



Unifying Capability Integration Analysis

Initial Insights

Leonard F. Kerzner
DRDC CORA

DRDC CORA TM 2011-022
September 2011

Defence R & D Canada
Centre for Operational Research and Analysis

Strategic Planning Operational Research Team
Chief of Force Development

Unifying Capability Integration Analysis – Initial Insights

Leonard Kerzner
DRDC CORA

Defence R&D Canada – CORA

Technical Memorandum
DRDC CORA TM 2011-022
September 2011

Principal Author

Original Signed By

Leonard.F. Kerzner

Strategic Planning Operational Research
DRDC CORA

Approved by

Original signed by Robert Burton

Robert M.H.Burton

Section Head (Joint and Common)

Approved for release by

Original signed by Paul Comeau

Paul Comeau

Chief Scientist

Defence R&D Canada – Centre for Operational Research and Analysis (CORA)

© Her Majesty the Queen in Right of Canada, as represented by the Minister of National Defence, 2011

© Sa Majesté la Reine (en droit du Canada), telle que représentée par le ministre de la Défense nationale, 2011

Abstract

In 2005, the Chief of Defence Staff (CDS) of the Canadian Forces (CF) mandated that Capability-Based Planning (CBP) be institutionalized as a part of a centrally driven, top-down approach to Force Development (FD) within the Department of National Defence (DND). For the last four years, military and defence analyst staffs have developed and implemented the first iteration of an end-to-end capability-based FD process. This process ultimately produces a list of programs necessary to meet the demands of the future security environment over the next 20 years. As part of the Capability Integration phase of FD, distinct methods are employed to determine the set of alternatives which maximizes CF capability that is affordable within the available Defence budget, financial risk, alternative priority and the implementation schedule for the selected alternatives. This report describes the results of a research initiative conducted in 2008 whose purpose was to construct a different optimization scheme which overcomes the processing and validation deficiencies that were observed during SCR work. At the same time, several other Capability Integration components have been incorporated into this revised optimization framework so that much more of the integration related activity, as well as the sensitivity analysis work so necessary to establish solution credibility, can be conducted in an affordable and timely fashion.

Resumé

En 2005, le Chef d'état-major de la Défense (CEMD) des Forces canadiennes (FC) a décrété l'institutionnalisation la planification fondée sur les capacités (PFC) dans le cadre d'une démarche centralisée et descendante de développement des forces (DF) au sein du ministère de la Défense nationale (MDN). Au cours des quatre dernières années, des militaires et des analystes en matière de défense ont travaillé à l'élaboration et à la mise en place de la première version d'un processus fondé sur les capacités de bout en bout. Ce processus a permis de dresser la liste des programmes nécessaires pour répondre aux besoins liés au contexte de sécurité des 20 prochaines années. Durant la phase d'intégration des capacités pour le DF, on emploie diverses méthodes distinctes afin de déterminer des façons d'optimiser les capacités des FC qui respectent le budget, le niveau de risque acceptable sur le plan financier, les priorités et le calendrier d'exécution de la Défense. Le présent rapport contient les résultats de la recherche menée en 2008 dans le but d'élaborer un nouveau cadre d'optimisation qui comblerait les lacunes en matière de traitement et de validation observées durant les travaux d'établissement de la Feuille de route des capacités stratégiques (FRCS). Aussi, plusieurs autres éléments d'intégration ont été incorporés au nouveau cadre d'optimisation, de sorte qu'une plus grande partie des activités d'intégration et d'analyse de sensibilité nécessaires pour établir la crédibilité d'une solution peuvent être menées de manière rapide et abordable.

Executive Summary

Unifying Capability Integration Analysis - Initial Insights,

**Leonard Kerzner; DRDC CORA TM 2011-022; Defence R&D Canada –
CORA; September 2011**

In 2005, the Chief of Defence Staff (CDS) of the Canadian Forces (CF) mandated that Capability-Based Planning (CBP) be institutionalized as a part of a centrally driven, top-down approach to Force Development (FD) within the Department of National Defence (DND). Over the last four years military and defence analyst staffs have developed and implemented the first iteration of an end-to-end capability-based FD process. Among the outcomes of this process is a list of the programs deemed necessary to meet the demands of the future security environment over the next 20 years and an associated schedule for their implementation. This capability plan is formally known as the Strategic Capability Roadmap (SCR).

The Canadian FD process is an integrated sequence of activities starting from the interpretation of government strategic guidance to the delivery of force elements for employment by Operational Commands. The FD process is conducted in three phases with three distinct objectives:

- Capability Planning - what the CF needs to be able to do;
- Capability Management - how well the CF will be able to meet its planning requirement; and
- Capability Integration - how the CF should adapt its plans to better meet its requirements.

Although each phase warrants discussion in its own right, the emphasis of this paper is on Capability Integration. In this phase, multiple capability alternatives are postulated for each identified capability deficiency. An optimization approach employing a genetic algorithm was used to search the solution space to identify the set of projects that provided the best total military capability possible for the available budgets. Once the best set of alternatives was identified, follow-on mathematical models were applied to assess the level of risk associated with the implementation of the capability solution and to explore potential development schedules.

A prioritized list of capability alternative components (projects) identified in the optimal solution was staffed through senior management boards and on 16 July 2008, the Defence Management Committee endorsed the project listing of the Strategic Capability Roadmap 1.0. The Committee also affirmed that from then on projects would be accepted into Investment/Implementation Plans on the basis of their assessment under this planning framework, rather than on the strength of stand-alone operational requirements arguments. This confirmed the intended role of centralized, joint Force Development in DND.

Notwithstanding the success of the first pass of the FD process, improvements were warranted. In the haste to produce the first SCR, some important factors could not be given their due attention. In addition, the degree of integration among the tools within the analytic framework could have been enhanced, thereby simplifying the analytic tasks that must be performed.

Work was conducted by the Strategic Planning Operational Research Team and Chief of Force Development to deal with these issues before their application for the next SCR cycle. Among these issues, the following were considered bottlenecks in the integration process:

- ♦ improving upon the optimization approach used for the selection of alternatives;
- ♦ expanding the variety of output visualizations; and
- ♦ assessing cost risk during the optimization phase.

The technical approach used for optimization in SCR 1.0 provided good quality solutions. However, the tool used was based on a licensed software package, which limited the number of possible users, and had lengthy run-times. Taken together these drawbacks limited the scope for the sensitivity analysis that could have been performed for SCR 1.0. Moreover, there had been no independent mechanism for verifying the optimality of proposed sets of solutions. The presence of a different approach to the conduct of the optimization would have enhanced the credibility of the solutions found.

The next deficiency in the Integration phase involved the post-processing of solution information. Separate models had been built quickly by different analysts for alternative prioritization, cost sensitivity analysis and scheduling. The challenges of working with such a mixed tool set led to little attention being given to output visualization for a broader user base and the inability to tie these analysis components into a seamless whole.

The final deficiency was methodological. It was noted that cost risk analysis was conducted as a separate activity in the integration phase, only after alternatives were prioritized within for a limited selection of solutions. While this activity in itself was not particularly time consuming, a better approach would have been to include the risk assessment as an ancillary process conducted within the optimization.

In response to the above deficiencies, an investigation was initiated to see what could be done to improve upon the user base, optimization turnaround and visualization issues and then other integration processing issues. One track focused on developing a more efficient way of using the tool employed in SCR 1.0. A second track investigated whether a conceptually simpler heuristic could be formulated and tailored to the problem at hand using software more readily available to the Department. Any other efficiency gains over the tool previously used would only be considered a bonus.

The discussion in this report centers on what has been achieved to date along this second track. It describes a set of analytical applications based on a single MSExcel workbook initially supplemented by the readily available @Risk simulation toolset, but which ultimately became self-contained through the insertion of Visual Basic for Applications macros that emulated @Risk functionality. These applications include:

- a custom heuristic for generating good solutions for the alternative selection problem;
- internal and post-processing applications that can be used to gain insight into the fine structure of proposed solutions such as robustness of a solution's set of

alternatives with respect to local changes in cost, capability delivery and the degree of closure associated with critical issues; and

- facilities to assess the impact of cost uncertainty with respect to the risk of solution delivery. This can be accomplished in two ways – either by specifying assurance constraints during the optimization phase or as ancillary results that are captured along with other solution information for post-processing reporting.

Tests of the new heuristic to date show that it performs as well as the commercial application used in SCR 1.0. However, the heuristic resolves the user base, sensitivity analysis and solution validation bottlenecks encountered in SCR 1.0 work.

The post-processing and visualisations represent an important development of the FD tool set as they provide decision makers with a visualisation of the potential trade-offs between ambition and risk within an financially constrained environment. There is potential to develop these tools further to provide decision makers the insights they need to develop the future SCR with an appropriate balance of capability and programmatic risk.

Sommaire

Unifying Capability Integration Analysis - Initial Insights

**Leonard Kerzner; DRDC CORA TM 2011-022; R & D pour la défense
Canada – CORA; Septembre 2011**

En 2005, le Chef d'état-major de la Défense (CEMD) des Forces canadiennes (FC) a décrété l'institutionnalisation la planification fondée sur les capacités (PFC) dans le cadre d'une démarche centralisée et descendante de développement des forces (DF) au sein du ministère de la Défense nationale (MDN). Au cours des quatre dernières années, des militaires et des analystes en matière de défense ont travaillé à l'élaboration et à la mise en place de la première version d'un processus fondé sur les capacités de bout en bout. Ce processus a permis de dresser la liste des programmes nécessaires pour répondre aux besoins liés au contexte de sécurité des 20 prochaines années et d'établir un calendrier d'exécution. Le plan qui découle de ce processus est officiellement connu sous le nom de Feuille de route des capacités stratégiques (FRCS) .

