuverability Degradation	Variable
Shelterable	
Maximum Altitude	25000 ft
Standard Conventional Load	L1, L2

Figure 15- Blue Aircraft Modeled for Scenario

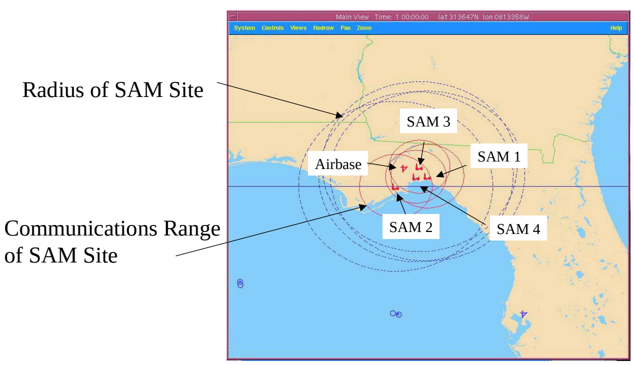

Figure 16- Red SAM Site Detail

Inputs and Outputs

The inputs to the scenario as provided by JHAPL are shown in Table 2. "PK" is the incremental probability of kill and "EK" represents the expected number of targets destroyed per hit. The minimum and maximum values bound the ranges for the variables, with the average of the two values being used as the midpoint, or baseline, value. The output Measures of Effectiveness are given in Table 3. Note that the number (percentage) of aircraft destroyed is the same as 1- (number of aircraft survived).

Table 2 - Inputs to Methodology Using ITEM

	Minimum	Maximum
PK of Cruise Missile versus SAM Fire Control System	0.65	0.95
Maneuverability Degradation of Strike 1 Aircraft	20%	60%
Maneuverability Degradation of Strike 2 Aircraft	40%	80%
EK of Strike 1 AC / SCL 1 Loadout versus Airbase Runway	1	2
EK of Strike 1 AC / SCL 2 Loadout versus Aircraft in Open	2	4
EK of Strike 2 AC / SCL 2 Loadout versus Aircraft in Open	1.5	2.5
PK of SAM-1 Missile versus Strike 1 Aircraft	0.1	0.3
PK of SAM-1 Missile versus Strike 2 Aircraft	0.1	0.4
PK of SAM-2 Missile versus Strike 1 Aircraft	0.1	0.2
PK of SAM-2 Missile versus Strike 2 Aircraft	0.1	0.3
Reliability (i.e. P(hit)) of SAM-1 Missile	0.55	0.85
Reliability (i.e. P(hit)) of SAM-2 Missile	0.55	0.85

Table 3- Outputs from Methodology Using ITEM

Measures of Effectiveness
Expected Number of Strike 1 Aircraft that are Destroyed
Expected Number of Strike 1 Aircraft that Survive
Expected Number of Strike 2 Aircraft that are Destroyed
Expected Number of Strike 2 Aircraft that Survive
Expected Number of SAM-2 Missiles Launched
Expected Number of Runways Destroyed by Aircraft Strikes
Expected Number of Aircraft in Open Destroyed by Aircraft Strikes

Preliminary Results

Preliminary results are shown in the form of a screening test. This is a first order fit between the input variables and the output variables, and a total of 129 cases were run. Pareto plots showing the magnitude of the contributions of the different input variables to the variability of the response is shown in Figure 17. For the expected number of aircraft destroyed for both Strike 1 and 2, it can be seen that the Pk of the cruise missile against the SAM fire control system is the highest variable contributor. This confirms the Day 1 objective: increasing the Pk value of the cruise missiles does indeed increase the survivability of the strike aircraft. (It is important to remember that the Pareto plot shows the greatest contributors to the *variability* of the response. In Figure 17, the Pk of the cruise missile is shown to contribute the highest amount to the expected number of aircraft destroyed. This is not to say that an *increase* in the Pk value *increases* the number of aircraft destroyed. In fact, the opposite is true, as shown in Figure 18.) Likewise, the Pk of the cruise missile has an overwhelming effect on the number of SAM2's launched against the strike aircraft. SAM1 was represented in the model, but because of the geometry involved, was in effect a non-player. The variable was kept in as a sort of sanity check; if SAM1 was to show an effect, the model would have been unrealistic, identifying an error somewhere in the model. Similarly, some of the results are obvious: the Pk of Strike 2 aircraft against aircraft in the open does indeed have the highest influence on the expected number of aircraft in the open destroyed. Again, this shows the model and system are behaving as expected.

