19 ISMOR

Improving Confidence in the Assessment of System Performance in Differing Scenarios

T D Clayton, Cardinal Consultants www.cardinal.inuk.com

1. Introduction

1.1 In assessing the effectiveness of a military system, such as a weapon or a communications system, the military context or 'scenario' within which the system is to be utilised is often of crucial importance in determining the results and conclusions of the study. Therefore, in order to be of value, a system effectiveness study must address scenario-dependent factors in a clear, direct and comprehensive manner. Without such studies, we increase the risks of developing and procuring systems and equipment which are non-optimal, and we risk having to deploy systems into situations where their effectiveness is poor or poorly understood. It is not satisfactory to run a combat model for two or three broadly similar scenarios and consider that an adequate means of addressing scenario effects.

1.2 Unfortunately, in most system effectiveness studies, accounting for the effects of different scenarios is one of the most difficult aspects of the work. Almost inevitably, combat models and simulations (and even wargames) are built with only one scenario (or a very limited range of scenarios) in mind and it is not feasible to build a series of adequately detailed models to represent a large range of widely differing scenarios.

1.3 This paper reviews some of the problems faced by an analyst in this area, and discusses some of the techniques which can help to address these problems in system effectiveness studies.

2. Problem Area 1 - Scenario Dependency of Input Data

2.1 Most combat models rely on large volumes of data to describe the performance of individual systems, such as the effectiveness of a weapon system. Unfortunately, many of those input data items themselves are (or should be) dependent on the scenario being considered.

2.2 For example, the performance of a sensor-aided munition might be significantly degraded by poor met-vis conditions or by ground conditions around the target which offer high false-target densities. Equipment reliability varies considerable between scenarios, e.g. diesel engines in sandy conditions and electronic equipment in wet conditions. Even data defining effects as fundamental as the lethality of a given warhead against a given target type can be significantly dependent on factors such as ground conditions and local features.

2.3 One of the problems is that different groups of analysts are generating results for different purposes. For example, those analysts responsible for conducting detailed terminal effectiveness studies are normally concerned primarily with comparing and optimising warhead designs or comparing and optimising vehicle armour and protection. They are normally less concerned with providing data for higher level models. Therefore, most terminal effectiveness models utilise a clean target description - i.e. the target vehicle only, placed on perfectly flat smooth ground. It would increase study costs by an order of magnitude to generate results for a range of different local conditions, such as rough ground, earthworks around the target vehicle, or adjacent trees or buildings. But it is really these more complex ('realistic') conditions which the combat modeller is attempting to study.

2.4 In addition, the Measures of Effectiveness used in low level studies are not always ideal for use in higher level studies. Again consider detailed terminal effectiveness studies which often use an overall lethality value (such as an MFK value) to represent the warhead performance. This might be satisfactory when simply comparing warhead designs, but the combat modeller is more interested in data which are clearly defined probabilities of specific levels of degradation of the target vehicle - e.g. the probability of complete disabling of the target's firepower capability.

2.5 This is a classic multi-disciplinary problem. To what extent is it the combat modeller's responsibility to ensure he understands the real meaning of the data provided to him, such as kill probabilities, and the assumptions inherent in their derivation, and to what extent is it the responsibility of the sub-system modeller to understand the requirements of the higher level models and to provide appropriate data and information on assumptions? In practice, it is often difficult for a specialist in one area to have a sufficiently detailed understanding of analysis models and techniques in other areas to ensure that all factors are understood and accounted for appropriately when results are used as data in other models.

2.6 The approach which has been used at CCL for some time is to appoint one or two senior and experienced analysts to manage the interface between the different groups. Their remit is to ensure they have a good understanding of analysis techniques and models at all levels of detail, from component and subsystem level to high level combat models and wargames, but in particular to understand the inherent assumptions and approximations and the associated scenario dependencies. Key responsibilities include:

- organising training schemes to help keep all analysts up to date with techniques in all areas,
- involvement in the design of models and analysis tools, especially in terms of inherent assumptions and data requirements,
- involvement in the planning of assessment studies, especially in terms of scenarios, assumptions, data sources and Measures of Effectiveness,
- acting as the conduit of data between groups, with responsibility for assessing the appropriateness of the data for the planned studies,
- involvement in the reporting of studies, particularly to ensure that underlying assumptions and scenarios are clearly reported alongside the results,
- maintenance of databases of system performance results, together with information on underlying assumptions and scenarios,
- maintenance of databases of 'scenario compensation factors' - i.e. data which can be used to correct for changes of scenario, where there is insufficient time to undertake a full study of a system under different conditions,
- development of models and analysis tools to support the above tasks, especially tools which assist in determining the sensitivity of results to underlying assumptions and estimating scenario dependencies.