Le processus canadien de DF est une suite intégrée d'opérations commençant par l'interprétation de l'orientation stratégique du gouvernement et qui mène à la mise en place d'éléments de force, qui sont ensuite employés par les commandements opérationnels. Le processus de DF se décline en trois phases ayant chacune un objectif distinct :

- Planification des capacités – ce que les FC doivent être en mesure d'accomplir;
- Gestion des capacités – la mesure dans laquelle les FC seront en mesure de combler leurs besoins en matière de planification;
- Intégration des capacités – de quelle façon les FC devraient-elles modifier leurs plans afin d'être en meilleure posture pour combler leurs besoins.

Bien que chacune de ces phases mériterait une explication détaillée, nous nous concentrerons surtout sur l'intégration des capacités. Cette phase consiste à trouver une solution à chaque lacune sur le plan des capacités. Pour ce faire, nous avons employé un modèle d'optimisation faisant appel à un algorithme génétique qui explore l'espace de solutions afin de déterminer les options offrant le meilleur ensemble de capacités militaires en tenant compte du budget établi. Une fois cette étape franchie, nous avons eu recours à des modèles mathématiques pour évaluer le niveau de risque associé à la mise en œuvre de la solution choisie et étudier les plans de développement possibles.

Par la suite, nous avons fait circuler une liste des options (projets) retenues comme solution optimale et classées par ordre de priorité parmi les comités de la haute direction. Le 16 juillet 2008, le Comité de gestion de la Défense a approuvé cette liste de projets, établissant ainsi la Feuille de route des capacités stratégique, version 1.0. Le Comité a également annoncé qu'à l'avenir, on déterminera les projets à inclure dans les plans d'investissement et d'exécution selon l'évaluation qui en aura été faite en vertu de ce cadre de planification, plutôt qu'en se basant sur la valeur des arguments indépendants relatifs aux besoins opérationnels. Cela confirmait le rôle attendu de la part de l'autorité centralisée de développement des forces au MDN.

Malgré le succès de la première phase de mise en place du processus de DF, des améliorations restaient à faire. Dans la précipitation à créer la première version de la FRCS, certains facteurs importants ont été négligés. De plus, il y aurait lieu d'accroître le degré d'intégration des outils au sein du cadre d'analyse, ce qui aurait simplifié l'exécution des analyses.

L'Unité de recherche opérationnelle en planification stratégique et le Chef – Développement des forces Work se sont penchés sur différents problèmes à résoudre avant la mise en chantier de la version suivante de la FRCS. Voici les questions qui étaient considérées comme les principaux obstacles au déroulement du processus d'intégration :

- amélioration de la démarche d'optimisation employée pour le choix de solutions;
- diversification des options de visualisation des extrants;
- évaluation des risques financiers associés à la phase d'optimisation.

La démarche technique d'optimisation utilisée dans l'élaboration de la FRCS 1.0 nous a permis de trouver de bonnes solutions. Cependant, comme l'outil employé faisait partie d'une suite logicielle sous licence, le nombre d'utilisateurs était limité. De plus, ses temps d'exécution étaient longs. Mis ensemble, ces inconvénients ont limité la portée de l'analyse de sensibilité qui aurait pu être effectuée lors de l'élaboration de la FRCS 1.0. Par ailleurs, nous ne disposions d'aucun mécanisme autonome pour la vérification de l'optimalité des ensembles de solutions proposés. Si nous avions fait appel à une méthode d'optimisation différente, les solutions trouvées auraient eu plus de crédibilité.

La seconde lacune de la phase d'intégration concernait le post-traitement des données. Différents analystes ont conçu rapidement des modèles distincts de priorisation des solutions, d'analyse de l'élasticité des coûts et de planification. Le fait d'employer un tel assortiment d'outils nous empêcher d'accorder suffisamment de temps pour visualiser les extrants d'un plus grand nombre d'utilisateurs et de regrouper ces éléments d'analyse en un ensemble homogène..

La dernière lacune était de nature méthodologique. Nous avons remarqué que l'analyse des risques financiers avait été menée en tant qu'activité indépendante durant la phase d'intégration et seulement après que l'on eût établi la liste des options prioritaires pour un nombre limité de solutions. Bien que cette activité n'exigeait pas en soi beaucoup de temps, il aurait été préférable de faire de l'évaluation des risques un processus auxiliaire mené dans le cadre de la démarche d'optimisation.

Afin de combler ces lacunes, nous nous sommes d'abord penchés sur les problèmes concernant l'utilisation, les délais d'exécution de l'optimisation et la visualisation, puis sur les autres problèmes liés à l'exécution de l'intégration. L'une des pistes de solution explorées concernait la conception d'une manière plus efficace d'utiliser le logiciel ayant servi à l'établissement de la FRCS 1.0. Nous avons également étudié la possibilité de formuler une approche heuristique plus simple et taillée sur mesure pour résoudre le problème en question en faisant appel à un logiciel auquel le Ministère aurait plus facilement accès. Tout autre gain d'efficacité par rapport à l'outil employé précédemment serait considéré comme un atout.

Dans le présent rapport, on traite surtout des résultats obtenus jusqu'à présent concernant cette seconde piste de solution. On décrit un ensemble d'applications analytiques conçues à partir d'un seul classeur Excel, qui avait d'abord été alimenté par la boîte à outils de simulation @Risk et qui, depuis, est devenu autonome grâce à l'insertion de macros Visual Basic imitant la fonctionnalité de @Risk. Ces applications comprennent les éléments suivants :

- une fonction heuristique sur mesure permettant d'apporter de bonnes solutions au problème de choix des options;
- des applications internes de post-traitement pouvant servir à obtenir un aperçu de la structure des solutions proposées; par exemple, la souplesse d'un ensemble d'options donné par rapport aux fluctuations des coûts, à la mise en œuvre des capacités et à la résolution des problèmes critiques;
- des ressources permettant d'évaluer les risques financiers associés à la mise en œuvre de la solution. Cela peut s'accomplir de deux façons : en cernant les difficultés liées à la vérification durant la phase d'optimisation ou en analysant les résultats auxiliaires enregistrés et toutes autres données liées à la solution aux fins de rédaction d'un rapport de post-traitement.

Jusqu'à présent, les essais de la nouvelle méthode heuristique ont montré que celle-ci fonctionne aussi bien que l'application commerciale employée lors de la production de la FRCS 1.0. De plus, la méthode heuristique résout les problèmes d'utilisation, d'analyse de la sensibilité et de validation des solutions que nous avons rencontrés pendant l'élaboration de la FRCS 1.0.

Le post-traitement et la visualisation sont d'importantes fonctions de l'ensemble d'outils de DF, puisqu'elles permettent aux décideurs d'avoir un aperçu des compromis possibles à faire afin de maintenir l'équilibre entre les ambitions et les risques dans un contexte financier restreint. On envisage la possibilité de continuer à développer ces outils afin d'aider les décideurs à concevoir la prochaine version de la FRCS en maintenant un niveau de risque acceptable concernant les capacités et les programmes.

Table of Contents

Abstract	i
Executive Summary	iii
Sommaire.	vi
Table of Contents	ix
List of Figures	xi
List of Tables.....	xii
Acknowledgements	xiii
1 Introduction.....	1
1.1 Background	1
1.2 Purpose	3
1.3 Document Organization.....	3
2 Implementation	4
2.1 Development of the Optimization Heuristic.....	4
2.2 Heuristic Details	4
2.2.1 Mutation.....	5
2.2.2 Crossover	6
2.3 Additional Features	9
3 Application	11
3.1 SCR 1.2 Context.....	11
3.2 Sensitivity to Population Constraint.....	11
3.3 Cost Risk Analysis	13
3.4 Technical/Implementation Risk Analysis.....	17
3.5 Solution Level Analysis	18
3.6 Ordering By FOM – Two Views.....	19
3.7 Ordering Alternative Persistence.....	20
3.8 Summary	21
4 Discussion.....	22
4.1 Overview	22
4.2 Directions For Future Development.....	23
4.2.1 Introduction.....	23
4.2.2 Enhancements for Contingency Management	23
4.2.3 Alternative Scheduling	24
4.2.4 Longer Term Developments	25
References	27
Annex A Capability Optimization in SCR 1.0	29
A.1 Optimization Problem.....	29

A.2 Optimization Approach.....	31
Annex B About @Risk.....	33
B.1 Risk Analysis	33
B.2 Previous Use of @Risk in SCR 1.0 Analysis.....	34
B.3 Use of @Risk in the Optimization Heuristic.....	38
Annex C Technical Details.....	41
C.1 Introduction.....	41
C.2 Grid Element Selection	41
C.3 Grid Element Selection Bias	41
C.4 Mutation.....	41
C.5 Crossover	42
C.6 Stopping Criteria.....	42
C.7 Improving Run Times	43
C.8 Managing FOM Sensitivity.....	44
C.9 Identifying the Pareto Frontier	44
Annex D Estimating Cost Risk from Information about Skewed Distributions	47
D.1 Introduction.....	47
D.2 Cumulants	47
D.2.1 Definition	47
D.2.2 Cumulants and Moments	48
D.2.3 Properties of Cumulants.....	48
D.2.3.1 Invariance.....	48
D.2.3.2 Homogeneity.....	49
D.2.3.3 Additivity	49
D.3 Application.....	50
D.3.1 Theoretical Approach	50
D.3.2 Numerical Example	51
Annex E Some Properties Of Skewed Triangular Distributions	55
E.1 Introduction	55
E.2 Elementary Properties	56
E.3 Percentiles	62
List of symbols/abbreviations/acronyms/initialisms	65