Figure 18 is the prediction profile for the screening test. While the Pareto plot identifies the chief contributors to the response, the prediction profile identifies magnitude and direction of the impact of the input variables, as well as shows the simultaneous impact of the other variables. As noted in the Pareto plots, the prediction profile shows that increasing the Pk of the cruise missile does indeed increase the survivability of both strike aircraft. In addition, the number of SAM2's launched is decreased. There is little change in the number of airbase runways destroyed or the number of aircraft in the open destroyed, which makes intuitive sense. A correlation is also seen between the degradation of the maneuverability of the strike aircraft and the number of these aircraft destroyed. Other results are checked for intuitive correctness.

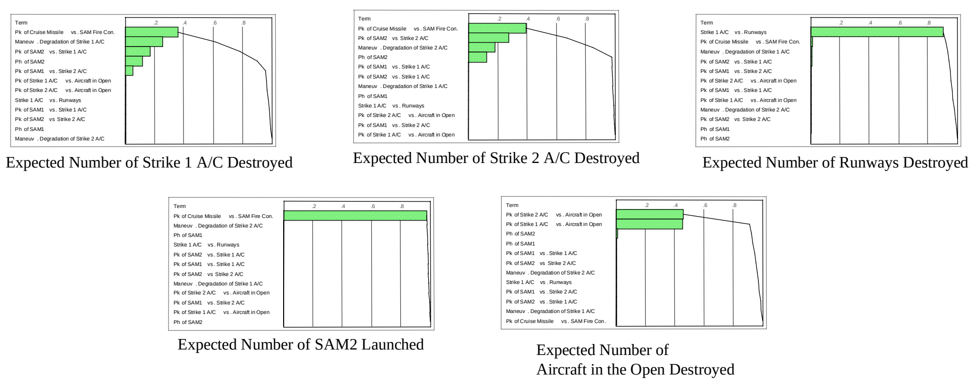

Figure 17- Pareto Plots for ITEM Screening Test

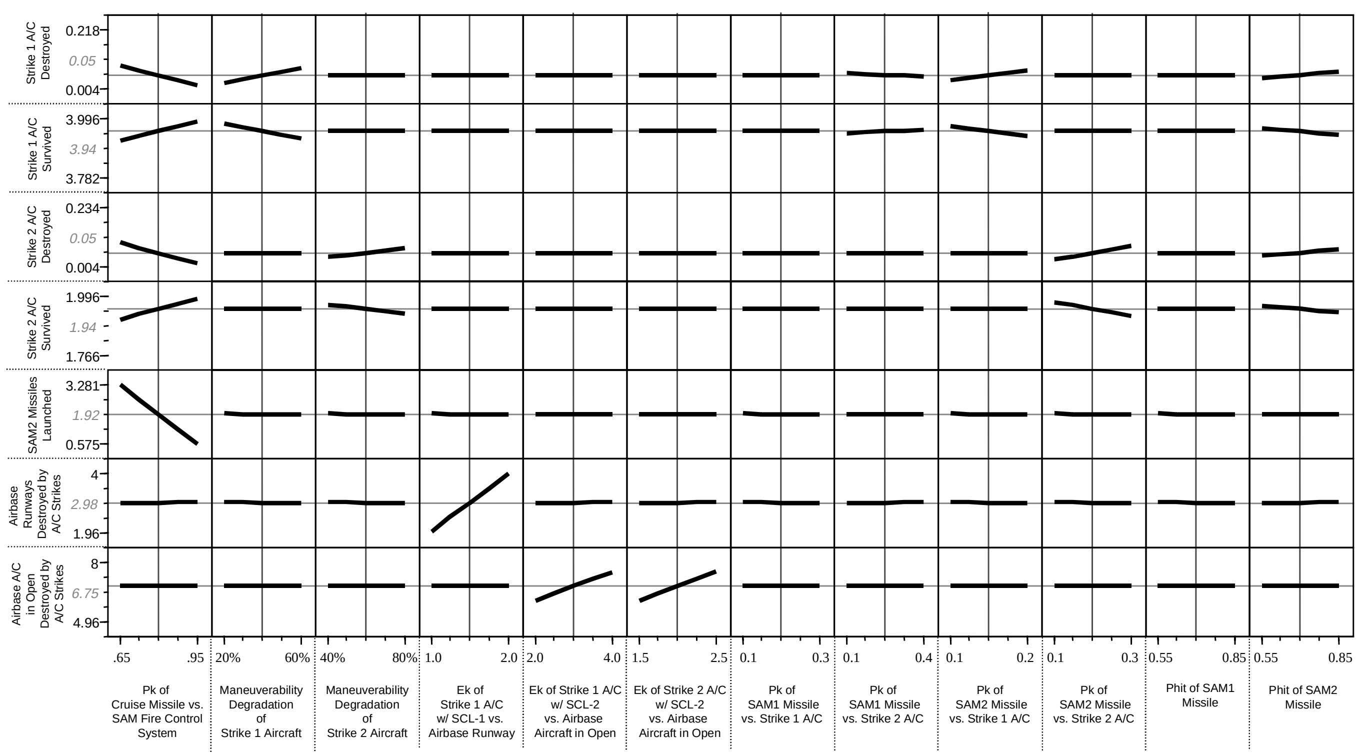

Figure 18- Prediction Profile for ITEM Screening Test

While the inputs and outputs to the screening test were rather simplistic, the analysis did serve as a proof of concept. It was shown that the current probabilistic method could be applied to a theater level code. Intuitive trends were confirmed, and issues were identified that need to be considered and solved in order to make an extrapolated probabilistic methodology a useful and beneficial tool.