2.7 Figure 2.7a illustrates the framework within which study results are stored, Scenario Compensation Factors are assessed, and results are made available as data for further studies. Figure 2.7b illustrates the framework within which the continued applicability of Scenario Compensation Factors is checked and adjusted.

Figure 2.7a - Data Interfacing Framework

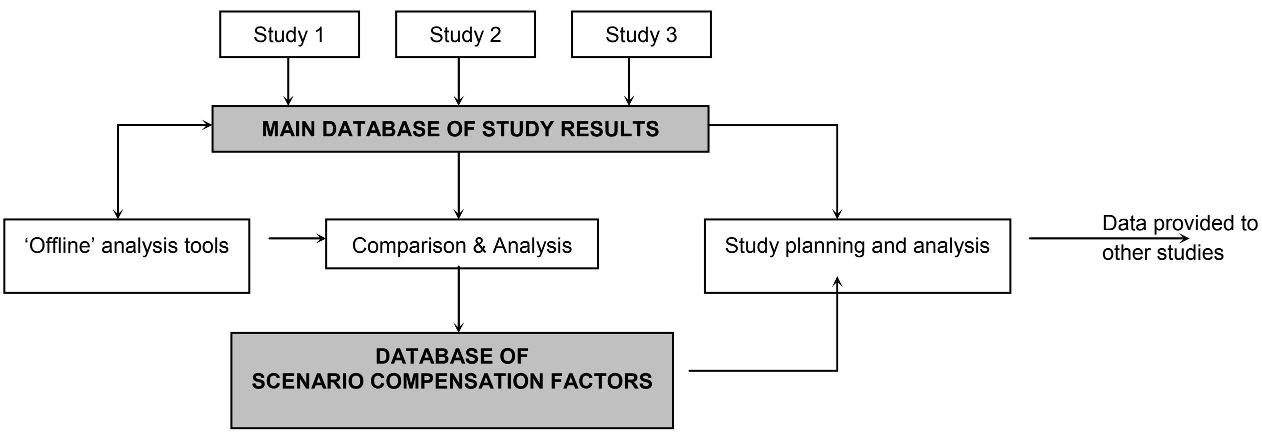
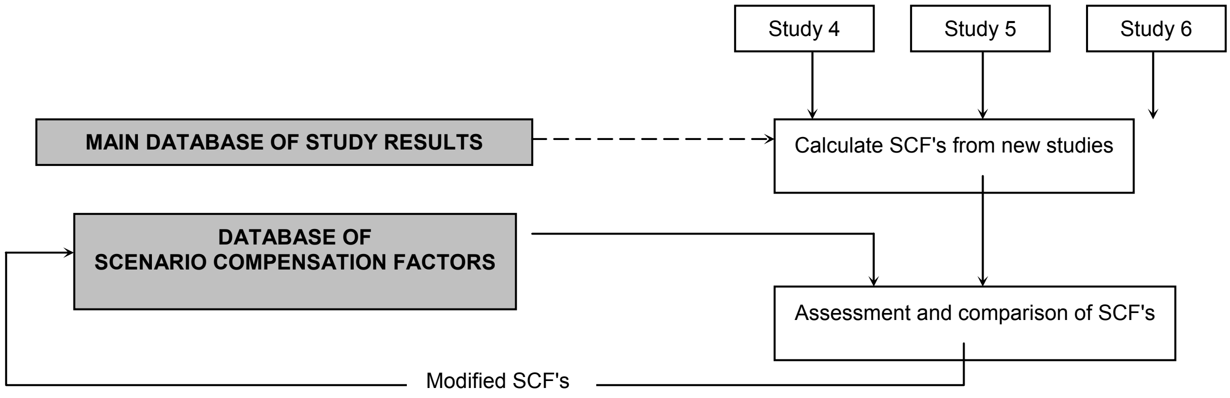


Figure 2.7b - Feedback Loop for Checking SCF's



3. Problem Area 2 - Selecting Scenarios to Assess

3.1 Analysts investigating system effectiveness are often doing so 5 or 10 or more years ahead of the in-service date of the proposed system. Good prediction of the actual scenarios into which we will have to deploy our forces in 5 to 10 years is not realistic. Predicting some of the scenarios might be feasible, but predicting a high proportion of them must be regarded as impossible. Consider some of the major campaigns in the last 20 years: the Falklands, the Gulf and Afghanistan. Those are three very different scenarios, none of which could have realistically been predicted 5 years before they took place, certainly not with enough confidence to base military procurement decisions on those predictions.