List of Figures

Figure 1: Illustration of Solution Mutation.....	6
Figure 2: Illustration of Solution Cross	7
Figure 3: Cost vs Performance for a Variety of Personnel Constraints.....	12
Figure 4: Cost Assurance Thresholds.....	14
Figure 5: Trade-off of FOM Delivery Potential with Assurance of Delivery	15
Figure 6: Efficiency Frontier Contrasted with Constrained Risk Profiles.....	17
Figure 7: Technical Risk Accepted	18
Figure 8: Issue Resolution	19
Figure 9: Stability of Alternative Orderings.....	20
Figure 10: Alternative Cost Probability Models.....	36
Figure 11: SCR Solution Financial Risk	37
Figure 12: SCR Solution Cumulative Risk Profile.....	38

List of Tables

Table 1: Illustration of Persistence of Alternatives in the Neighbourhood of a Solution.....	21
Table 2: Cost Uncertainty Scale	35
Table 3: Solution Scenario Scorecard	45
Table 4: Cumulants Associated With Selected Cost Uncertainty Bounds	50
Table 5: Sample Cumulants.....	52
Table 6: Central Moments Corresponding to Sample Cumulants	52
Table 7: Normalized Central Moments Corresponding to Central Moments.....	52
Table 8: Normalized Cumulants Corresponding to Normalized Central Moments	52
Table 9: Expected Values Associated With Specified Risk Ranges	58
Table 10: Variance Associated With Specified Risk Ranges.....	59

Acknowledgements

The author would like to acknowledge the contributions and insights provided by Dr. Roman Petryk whose intimate knowledge of the alternative optimization process and the calculation framework for the figure of merit used in SCR 1.0 and SCR 1.2 greatly reduced application development time. He would also like to thank Mr. Mark Rempel for his insights into genetic algorithms which spurred the author's interest to investigate a different approach to conducting the optimization. He would also like to thank Dr Ben Taylor, T/L SPORT and LCol Brian Newman D Mil CM (Integration) for their editorial comments during the development of this report and for the time they made available to work with the author in the design of meaningful output visualizations.

This page intentionally left blank

1 Introduction

1.1 Background

In 2005, the Chief of Defence Staff (CDS) of the Canadian Forces (CF) mandated that Capability-Based Planning (CBP), an approach to force planning that has been well documented and developed by CORA over the past 10 years [References 1 – 8], be institutionalized as a part of a centrally driven, top-down approach to Force Development (FD) within the Department of National Defence (DND). Subsequently, military and defence analyst staffs have developed and implemented the first iteration of an end-to-end capability-based FD process. This process ultimately produces a list of programs necessary to meet the demands of the future security environment over the next 20 years and their implementation schedule. This plan is formally known as the Strategic Capability Roadmap (SCR).

The Canadian FD process is an integrated sequence of activities starting from government strategic guidance to the delivery of force elements for employment by Operational Commands. It is conducted in three phases, each with a distinct emphasis:

- Capability Planning - what the CF needs to be able to do;
- Capability Management - how well the CF will be able to meet its requirements; and
- Capability Integration - how the CF should adapt its plans to better meet its requirements.

Although each phase warrants discussion in its own right, this report concentrates on Capability Integration. In this phase, multiple capability alternatives are postulated for each deficiency identified through the activity conducted in preceding phases. Alternatives can involve acquiring new equipment, changing tactics, techniques and procedures, upgrading current equipment, defining new concepts of operation, changing the role of existing equipment or combinations of these options. Alternatives can completely or partially resolve a deficiency and can vary widely in resource demands. Each alternative is fully defined in terms of the cost, lifecycle, personnel requirements, deficiency closure and implementation risk.

The most challenging part of Capability Integration is determining the set of capability alternatives that will maximize CF capability while remaining affordable. An optimization model, more fully described in Annex A, was developed which attempts to maximize the value of deficiencies resolved, maximize compliance with the Objective Force (a concept for the ideal future CF), minimize additional personnel requirements, minimize cost and minimize implementation risk. A genetic algorithm was applied to search the solution space to identify the one that best met the goals of the optimization while abiding by the imposed constraints. Once that solution was identified, mathematical models were applied to assess the level of implementation risk assumed and to generate a development schedule.

A prioritized list of the capability alternative components (projects) identified in the optimal solution has been staffed through senior management boards. On 16 July 2008, the Defence Management Committee endorsed the project listing of the Strategic Capability Roadmap 1.0. The Committee also affirmed that from then on projects would be accepted into Investment/Implementation Plans on the basis of their assessment under this planning framework,

rather than on the strength of stand-alone operational requirements arguments. This confirms the future role of centralized, joint Force Development in DND.

Notwithstanding the success of the FD process, improvements are warranted. In the haste to produce the first SCR, some important factors could not be given their due attention. In addition, the degree of integration among the tools within the analytic framework could have been enhanced, thereby simplifying the analytic tasks that follow once an optimal set of alternatives is identified.

Work was conducted in the Strategic Planning Operational Research Team (SPORT) and Chief of Force Development (CFD) to deal with these issues before their application for the next SCR cycle which is estimated to be in the 2011/2012 timeframe. Among the activities associated with this work, research was initiated to improve upon the approaches used for the optimization of alternatives, enhance user oriented visualizations and more closely integrate cost and implementation risk, alternative prioritization and scheduling into a single analytic process. These were considered process bottlenecks during Integration Phase of the SCR development.

While it did provide the required high quality solutions, a noted drawback of the SCR 1.0 approach was the limited ability provided to ordinarily-trained military staff officers to perform the optimization. Since SCR 1.0 was an innovative activity, there was little partner concern with the degree of Operational Research support dedicated to the effort, both in terms of personnel and the specialty software required. However, future SCR activity will be expected to be supported by regular trained CFD military personnel with minimal OR involvement.

Two issues required immediate resolution in order to address this drawback: the limited number of licensed copies of the optimization software available to DND; and the processing time to generate a set of optimal alternative packages for the cost range of interest. The first limited the number of platforms that could be dedicated to conduct the optimization. The second also constrained the number of runs that could be performed during the extremely limited timeframe that was made available for the analysis. Taken together these issues limited the scope for the sensitivity analysis that could have been performed for SCR 1.0. Furthermore, there had been no independent mechanism for verifying the optimality properties of proposed sets of solutions. The conduct of a parallel analysis using a different optimization approach would have enhanced the credibility of the optimal solution found by the SCR analysis team.

The next deficiency in the Integration phase involved the post-processing of solution information. Separate models were required for alternative prioritization, cost sensitivity analysis and scheduling, each developed by a different SPORT analyst, each requiring different protocols for information passing and handling. This was caused by the rather compressed timeframe available to do this post-processing work, which only could be resolved by designing the processing segments in parallel without complete foreknowledge of what was expected of results reporting. This led to little attention given to tie these post-processing components into a coherent whole.

In response, a two-tracked investigation was initiated to see what could be done to first improve upon the user base and optimization turnaround issues and then other integration processing issues as they arose. One track focused on mastering the tool used in SCR 1.0 so that either through problem reformulation or the identification and adoption of a more efficient combination of control parameters processing times could be shortened. A second track investigated whether a conceptually simpler heuristic could be formulated, tailored to the problem at hand using software

more readily available to the Department. The advantage of this would be that more users could be conducting a greater number of sensitivity cases during the available analysis period. Furthermore, if the running time for optimization problem could be significantly shortened even greater opportunity could be had for sensitivity analysis work. Lastly, the solutions identified by the SCR 1.0 and any tool developed on this track can be assessed in the other tool to verify their optimality properties.

1.2 Purpose

The discussion in this report centers on what has been achieved to date along the second track. In particular, it describes a set of analytical applications including a custom heuristic for generating good solutions for the alternative selection problem. Beyond this, there are several internal and post processing applications that can be used to gain insight into the fine structure of proposed solutions. These examine the robustness of a solution's set of alternatives with respect to local changes in cost, Figure of Merit (FOM) delivery, the impact of cost uncertainty in respect to the risk of solution delivery and the degree of closure associated with critical issues. These are described in greater detail later in the development, along with sample outputs depicting the kind of information that can be generated by these applications.

As of writing, there is scope for further development of this tool to allow for a broader range of post-processing work. One aim is to define a time precedence order for the alternatives that comprise the solution which would immediately establish an outline of the schedule for the projects associated with these alternatives. This may come to mean enfolded the Roadmap aspect of the SCR as part of the alternative optimization activity. This is desirable since it is the lack of a consistent and unbiased procedure to define such a schedule that impedes greater cooperation between the Force Development, Programming and Force Generation communities in the delivery of timely and appropriate military capability.

1.3 Document Organization

The main body of this report describes the background of the research requirement leading to the development of the series of applications discussed in this report. It outlines what has been accomplished followed by details about how components were implemented. Next, examples of the kinds of results visualization made possible through these applications are presented. The main body concludes with an overview of the significance of this development, with suggested insights into two areas that provide for a prototype alternative selection and scheduling system when fully implemented.