Future Work

Now that the probabilistic methodology has been shown to work in principle, the key issues that differentiate a vehicle level code from a theater level code can be addressed in earnest, with the result being a cohesive new probabilistic methodology to assess system effectiveness. A formulation for the incorporation of Time Dependent Response Surface Equations (TDRSEs) needs to be actualized, and a middle level code, relating vehicle level outputs to theater level inputs, needs to be identified and incorporated. Finally, the new methodology will be be used to conduct a representative system analysis.

Acknowledgements

This work was supported by the Office of Naval Research under contract number N00014-00-10132, and under the direction of Katherine Drew, Program Manager of the ONR Affordability Program. Scenario support, as well as inputs and outputs, were provided by Johns Hopkins Applied Physics Laboratory.

[1] Blanchard, Benjamin S., Logisitics Engineering and Management, 5th Edition, Prentice Hall Inc., 1998.

[2] Tillman, F.A., Hwang, C.L., and Kuo, Way, "System Effectiveness Models: An Annotated Bibliography", IEEE Transactions of Reliability, Vol. R-29, No. 4, October 1980, pp 295-304.

[3] ARINC Res. Corp., Reliability Engineering. Englewood Cliffs, NJ: Prentice Hall 1964.

[4] *Definitions of Effectiveness Terms for Reliability, Maintainability, Human Factors, and Safety*, MIL-STD-721B, US Department of Defense. 1966 Aug.

[5] Tillman, F.A., Hwang, C.L., and Kuo, Way, "System Effectiveness Models: An Annotated Bibliography", IEEE Transactions of Reliability, Vol. R-29, No. 4, October 1980, pp 295-304.

[6] Mavris, D.N., Baker, A.P., Schrage, D.P., "Simultaneous Assessment of Requirements and Technologies in Rotorcraft Design", Proceedings of the 56th Annual Forum of the American Helicopter Society, Virginia Beach, VA, May 2-4, 2000.

[7] Zink, P. S., D. A. DeLaurentis, M. A. Hale, V. V. Volovoi, D. P. Schrage, J. I. Craig, R. E. Fulton, F. Mistree, D. N. Mavris, W. Chen, P. J. Röhl, K. E. Lewis, P. N. Koch, C. E.S. Cesnik, T. W. Simpson, "New Approaches to High Speed Civil Transport Multidisciplinary Design and Optimization," 2000 IEEE Aerospace Conference, March 18-25, 2000, Big Sky, Montana. IEEE-351-00.

[8]Mavris, D.N., Baker, A.P., Schrage, D.P., "Technology Infusion and Resource Allocation for a Civil Tiltrotor," Proceedings of the AHS Vertical Lift Aircraft Design Conference, San Francisco, CA, January 19-21, 2000.

[9] Mavris, D.N., Soban, D.S., and Largent, M.C., "An Application of a Technology Impact Forecasting (TIF) Method to an Uninhabited Combat Aerial Vehicle", 1999 AIAA/SAE World Aviation Congress, San Francisco, CA, Oct 19-21, 1999. SAE/AIAA 1999-01-5633.

[10] Mavris, D.N., Soban, D.S., and Largent, M.C., "An Application of a Technology Impact Forecasting (TIF) Method to an Uninhabited Combat Aerial Vehicle", 1999 AIAA/SAE World Aviation Congress, San Francisco, CA, Oct 19-21, 1999. SAE/AIAA 1999-01-5633

[11] Box, G.E.P., Draper, N.R., Empirical Model Building and Response Surfaces, John Wiley & Sons, New York, 1987.

[12]Montegomery, D.C., Design and Analysis of Experiments, 3rd Ed., John Wiley & Sons, New York, 1991.

[13] SAS Institute Inc,. *JMP, Computer Program and Users Manual*, Cary, NC, 1994

[14]"Integrated Theater Engagement Model (ITEM) Technical Manual", Version 8.3, Science Applications International Corporation, 10260 Campus Point DriveSan Diego, California 92121 November 5, 1999

[15]Modified from diagram in Reference 14.