3.2 Of course, there is far more work being undertaken at present on predicting potential scenarios than there was 20 years ago. And of course this work is worthwhile. However, analysts who are studying system effectiveness should (in the view of this author) address wider scenarios than the ones we know about now.

3.3 One approach we might consider is to use 'generic' scenarios. That is, rather than considering a very specific detailed scenario with a specific threat in a specific part of the world, consider generic scenarios defined by possible values of a small number (say 25 to 50) of key parameters. Such key parameters would include terrain type, climate, geography and politics, opposing force strength and tactics, opposing force capability and technology, and so on. Table 1 (page 10) shows a simplified framework for generic scenarios using 25+1 key parameters. But if all parameters are treated as occurring independently, we have a huge number of possible generic scenarios; in a 25+1 parameter framework, assuming 3 possible values for each parameter, the number of possible scenarios is 3^{26} - i.e. approximately 10^{12}. Even if we could determine which combinations represented the most likely 1% of these, we still cannot study 10^{10} different scenarios individually.

3.4 A potential solution to this problem is to "decouple" the scenario parameters. That is, make the assumption that the effect of a given scenario parameter on the performance of the systems under investigation can be assessed independently of the value of the other scenario parameters. Then we can conduct 26 studies instead of billions.

3.5 As with any assumption, this approach must be clearly understood, and its validity assessed in relation to each individual study. In some cases, it may not be possible to decouple the parameters entirely; it may be that the effects of small group (2 or 3) of parameters are likely to be significantly interdependent. But the decoupling approach can still be used, simply by defining this group of parameters as a new single parameter (a "meta-parameter") which may take a larger number of separate values or states.

3.6 An example of a study in which this approach has been used is presented in section 5 below.

4. Problem Area 3 - Modelling the Effects of Widely Differing Scenarios

4.1 The majority of the combat models and wargames we have used in recent times are designed to represent only one type of scenario - e.g. a major ground-air campaign against a highly capable threat in an environment typical of central Europe. Models and wargames like that are perhaps not very useful in studying campaigns such as the Falklands or Afghanistan. And there are huge problems in trying to adapt models to scenarios for which they were not originally designed, mainly because most models were never designed to investigate widely differing scenario-dependent factors. And that is partly because scenario-dependent effects are the most difficult to represent adequately. Some simple examples of this are as follows.

4.2 Terrain is a classic example. One of the primary effects of differing terrain types is on lines of sight, which in turn is critical in many situations. For this reason, many models include detailed line-of-sight calculations, but this does not mean that they adequately address differing line-of-sight effects. For example, it is not appropriate to place units arbitrarily on some digital terrain and then calculate lines-of-sight. In reality, a military commander would position his units (as far as possible) to ensure that they have adequate lines of sight within their potential engagement envelopes. So, when a model is calculating lines-of-sight, it is also making a number of additional assumptions about the flexibility the units have to move to alternative vantage points. This is the real scenario-dependent factor and it is essential to analyse and understand exactly what assumptions are being made.

4.3 It is also very difficult to model the effects of widely differing environment and ground type on force mobility. Of course, detailed mobility models can investigate the effect of differing ground conditions on the mobility of a specific type of vehicle, but that is only a small part of what we would need to do in a combat model. We should really represent all the decisions which military commanders could make when a unit is to move. A model which makes implicit assumptions that mobility is about vehicles moving along roads would not be useful when studying scenarios in which there are no roads, or when the best decision is to use helicopters to transport jeeps around.

4.4 A third example is the effect of our own knowledge of the opposing forces. A campaign in which we have high confidence that we know, with good accuracy the strength, capability and even locations of the opposing force is very different from one in which we know little about the enemy numbers, weapons and deployments. Representing both situations within a single combat model is not likely to be feasible.