This document has five Annexes. Annex A provides an overview of the analytic formulation of the capability optimization problem and the approach used to tackle it in SCR 1.0. Annex B discusses the simulation engine used for the optimization component of this application, its previous application in SCR work and how dependent the application is upon it. Annex C touches on some technical issues regarding the heuristic that may be of interest to future system users and developers. Annex D presents a method of cost risk assessment applicable to general classes of cost distributions. Annex E presents some basic facts about triangular distributions which suffice to enable application of the method of Annex D. This permits the optimization to be conducted in consideration of cost risk constraints – which was not considered during SCR 1.0.

2 Implementation

2.1 Development of the Optimization Heuristic

To address the shortcomings of the SCR 1.0 approach to optimizing alternatives, the author investigated whether it was possible to develop a more user accessible method. The aim here was not necessarily to find the “best” method of conducting the optimization but to possibly discover an approach that could be used by regularly trained military officers that performed as well if not better than that used for SCR 1.0. The outcome of this investigation was initially an application of the @Risk framework marketed by Palisades Systems for creating simulations from MSeExcel workbooks. At the time of development @Risk was approximately one third the price of the toolset used in SCR 1.0, enabling the Department at a minimum to triple its processing capacity with respect to Capability Integration. Moreover, DND owned over twenty copies of this software that could be made available for surge analyses of the type encountered in SCR work. The workbook underlying the framework is for the most part identical to that used as the calculation underpinning the previous analysis. Any difference between the two workbooks was the result of the addition of a control section to assist the user in defining simulation characteristics at various points during the analysis as well as a visualization section which is used to monitor simulation progress and to organize processed information into user oriented reports. For further detail about @Risk uses and capabilities, previous applications of @Risk in SCR work and the product features in the heuristic, refer to Annex B.

Further along in the development, it was noted that most @Risk functions used could be replaced by either inherent MSeExcel functions or simple combinations of such functions. Simulation control and accumulation of statistics, the reason for @Risk in the first place, was emulated through an embedded MSeExcel macro. As a result, the heuristic was modified so that it was entirely supportable on any platform that can run MSeExcel. Since this application is a Departmental standard, accordingly the alternative heuristic developed can be run from almost any DND workstation.

2.2 Heuristic Details

One way of generating approximate solutions for problems with large solution spaces is to apply directed random search techniques. This works by developing a random combination of alternatives, calculating the resulting FOM, verifying that problem constraints are met and retaining for further investigation only those combinations that perform best. Should another solution be found that outperforms the best solution found to date, then the new solution should replace the former best solution as the best solution found to date.

While conceptually simple, this approach tends to be wasteful of computing resources. This is the equivalent of searching for a needle in a haystack by sticking one’s arm in from a random direction, for a random depth and hoping to grab the needle. This approach is further at a disadvantage for one is often more interested in the set of good solutions in a certain range about the cost target, especially if these solutions could show improvement in some of the ancillary variables under examination. For example, a solution with slightly higher cost within the vicinity

of the desired cost could be accepted if it were associated with lower technological or financial risk, resolves more issues and has a lower personnel demand.

After an initial and, given the above, unsurprisingly disappointing trial involving pure directed random search, the heuristic was restructured to concentrate on the simultaneous solution of a grid of cost constrained problems about a user specified point of interest. A cost grid is comprised of a series of equally spaced, contiguous ranges covering the cost span to be investigated. A grid element is associated with a local range of costs and an interim solution whose cost lies within that range which provides the best FOM of such solutions found to that point in the optimization. The heuristic for any given analysis begins by first zeroing out all stored solution information throughout the cost grid. Next, an element of the cost grid is randomly selected for processing. The solution currently stored for the grid element, initially a zero vector, is extracted and is subjected to further improvement applying custom mutation and crossover processes that emulate those employed in genetic algorithms¹. Essentially, these are the same processes that are used for solution optimization in SCR 1.0, with the exception that the user is given a stronger set of process controls to use these methods in the custom heuristic. Moreover the details of the implementation of mutation and crossover processes in the heuristic are available for inspection by the user community, which was not the case for the method used in SCR 1.0.

2.2.1 Mutation

Mutation as defined in this paper means the alteration of a tentatively assessed good combination of alternatives defining a solution for a given cost by replacing a certain fraction of the selected alternatives by other available alternatives. Figure 1 below illustrates more concretely what happens in mutation. The Figure shows that each of 10 mutation loci has up to six alternatives identified by a number ranging from zero to five. Think of a mutation locus as a deficiency requiring a decision about its alternatives. Each solution associated with a grid element is characterized by a selection of an alternative at each locus. The particular selection for the example is shown under the column Interim Solution. The mutation procedure first involves identifying loci to be changed. This is done by random draw against the specified mutation rate for each locus. In this example, loci 04 and 09 have been selected to undergo mutation. This has been designated by the asterisk lying in the column in the middle of the Figure. The next step is to select the alternatives that will be associated at these loci. This is done randomly from the alternatives not associated with interim solution at the loci. In this example, alternative 2 in locus 04 is replaced by alternative 0 while alternative 0 in locus 09 is replaced by alternative 4. The characteristics of the mutated solution are identified in Figure 1 by the column labelled Mutated Solution.

In practice, there can be restrictions at a locus as to which alternatives are feasible for it. Accordingly the selection process for alternatives in mutation has to accommodate these restrictions. The heuristic can easily deal with this having the user explicitly assign zero probability for selection of unavailable alternatives.

¹ A genetic algorithm, [http://en.wikipedia.org/wiki/Genetic_algorithm], is a search heuristic that mimics the process of natural evolution. Genetic algorithms belong to the larger class of evolutionary algorithms, which generate solutions to optimization problems using techniques inspired by natural evolution, such as inheritance, mutation, selection, and crossover. Both classes of heuristics are routinely used to generate useful solutions to optimization and search problems.

Should a mutated solution rate better with respect to FOM than its generator and lie within the cost span of the grid element of its generation, then the mutated solution is preferred. This means that the mutation solution replaces that of its generator for that grid element. Alternatively, a suggested mutated solution may no longer satisfy the constraints met by the solution that spawned it. Should this arise through overall infeasibility, then the solution is rejected and new grid element is selected to provide a solution generator for the next stage of the process. If, on the other hand, the mutated solution is otherwise feasible but its cost does not lie within the cost range of its progenitor, then a FOM comparison is made between this solution and the solution associated with the grid element whose local range spans the mutated solution's cost. If the mutated solution's FOM is greater, the mutated solution becomes the interim solution for its associated grid element.

The mutation process has been optimized so that if a deficiency's tentatively selected alternative is to be mutated, a definitely distinct alternative will be chosen in its place. This requires more sophisticated processing than implementing mutation by pure random selection among all possible acceptable alternatives - which allows the re-entry of alternatives into the mutated solution that had been identified to be discarded. The extra processing to ensure the discard of a designated alternative is compensated by the more rapid movement towards heuristic convergence in most analysis situations tested to date.

Interim										Mutated										
Solution										Solution										
Locus 01	0	1	2	3	4	5				0	1	2	3	4	5					
Locus 02				0	1	2	3	4	5				0	1	2	3	4	5		
Locus 03	0	1	2	3	4	5				0	1	2	3	4	5					
Locus 04				0	1	2	3	4	5	*					0	1	2	3	4	5
Locus 05		0	1	2	3	4	5					0	1	2	3	4	5			
Locus 06		0	1	2	3	4	5					0	1	2	3	4	5			
Locus 07	0	1	2	3	4	5					0	1	2	3	4	5				
Locus 08			0	1	2	3	4	5					0	1	2	3	4	5		
Locus 09					0	1	2	3	4	5	*		0	1	2	3	4	5		
Locus 10		0	1	2	3	4	5					0	1	2	3	4	5			

Figure 1: Illustration of Solution Mutation

2.2.2 Crossover

Crossover as used in this paper is a process that creates one or more derived solutions through the transfer of the alternative selections for each particular deficiency between a pair of interim solutions maintained in the cost grid. Once a grid element is selected for a crossover comparison, the best solutions associated with the grid element's nearest neighbours (in the grid) generate a hybrid trial solution which is then compared to the best solution for the original grid element. Parent solutions are typically those retained for the grid elements whose cost ranges lies one increment above and below that of selected the grid element. A recent extension of this procedure allows the grid element distance between parents to be randomized within a small range for grid element offsets. The hybrid is constructed by selecting for each deficiency, the alternative

associated with one of its parents at a rate of 50% of any particular parent's alternatives. To add further variety to hybrid, the heuristic allows mutation of the cross at an effective rate of no more than five percent at each mutation locus. The idea behind cross is illustrated in Figure 2.

Suppose a grid element has been selected for solution generation. Suppose further that the offset for the cross is set to be one unit. This means that the parents are the interim solutions associated with the grid element whose range is offset one increment above and below that of the selected grid element. The respective alternatives for each locus are shown in the second and third columns of the Figure. The hybrid is generated by selecting for each locus the alternative associated with one of its parents at a rate of 50% from each parent. Where the alternatives are identical at a locus, which occurs quite often, the hybrid would share that alternative. Where these are different, the described selection scheme applies. In this case, the hybrid at locus 04 has taken its alternative from Parent2 while at locus 09 the hybrid has taken its alternative from Parent1. This is highlighted in the Figure by the cells with a lime green background.

	Parent 1	Parent2	Hybrid
Locus 01	4	4	4
Locus 02	3	1	3
Locus 03	5	5	5
Locus 04	2	0	0
Locus 05	3	3	3
Locus 06	0	0	0
Locus 07	4	1	1
Locus 08	2	2	2
Locus 09	0	4	0
Locus 10	1	1	1

Figure 2: Illustration of Solution Cross

As indicated earlier in the discussion concerning mutated solutions, a crossover-developed solution may violate the range constraint associated with the grid element that motivated its construction. Nonetheless, if the candidate solution's cost is within the cost range for some grid element, then a comparison can be made between that element's interim solution and the candidate. Once again, if the hybrid solution is seen to provide an improvement upon the FOM associated the grid element's interim solution, the candidate is then adopted as the interim solution for that grid element. Otherwise, it is discarded and a new candidate is generated.

Whether a candidate for comparison has been generated through mutation or crossover, a preliminary evaluation of the FOM and its other ancillary attributes is conducted. It is at this stage that constraint feasibility is tested for the solution. Possible constraints built into the heuristic include personnel limits, cost range span, risk associated with cost uncertainty, implantation risk and objective force compliance. Any solution found violating a constraint is immediately designated infeasible and is subsequently excluded from comparison processing. The heuristic then continues by selecting a new grid element for solution generation.

Once a solution has been identified which outperforms an interim solution associated with a cost grid element and the subsequent replacement of the latter solution by the former effected, a follow-on calculation is made to determine whether the solution lies on the efficient frontier of the current set of solutions. A solution is efficient when no other solution meets or exceeds its FOM at lower cost. The set of efficient solutions identifies the efficient frontier. The efficient frontier thus outlines the “best” that can be achieved during the optimization where the range grid consists of a single element with an upper bound at the cost of interest.

If a solution is identified as being on this frontier, the weight for future selection of this grid element is reduced so that grid elements associated with dominated solutions are given greater opportunity for selection in succeeding simulation passes. In effect, the heuristic favours the replacement of dominated solutions with efficient ones wherever possible. If later in the heuristic, the solution becomes dominated, then the solution’s weight is reset to that applicable to all dominated solutions at that stage in the algorithm. The specific bias applied to dominated solutions depends on the percentage of range grid elements whose interim solutions are non-dominated. This means that at the beginning of the heuristic, grid element selection strongly prefers those whose interim solutions are not dominated. At higher densities of non-dominated solutions the bias is nearly zero – meaning effectively that all grid elements are given nearly equal opportunity for selection.

As a result of the greater necessity of comparison of a mutated or crossed solution to some point in the solution grid, this heuristic will tend to reassign current solutions more frequently than if a series of strictly cost constrained problems is tackled. This accounts for the high quality outline of the efficiency frontier that was found within 500 generations for a grid of 100 points.

There does not appear to be a good a priori stopping condition for the heuristic. For this heuristic, a generation is considered the application of the basic simulation step for the number of elements in the range grid. Where the grid contains 64 range elements, the generation size is to be 64 simulation iterations. Similarly for a grid with 100 range elements, the generation size is 100. For the sample application at hand, with a grid of 100 range elements the heuristic has run for 3000 generations with at least one value on the grid showing change. At this stage of the heuristic changes most often occur on the fourth significant digit of the FOM. This would be barely detectable in the comparison charts that are used to monitor algorithm progress. A good compromise for practical application is 1500 generations. Run times to this level of detail take three hours on a laptop personal computer. A short discussion about how to determine when to stop the simulation is contained in Annex B.

The heuristic can be expanded so that a much finer grid of range elements can be examined. Indeed trials have been conducted on grids of up to four hundred elements. It should be noted however that run times will increase commensurately with the change of refinement. For the limited number of examples examined to date, the time taken to identify solutions close to the optimum appears to increase linearly. Annex B contains a discussion about what can be done to improve run times for such cases.

2.3 Additional Features

Beyond determining which are the best solutions found in the grid, the heuristic also keeps track of several ancillary solution attributes. In particular, a solution is associated with a nominal cost, its expected cost, its cost standard deviation, total military and total civilian personnel requirements, a measure of technological risk, a measure of objective force compliance² and a measure of effective issue closure³.

The tool in which this heuristic is embedded has several post-processing modes from which other useful information can be extracted. Emphasis is initially on examining cost and technological risk across the set of solutions found.

The cost risk associated with a solution can be estimated and the solution's impact on the accepted extent of issue closure assessed. In this case, the nominal costs for alternative, that is the cost that comes with alternative specifications reflecting the most likely cost of the alternative, also come with upper and lower percentages for cost uncertainty analysis. These uncertainty bounds are typically of a form of x percent below the nominal value and y percent above. As was the case for SCR 1.0 (see Annex B), this information is translated into the determining characteristics of a triangular distribution of the alternative cost that is compatible with the nominal and bound information [Reference 12, 14]. Accordingly, since the solution cost is the sum of the costs of its component alternatives, it is thereby possible to use the characteristics of these component distributions to determine the mean (average) and the variance of the solution's cost.

With this information, the law of large numbers can be applied to the resultant distribution so that the probability that the solution cost can exceed a fixed cost threshold can be approximated. Conversely, thresholds can be defined so that the cost of implementing a desired solution can be assigned a specific probability of cost risk. The mechanism for translating triangular distribution information to do this is discussed in two steps - Annex D to translate the characteristics of a much broader set of distributional information into a risk calculation and Annex E to develop these characteristics for triangular distributions.

While the appropriateness of a triangular representation for cost uncertainty as given from user information can be questioned, it is not essential to the demonstration that cost risk assessment can be conducted at the same time and with the same tools as the optimization without the need for the separate analysis phase used in SCR 1.0.

Indeed, Annex D outlines how much broader classes of uncertainty distributions, specifically any for which the first five central moments can be calculated, can be applied to a method extension. The same extension also covers cases where the number of resolvable issues is sufficiently small

² Objective Force Compliance is a rating scale that encapsulates the degree to which the alternatives selected for a proposed solution align with the CF vision for its future. This is more fully described in Reference 14.

³ Issue Closure is a rating scale by deficiency of the degree to which these can be resolved by the selected alternative in a given solution. This measure includes the impact capability interdependencies. Once again, this is more fully described in Reference 14.

so that normal approximation of cost risk is of dubious value. For this paper, the extension was not used since the number of alternatives resolved in even the lowest cost solutions examined exceeded the threshold for the use of a normal approximation about the expected value of the solution cost.

The derived cost thresholds are plotted against nominal cost in order to gain a sense of the impact of this component of risk on solution affordability. Similarly, a plot of technological risk associated with these can be portrayed to monitor what degree of risk mitigation is in order.

Some of these additional analysis modes are geared to a specific solution. For example a plot can be generated which portrays the issue value against the degree of issue closure associated with any particular selected solution. This chart, shown as Figure 8 in the next section, can be thought of as a risk management profile plot in that successful issue resolution should address high value issues with a high degree of closure with less emphasis shown in closing lower valued issues.

The tool developed also can be used in other aspects of solution analysis. For example, the alternatives in a solution can be rated by the degree of contribution to the FOM achieved by each component. This is accomplished by successively removing the least significant alternative contributing to the FOM, thereby retaining the most significant combinations of alternatives. It should be noted that this is not the same combination that can be derived by ordering the alternatives by direct FOM contribution. The former ordering can differ from this order since the impact of the selection on other issues becomes a factor in the ordering. This will be demonstrated in a sample chart in the next section.

Another way of assigning precedence to the alternatives in a solution is through the persistence of the alternative along a range of neighbouring cost intervals. It is evident that an alternative that occurs in a greater percentage of the solutions in a range would likely be preferred over alternatives that occur less frequently.

The advantage of assigning an order to alternatives is that it organizes solution implementation. With other factors treated equal, one should consider implementing a solution so that the greatest contribution can be had to dealing with the identified issues. Another advantage of imposing an order on the alternatives relates to managing cost risk. Should the cost of implementing the solution exceed planning thresholds, the ordering identifies the capability deficiencies which should be deferred in the sense that these would have the least impact on the FOM of what could be affordably delivered. Since prioritization is highly dependent on the solution selected for analysis, a specific example of the kinds of output that can be obtained has not been provided in detail. Nonetheless, an indication of what can be done can be seen in the comparison of the two deficiency rankings indicated in Figure 9 of the next section and a previously published chart taken from Ref 14 but reproduced for this document as Figure B -3.

3 Application

3.1 SCR 1.2 Context

The heuristic outlined in the previous section was applied to the SCR 1.2 analysis, sponsored by the Chief of Force Development, which was conducted in the summer of 2009. The objective of the analysis was to revalidate the assumptions underpinning previous SCR work and to update its results in light of government announcements which had closed a number of the issues that were present during SCR 1.0. SCR 1.2 preparatory work gave CFD staff an opportunity to review assumptions, include the impact of scenario analysis conducted in the interim and to expand analytical tools facilitating the processing and visualization of outcomes. Moreover, in SCR 1.0, the deficiencies in certain capability families were treated as resolved through the selection of specific alternatives. The cost of these alternatives was deducted from the overall SCR 1.0 cost envelope to define a net cost. The optimization then was conducted for this net cost to best resolve deficiencies over the remaining capabilities. In SCR 1.2, these capability families were now put on analytic par with the retained elements of other capability domains – with the optimization conducted over the full SCR 1.0 cost envelope and with the possibility of other or no alternatives for these deficiencies selected as part of the solution.