4.5 The point is that combat models do not really model the decisions which military commanders make during combat operations. The designer of the model effectively made most of those decisions when designing the model. That is, for the specific type of scenario being assumed by the model, it is also assumed that the commanders on the ground would make specific types of decisions - e.g. which group of tanks to attack with this weapon or which road to move along with this unit. Models are therefore not good at decisions such as "get the Americans to airlift a force into a position behind this high-risk area."

4.6 The point is <u>not</u> that it is difficult to model any one situation, but that it is very difficult to build a model with sufficient flexibility (and avoidance of implicit assumptions) to model widely differing scenario types.

4.7 One approach is therefore to use models (or, at least, computer software) to undertake the parts of the task which they are good at (i.e. doing the calculations and concatenating probability distributions), and leave the other parts of the task (such as assessing potential military decisions) to the human.

4.8 An example of a study using this approach is presented in section 5 below.

5. Example Study - Comparison of Two Potential Cannon Systems

5.1 This section summarises how the principles discussed above were recently employed in a study conducted by Cardinal Consultants. Certain details and data have been changed for this Unclassified paper.

5.2 The objective of the study was to compare the Operational Effectiveness of two potential candidates for a cannon system to be retro fitted to light armoured vehicles. The results of this comparison were to be a major input to the procurement decision.

5.3 The following data were provided to define the two systems:

	System A	System B
Weight of gun	High	Low
No. of rounds carried on vehicle	Low	High
Time to open fire	Medium	Low
Rate of fire	Low	High
Fire on the move	No	Yes
Max effective range	6 km	4 km
Accuracy	2 mil	3 mil
Dispersion - 1 km	2.0 m sd	4.0 m sd
Dispersion - 2 km	5.0 m sd	10.0 m sd
Dispersion - 5 km	12.0 m sd	(n/a - out of range)
Round penetration capability	High	Medium
Single round Pk - Truck	0.06	0.04
Single round Pk - LAV	0.04	0.025
Single round Pk - Tank	0.015	0
Affordable number of systems	50	100
Affordable number of rounds	50,000	250,000

5.4 The initial study plan was in two parts:

a. Build a simple simulation of an engagement to determine the numbers of rounds required to kill a target, as a function of engagement range, taking aiming errors and dispersion into account.

b. Use an existing force-on-force model to study effectiveness in two or three different scenarios.

5.5 The actual study proceeded as follows:

5.5.1 <u>Stage 1 - Data Review</u>

On reviewing the data provided, a series of problems were identified, particularly relating to the scenario dependency of the data. For example, the following problems were identified concerning the 'Single Round Pk' values.

a. For cannon systems, there is a significant probability that more than one round will strike the target and for some target types, independent (binomial) combination of SSKP's is inappropriate and could significantly favour the larger calibre system.

b. The data provided were Whittaker weighted by azimuth (i.e. weighted over frontal arc of the target). This is unlikely to be appropriate in all scenarios and assumption of a frontal attack on the target would favour the larger calibre system.

c. No information on range dependency of penetration capability was provided, nor was any information provided about the assumed range or the assumed impact velocity of the rounds. (It is therefore not clear which system is being favoured).

d. Data for three target types were provided, but in some scenarios it is likely that other target types would be realistic and feasible targets for the cannon system. Such additional target types might include helicopters, personnel, radar and communications installations, and buildings. It was felt that omission of these target types could potential favour the larger calibre system.

In view of these concerns, a more comprehensive terminal effectiveness study was conducted, in concert with use of the engagement model, to determine the lethality of a burst of rounds against a wider range of target types, and tabulating results as functions of target azimuth and engagement range.

5.5.2 Stage 2 - Review of model applicability

Secondly, a review was undertaken of the applicability of the existing force-on-force model and the applicability of the three proposed scenarios. A series of potential problems were identified, including the following.

a. The model would always choose to engage the available (within range) target vehicles with the highest military worth. Thus System B would be modelled as engaging tanks, even when it has a very low effectiveness against that target. (This favours System A).

b. The model implicitly assumed that Blue units would fire only at land vehicles and not at other targets such as helicopters and personnel. (This favours System A).

c. The three proposed scenarios varied little in terrain type, tending to give fairly long engagement ranges. (This favours System A).

d. No variations in met-vis were proposed, tending to give fairly long engagement ranges. (This favours System A).

e. No night conditions were proposed, tending to give fairly long engagement ranges. (This favours System A).

f. It was proposed that the same positions of our own units should be used in the model runs for both systems. This fails to account for the fact that commanders would, where possible, position shorter range systems closer to the likely locations of targets. (This favours System A).

g. The model implicitly assumed that Blue units are static when firing; this fails to give credit to a system which can fire on the move. (This favours System A).

h. Blue vehicles were assumed always to move along roads or firm terrain. This fails to give credit to the lighter and more manoeuvrable vehicles equipped with the lighter cannon system. (This favours System A).

i. The proposed model runs would use the same number of Blue units, whether equipped with System A or System B. This fails to give credit to the fact that more units can be deployed with the cheaper system, giving the military commanders more flexibility. (This favours System A).