As indicated, the scope of SCR 1.2 was limited to issues that could not be addressed by previous government announcements or stated intent. What remained to be examined were alternatives often comprising of many project initiatives at speculative capital status. In addition, a stricter approach was taken in SCR 1.2 with respect to personnel. In SCR 1.0, total personnel costs were understated in that these costs were first counted in a solution provided the number of personnel in that solution exceeded a specified threshold. This threshold reflected the residual capacity of the CF Force Growth anticipated in the planning horizon to operate and support the alternatives selected in a SCR 1.0 solution. For SCR 1.2, it was necessary to give complete visibility to the solution's personnel cost. Moreover, there had been concern in the analytic framework about the achievability of the SCR 1.0 program since guidance at the time directed that full available budget was to be used to define the best performing set of alternatives. Considering that some alternatives during SCR 1.0 work had come with rather large cost uncertainties, this meant that not as much potential capability could be realized since there would be substantial risk that the components of the solution could not be delivered since the solution total cost at delivery would exceed the given accrual funding and operating budget envelopes. To forestall this possibility in SCR 1.2 work, solutions would be considered for implementation if they were associated with a sufficient degree of program delivery assurance. The calculation of the solution's assurance was made into an ancillary value associated with the solution so that assurance constraint could be made part of the optimization. Lastly it should be noted that the cost uncertainty profiles used in SCR 1.0, specified at the time by triangular distributions, were retained for this analysis.

3.2 Sensitivity to Population Constraint

The first area of concern involved the extent to which it was possible to deliver closure against the unresolved issues over a user-specified range of costs with specific caps on the numbers of military personnel. Four cases were examined with caps of 2000, 3000, 4000 and an unlimited number of additional personnel. In each case, it was necessary to map out the efficient frontier to

a high degree of fidelity and then contrast the anticipated performance across the cost span. The outcome of this comparison using the heuristic is shown in Figure 3. Each point on the Figure corresponds to a feasible non-dominated solution for the appropriate constrained problem. The vertical axis represents the FOM achievable for the SCR 1.2 over an expansive cost range. In what follows scale values for axes have largely been suppressed in order to make the model outcomes available to the wider defence analytic community. FOM is the same figure of merit used in SCR 1.0 work and is the objective function for the optimization described in Annex A.

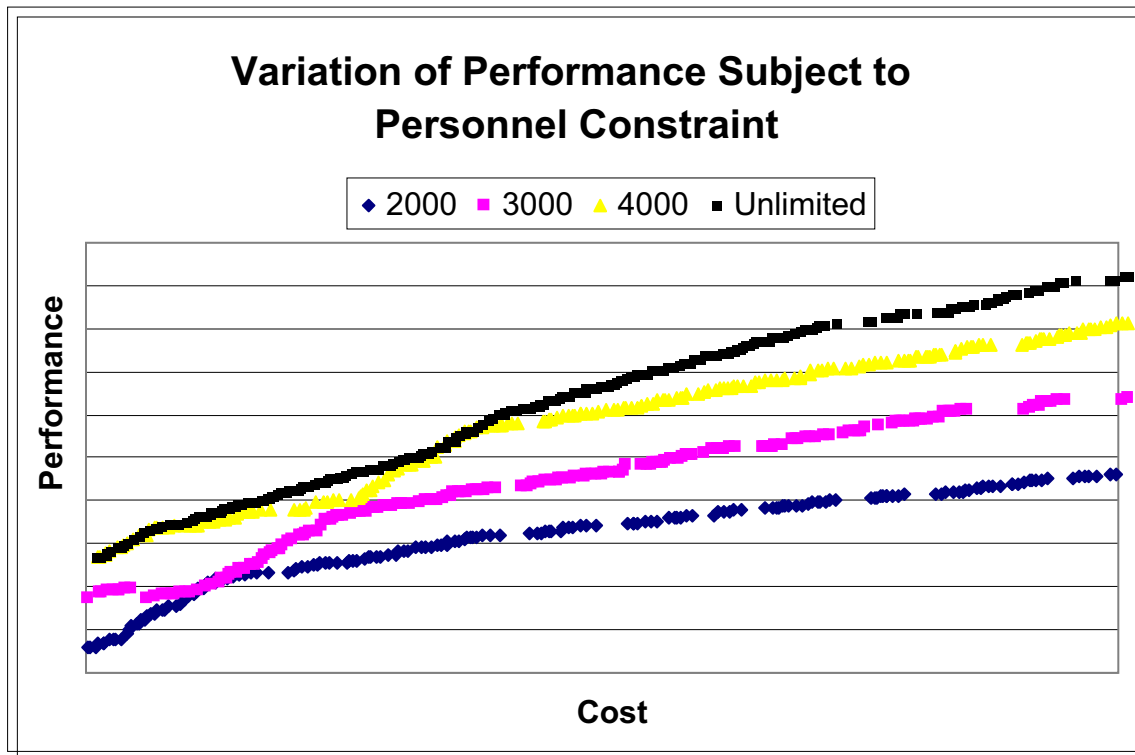


Figure 3: Cost vs Performance for a Variety of Personnel Constraints

As is to be expected, there is definite advantage in delivery of the performance metric with increasing personnel numbers. One notes however that there are particular solutions having lower personnel numbers that deliver almost as much, though not quite the same, performance as can be had with higher personnel numbers. This is seen in the near convergence of the blue diamond and purple square lines, and the purple square and yellow diamond lines towards the left of the figure. The same occurs with the pair of lines comprised of black diamond and yellow triangular points respectively slightly to the right of the first two points but in this case, the convergence refers to the solutions for the unlimited case meeting those of the 4000 personnel constraint. This pattern has also been observed using the SCR 1.0 optimization methodology on the SCR 1.2 dataset. This pattern indicates that one or more technologically centered alternatives can be adopted into more highly constrained cases that trade-off against extra permitted personnel associated with lesser constrained cases at the noted cost levels.

3.3 Cost Risk Analysis

Figure 4 below shows how total cost at specified levels of risk is related to the nominal cost of each optimized solution through the cost uncertainties of each of its component alternatives. The horizontal axis represents nominal cost while the vertical axis represents possible delivery cost. The value of the y-coordinate of a point on the chart associated with a given assurance level indicates that there is a degree of certainty that whole package can be delivered at or exceeding that y-coordinate for the optimal package associated with the cost given by the point's x-coordinate. The graph has been scaled so that the graph vertex is at a common nominal cost and the unit scales on the axes are identical. One notes that regardless of risk, the assurance lines generally slope upward, that is for any degree of program delivery assurance, there is first a requirement to establish an appropriate cost buffer and secondly this buffer increases with increasing nominal cost.

While there is a general tendency for costs at a fixed assurance level to increase with nominal cost, there can be circumstances where this may not strictly follow. It can happen that a solution can have a greater cost uncertainty than an adjacent solution with higher nominal cost. It all depends on the uncertainty profiles associated with the alternatives selected for each solution. These profiles have no direct impact on the outcome of the optimization process for identifying good performing sets of alternatives for a given nominal cost.

Note that a jump occurs in the span between the 1% and 99% FOM delivery assurance points toward the left hand side of the chart. The jump is a consequence of a change in the nature of the solutions on the efficiency frontier that arises from the sample data. Efficient solutions whose costs lie to the right of the jump are associated with alternatives that have larger cost uncertainties than those to the left. In particular, all represented solutions with nominal costs to the right of the jump satisfy the 3000 personnel constraint whereas those to the left have their personnel number below the constraint. For solutions to the right of the jump it was necessary to select effective alternatives having greater cost uncertainty so that personnel total for the solution would not exceed the constraint to compensate for rejected alternatives that were as or more effective but were associated with less cost risk. This change in the nature of portrayed solutions is also reflected in the first kink in the associated FOM line in Figure 3, in the break in the risk curve in Figure 5 and the divergence pattern of cost assurance curves from their FOM envelope in Figure 6.

Charts such as Figure 4 are very important for program planning. Suppose a maximum cost threshold has to be established which is to provide for a specific assurance level of program delivery. One looks up the threshold cost along the vertical axis and projects it horizontally until it intersects the appropriate assurance line. From the point of intersection, one projects vertically downward to determine the associated nominal cost on the horizontal axis. One then uses the efficient solution most closely associated with the nominal cost for planning purposes. Due to the law of large numbers as applied to the cost uncertainties for the sample dataset, set of points at 50% assurance corresponds to the expected value of the solution for a point with specified nominal cost. This is also the sum of all the expected delivery costs for the alternatives in the solution. The difference between the expected and the nominal cost represents an internal contingency accepted for the solution. The funding for this difference could be available to capability managers for trickle release as planned nominal costs for alternative delivery can no longer be met. The buffer between the selected assurance level and this expected cost represents

an external contingency that has to be provided for the collective of all alternatives if the package has to deliver at the assured risk. This contingency could be managed as a VCDS reserve made available to ensure delivery of critical alternatives once their expenditure exceeds expected costs.