In view of the above difficulties in achieving a fair comparison between the two systems, it was agreed that a different approach should used.

5.5.3 Stage 3 - Development of a Study-Specific Assessment Tool

A software tool was developed specifically for this study whose functions were:

a. to input, store and display data defining the key weapon parameters and data defining a series of probability distributions to represent scenario-dependent effects,

b. to carry out the calculation of a defined series of Measures of Effectiveness, using the input data and probability distributions,

c. to maximise the visibility of the data and assumptions and to facilitate the direct comparison of the two systems.

The 'CST' tool is described briefly below.

5.5.4 Stage 4 - Conduct of the Study

The study consisted of a series of sub-studies, or assessments, each of which was designed to provide the data and distributions for the CST tool. Each such assessment used a variety of methods for deriving appropriate data, including use of other models, hand calculations, and Military Judgement. The Scenario Parameters framework (see Table 1) was used as a checklist at each stage to attempt to ensure that all relevant scenario-dependent factors were addressed.

Each assessment was fully documented, and the results subject to scrutiny prior to undertaking the final runs of the CST tool.

6 The CST Tool

6.1 The tool was developed in Visual Basic for speed of development and validation, and given that the numerical processing requirements were not high. Key features are as follows:

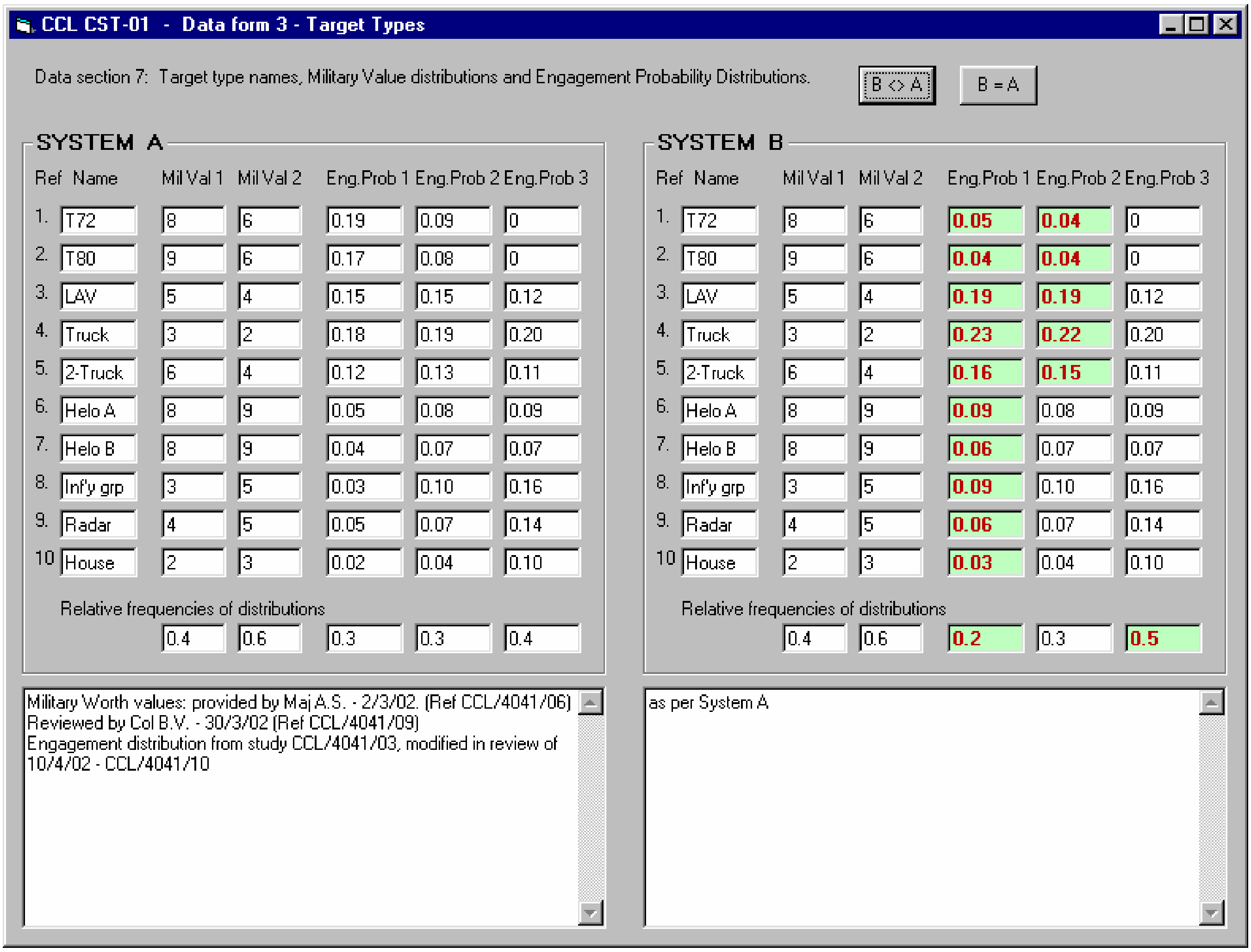
1. The tool is specifically designed to conduct a comparison of two systems. In order to maximise the clarity of data differences between the two systems, data for both systems appear on each data screen, and differences in values are highlighted. (see Fig. 6.1)