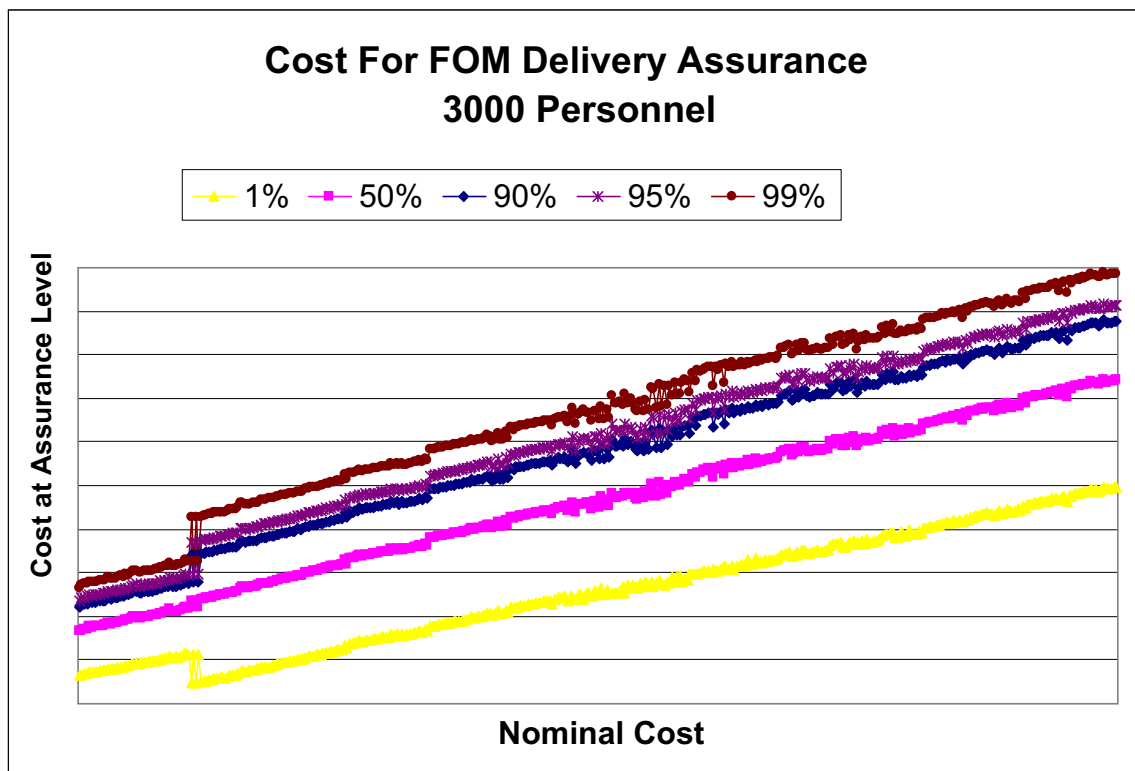


Figure 4: Cost Assurance Thresholds

Another item to note in the above Figure is that despite the fact that the nominal cost for an alternative always lies between its maximum and its minimum, the nominal costs for a solution often lies well below the 1% delivery assurance line. This should not be unexpected as the rightward skewness of individual cost distributions assign greater probability to costs above the nominal than otherwise. As this is applied consistently, it would not be unreasonable that alternative delivery at the solution's nominal cost can be achieved at such low probabilities.

Figure 5 integrates the issue of FOM potential delivery with risk of solution delivery at nominal cost. Implicit in this chart is a fixed cost threshold that is used as an upper bound of FOM delivery. Specific FOM values are related to this upper bound as a percentage. The line defined by pink squares accordingly shows the percentage of the maximal Figure of Merit that can be delivered by the heuristic solution associated with the nominal cost. For simplicity of labelling, the percentage associated with a point on this curve is called its FOM Achievability Potential, simplified to Achievability, at a given nominal cost. For the range of interest, it is noted that these percentages exceed 93% of the potential that can be had at the cost threshold. As nominal cost is increased, this potential increases since the associated deliverable FOM increases with increasing cost.

The blue line shows the program delivery assurance, simplified to Delivery, associated with the heuristic solutions for the range of nominal costs. As nominal costs are increased, the solutions found contain increasing numbers of alternatives with much wider bands of cost uncertainty. This severely compromises the ability to deliver the solution at the given cost threshold. This is indicated by the declining assurance for package delivery as nominal costs are increased.

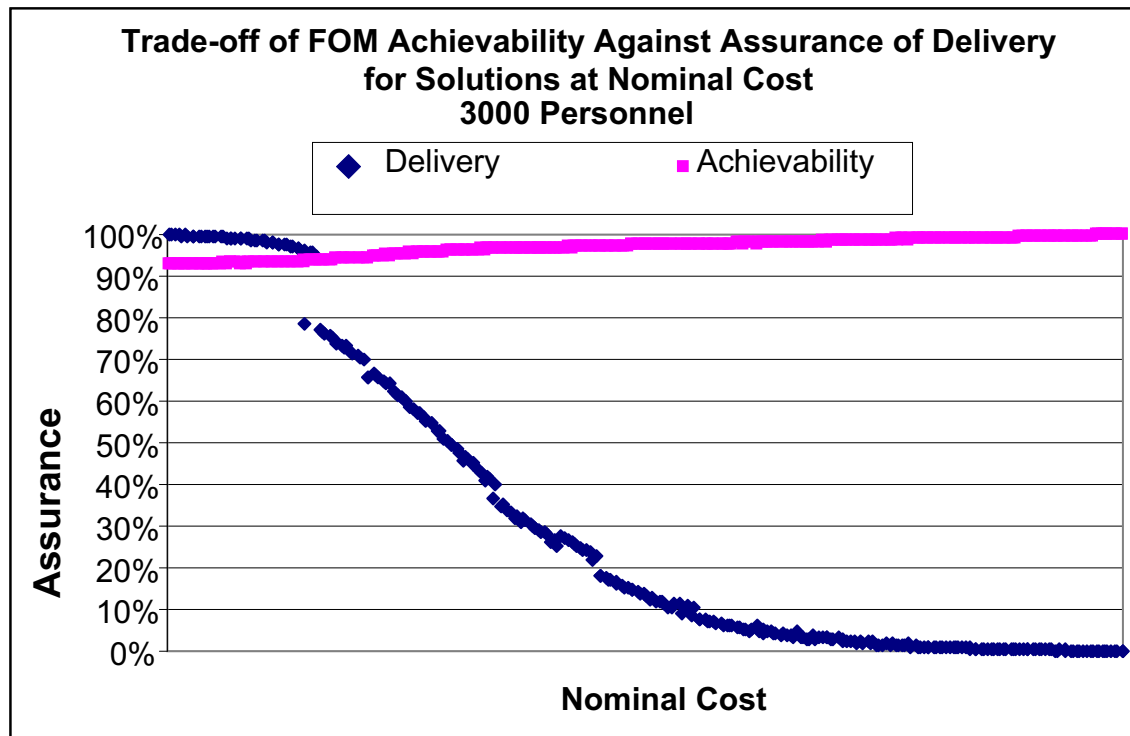


Figure 5: Trade-off of FOM Delivery Potential with Assurance of Delivery

Note that the abrupt risk increase phenomenon shown in Figure 4 is reflected in Figure 5 by sudden decrease in delivery assurance. The chart reflects the assurance value associated with the best FOM producing solutions for the range of nominal costs examined. To the left of the break, solutions represented lie below the personnel cap of 3000 personnel. The FOM for solutions in this range can be achieved through low risk but personnel intensive sets of alternative. For nominal cost values to the right of the break, no more of these alternatives can be part of their respective solutions since their personnel level meets its cap. Accordingly, best FOM solutions in this range need to trade off the possibility of greater FOM (as might be realized in the black line in Figure 4) with the acceptance of greater cost risk to remain feasible. This qualitative transition accounts for the break and the rapid decline in delivery assurance. Also note that some points in the vicinity of the break in the Delivery curve are reported to have lower Delivery than a neighbouring point at higher nominal cost. This is a consequence of the fact that the optimization procedure had to be terminated before very fine resolution of the risk/cost profile efficient frontier could be attained.

If one were to use the information portrayed in the chart for implementation, it would be strongly suggested that those solutions with nominal costs at the low end of the planning range should be

adopted. While this means compromising on potential force effectiveness as measured by the FOM, this is made up for the high associated assurance for package achievability.

In general, where phenomena generated by the encountering constraints over the cost range represented can be avoided, the two curves shown would both transition with nominal cost without any breaks and intersect at a unique point. For that point's nominal cost, Achievability and Delivery would share a common value. One would either select that particular solution or trade-off Achievability against Delivery until an acceptable balance of these characteristics is determined.

Another way of investigating risk is to contrast one of the delineated efficiency frontiers in Figure 3, specifically the one which is associated with the 3000 personnel constraint with a series of cost profiles arising from the optimization under the constraint that solutions are to have specified assurance levels of delivery for a common cost threshold. This is shown in Figure 6.

The top line in the Figure depicts the FOM associated with the range of costs considered in the optimization when assurance is unconstrained. The remaining lines in the Figure depict the best FOM achievable when cost assurance is constrained to be no less than the value indicated for the line in the legend. For example, the series of points in dark blue, representing 95% delivery assurance, initially tracks the efficient frontier, however these have been overlaid by the brown points on the frontier. Beyond a certain cost, these curves diverge with the constrained case tracing a curve distinctly below the efficient frontier. The same divergent pattern applies to the other cost assurance threshold curves except that they deviate from the frontier at higher nominal cost.

Accordingly, one observes that for lower cost values, the solutions associated with the efficient frontier, the line in brown, have a high assurance of being delivered at a cost below the threshold value. From the first kink in the brown line and beyond, the Figure indicates that solutions on the line have an assurance of delivery below 50%. At the point of intersection of the brown line and the purple line and beyond the assurance of delivery at the threshold cost is below 5%.