2. Stored with the data are notes and references detailing the individual assessments which have provided the data for this study. (see Fig. 6.1) This is regarded as an important aspect of ensuring that data and assumptions are clear and open to scrutiny. Indeed, the results of the scrutiny process itself can (should) be recorded.

3. All data are stored automatically when results are generated, giving a full audit trail throughout a study.

4. All data are subject to automatic validity checks before results are calculated.

<u>Figure 6.1 - Example of a Data Handling Screen</u>



6.2 The tool calculates a series of Measures of Effectiveness, using Monte Carlo methods, from the following input data:

a. Target type engagement distributions,

b. Military worth values for each target type,

c. Kill probabilities for a burst of rounds for each target type as a function of azimuth and range,

d. Engagement range distributions as a function of target type,

6.3 The Measures of Effectiveness include:

a. Military Worth of kills per burst,

b. Military Worth of kills per vehicle ammunition load,

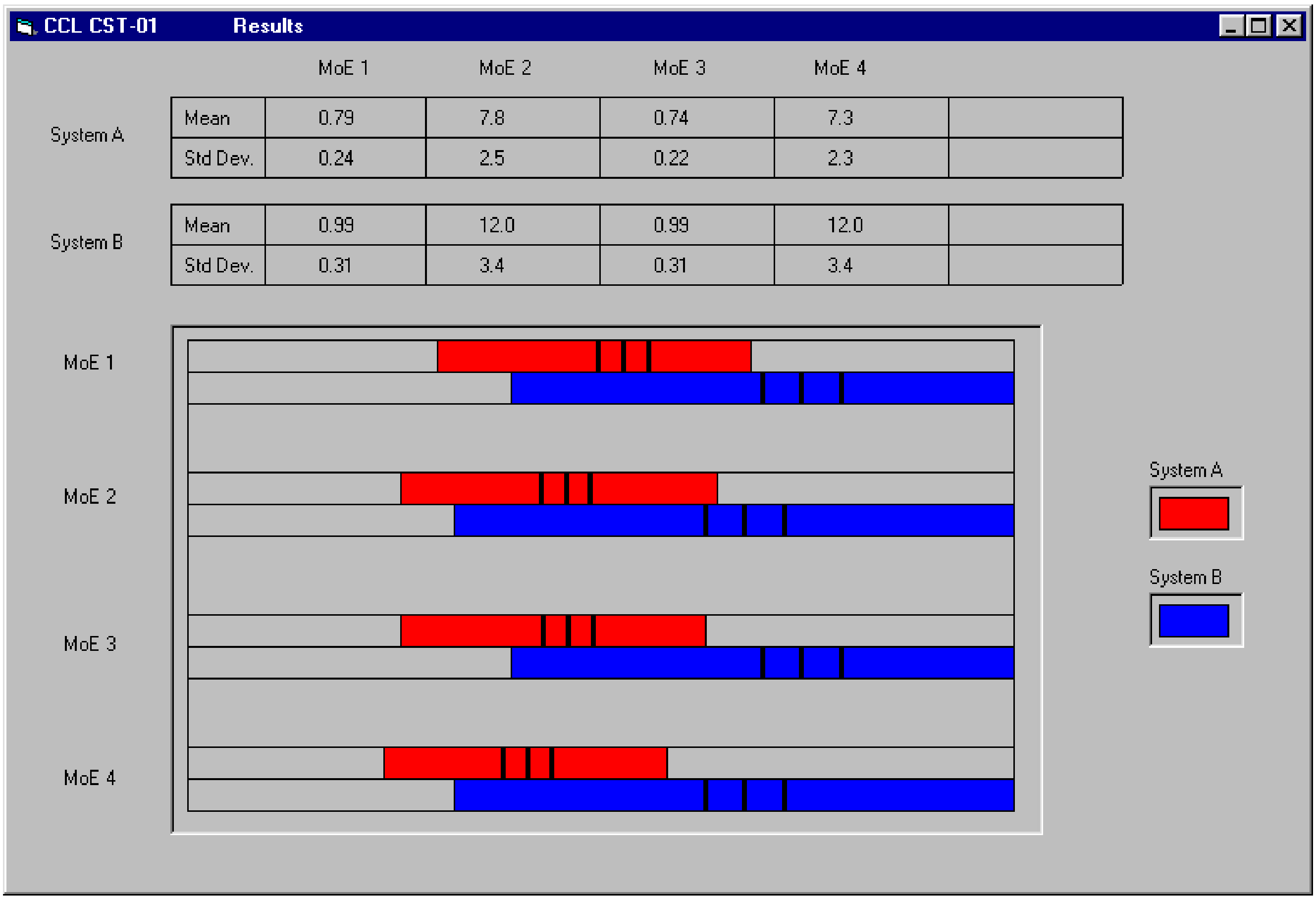
6.4 In addition to the numerical calculations, the tool allows each MoE to be modified by input adjustment factors representing the Military Judgement of some aspects not represented by the data. In the study, Military Judgement was used to place a value on:

a. The ability of System B to fire on the move,

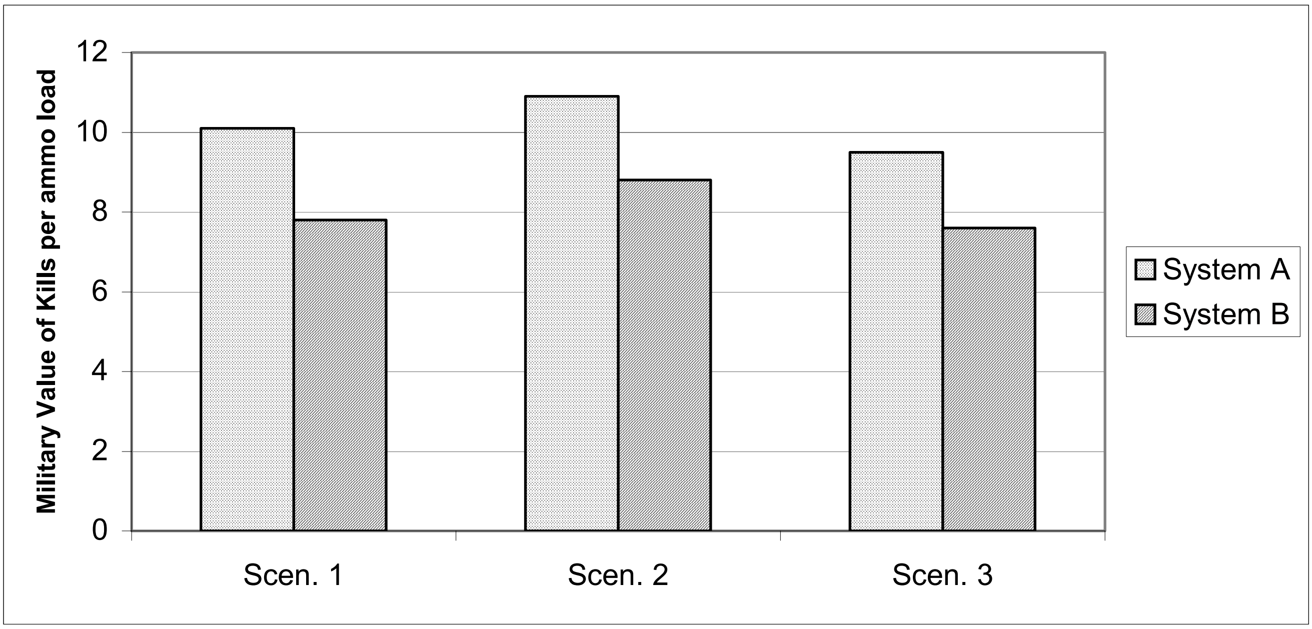
b. The increased manoeuvrability of the lighter system.

7. Results of the Example Study

7.1 The results of the sample study are illustrated in Figure 7.1, and show that System B is preferred, based on these Measures of Effectiveness.



7.2 The results contrast sharply with the results from running the originally planned model, which was expected to favour System A.



8. Summary and Conclusions

8.1 Assessment of the performance of military systems can be highly dependent on the scenarios being assumed. Since the actual scenarios into which systems will be deployed in 5 to 10 years cannot be predicted reliably, it is important to address scenario dependent parameters in a comprehensive manner.

8.2 Examples have been described of methods of ensuring that scenario dependencies are considered correctly, including:

- establishment of specific responsibilities for understanding and assessing scenario dependency of data transferred between modelling levels, and use of supporting knowledge databases,
- use of checklists of scenario parameters to assist in addressing a wide a range of factors as possible,
- detailed analysis of methods and assumptions used in combat models to identify potential sources of bias in the results, especially when comparing two similar systems,
- conduct of subsidiary studies and assessments to generate intermediate results which can be subject to scrutiny and validation,
- use of simple tools to combine intermediate results into overall Measures of Effectiveness whilst maximising the visibility of the underlying data and assumptions,
- use of Military Judgement, combined with use of intermediate results, to address factors which are not feasible to assess using modelling methods.

8.3 It is suggested that the combined use of the above methods can, in many cases, generate more reliable assessments of system performance than would be generated using existing combat models.

Climate	
Temperature	Precipitation
Hot and humid	Frequent heavy
Hot and dry	Infrequent heavy
Temperate	Moderate
Cold	Low
Severely cold	None

Ground		
Vegetation	Topology	Roads
Jungle	Mountains	Good
Swamp	Hilly	Medium
Temp. woodland	Mixed	Poor/none
Scrub	Flat	
Grass (open)		
Mixed		
None - desert		
None - ice		
Urban / suburban		

Geography and Politics		
Geog. isolation	Neighb. countries	Local civilian popul'n
Very isolated	Hostile	Hostile (high danger)
Medium	Neutral	Hostile (low danger)
Easily accessible	Mixed	Neutral
	Friendly	Friendly
	Actively supportive	Actively supportive

Opposing Ground Personnel	
Numbers	Predominant Types
High	Professional
Medium	Fanatic
Low	Mixed
	Amateur - volunteers
	Amateur - conscripts

Opposing max. capability
Max capability
L.R. Nuclear
S.R. Nuclear
L.R. ChemBio
S.R. ChemBio
L.R. Conventional - Heavy
S.R. Conventional - Heavy
Light only

Opposing force ground equipment		
Technology	Numbers	Own Intell.
Mostly high tech.	High	Very good
Limited high tech.	Medium	Fair
Medium	Low	Little
Low		

Opposing force posture/deployment	
Posture	Detectability
Fully defensive - well prepared positions	Unit positions not well known and difficult to detect
Fully defensive - partially prepared positions	Some units known, some detectable
Mainly defensive with some forays	Well known positions and easily detectable
Mainly attacking	

Opposing force Air Capability (unsurpressed)			
Range of Aircraft types	Level of Technology	Numbers	Own Intell.
High	High	High	Good
Medium	Medium	Medium	Fair
Low	Low	Low	Poor
None			

Opposing force anti-air capability (unsurpressed)		
Number of units	Capability	Own Intell.
High	High	High
Medium	Medium	Medium
Low	Low	Low

Maritime	
Involvement	Opposing capability
Strategic	High
Tactical	Low
None	

BLUE ROLE
Nature of own role in campaign
Peace keeping / enforcing
Combat role - defensive
Combat role - hunt and kill

Table 1

Example of Simplified Framework for Generic Scenarios ("25+1" parameters)