Noting the left hand side of the Figure but to the right of the first kink of the brown line, there are solutions at certain costs that can be delivered with a greater degree of assurance than those on the efficient frontier but whose Figure of Merit is distinctly below what could be delivered at that cost for a lesser degree of cost assurance. This indicates that if cost risk assurance becomes a significant factor in decision-making, then it may be necessary to select solutions for certain ranges of costs that deliver less FOM than otherwise. Accordingly the decision maker must assess whether the amount of FOM traded-off is worth the gain in assurance. This may require a finer analysis of the contents of solutions at comparable nominal cost to assess whether an adequate degree of closure for critical deficiencies are achieved.

The phenomenon at the right of the profile by the 5% line can be understood as follows. There appear to be sets of alternatives that can be selected at relatively small cost and are associated with relatively small financial risk. However, these can only be found in solutions provided higher cost, financially risky but otherwise more effective alternatives are displaced from nearby solutions. As several of these are discarded, one or more of these sets can be integrated into solutions that satisfy the 5% assurance threshold. If the cost of alternatives discarded is large enough, FOM synergies associated with the complete adoption of one such set take effect. This is

shown by the increase of FOM with cost in four sections along this curve. At higher costs, alternatives have to be selected that remain cost feasible but compromise the integrity of these aforementioned alternative sets. This results in those sections of the curve where FOM declines with nominal cost. For a particular range of nominal cost no combination of alternatives can be found that satisfies the assurance constraint although such combinations can be found at higher costs. In all, the phenomenon just described arises from the non-linear nature of the FOM calculation described in Annex A as well as the requirement that the personnel and cost assurance constraints be met.

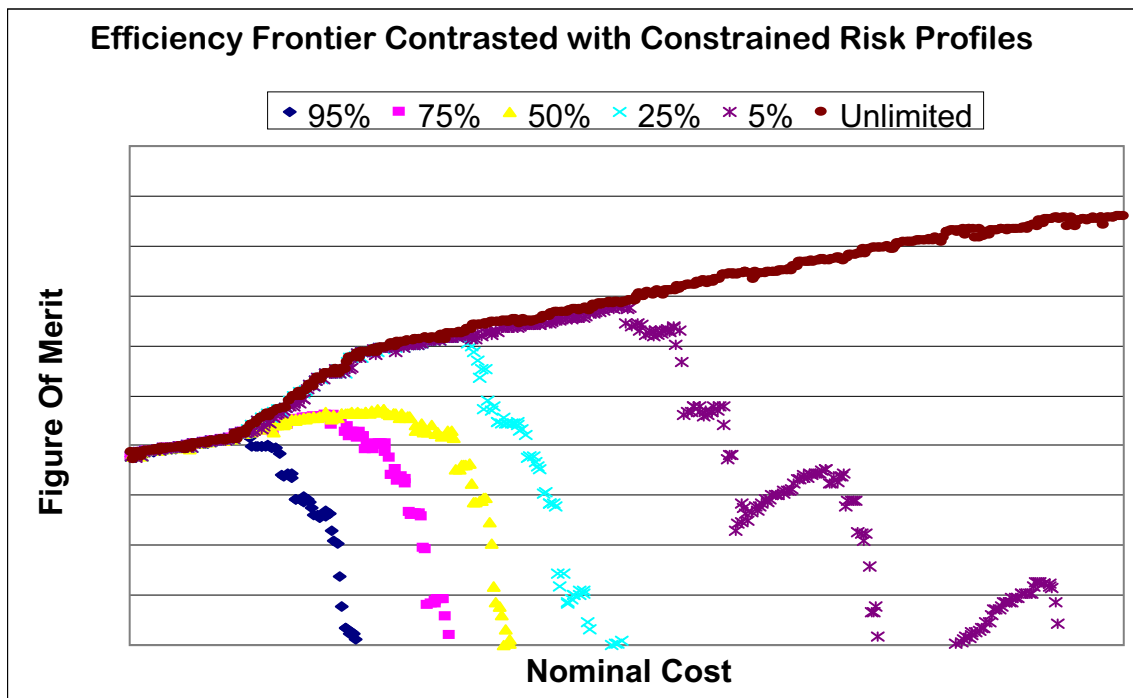


Figure 6: Efficiency Frontier Contrasted with Constrained Risk Profiles

3.4 Technical/Implementation Risk Analysis

The next Figure relates the technical and implementation risk associated with optimal solutions across the cost span. In this case, the scores given in the risk assessment were linearly scaled to a zero base, with a maximum score in all categories given a value 1. Note that the solutions on the efficient frontier show a small increasing drift from a starting score of .25 to somewhere between .27 and .28 on average. This means that solutions at higher cost levels accept higher technical risk. Nevertheless, the given score, when translated back to risk rating indicate that optimal combinations of alternatives are associated with an average implementation risk at the upper end of the low risk category throughout the range of nominal costs examined.

A more detailed analysis could be had by providing confidence risk range bands about each point identifying aspects of the range of scores associated with alternatives. Some of this was done in the SCR 1.0 final report [Reference 14].

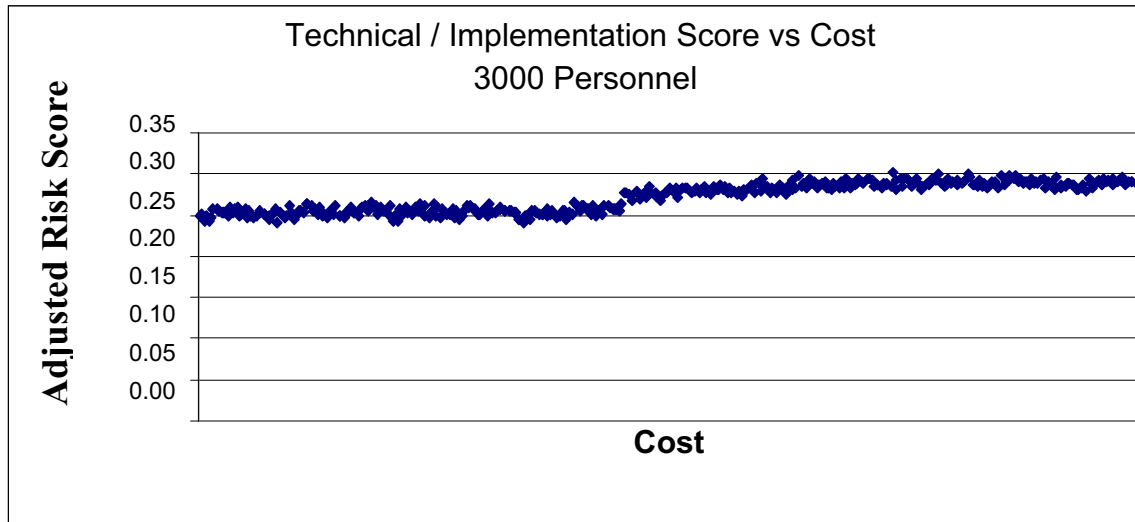


Figure 7: Technical Risk Accepted

3.5 Solution Level Analysis

The next chart emphasizes some of the characteristics that can be gleaned from point solutions. One of the necessary characteristics of a good solution should be the degree to which important issues are addressed. The following figure provides a representative plot of the importance of issues, shown as Raw Score along the horizontal axis, denoting the degree unfulfilled system deficiency with no alternatives selected and the degree of closure of this deficiency achieved by the solution along the vertical axis descending. This format is chosen to give an impression of a risk control chart. A properly managed portfolio should provide a high degree of closure for highly valued issues, with unresolved issues being of lesser importance. In contrast, if important issues remain unaddressed, one should find these located toward the top right of the figure. Now by the nature of the optimization, the links between issue closure of important items and those of lesser important items may come into play so that lower valued issues are addressed to ensure that higher valued issues are better addressed. This can account for the lower valued issues clustered toward the bottom left of the chart. Overall, for the point considered, most of the important issues, with raw score greater than 2.00 seem to be addressed. Note that this chart indicates the capability owner of the issue raised by colour coding.

It should also be noted that some points of the chart are superimposed by the nature of sharing common Raw Score and closure. As issues start being closed some of these superimposed points should become visible.

The real value of this chart is in relating point to point. In this case, the nature of change in the overall FOM for increased cost or personnel levels is detectible by noting the changes at the issue level. This provides the sponsor community with feedback as to where the heuristic solutions have been assigning resources for the greatest impact on FOM. Similarly, the chart format can be used to assess how higher assurance solutions compare with lower assurance solutions at a given solution cost if the FOM has to be compromised.



Figure 8: Issue Resolution

3.6 Ordering By FOM – Two Views

The next Figure compares two types of alternative ordering in a sample solution. In the first, alternatives in the solution are ordered in terms of the greatest contribution to solution FOM. Since solution FOM is comprised of the individual contributions of alternative contributions, these can be set in descending order to discover what in the solution contributes most. For the sample data in question this is shown along the horizontal axis. Another way of imposing a FOM ordering is to successively delete least significant contributors to the FOM, so as to retain as much FOM as possible in the remaining subset. When this order is imposed on the alternatives in a solution a different ordering arises. The order associated with the second method is shown along the vertical axis. The points shown in the chart compares the values of the rank positions for each issue of the sample problem.

Of note is the general correlation between the two orderings. Of greater significance though, is the presence in the fifth and sixth positions in the vertical ordering of items ranked comparatively low, 23rd and 33rd in the horizontal ordering. This means that these lesser significant items in the first instance must contribute to the overall solution as major enablers to resolve issue deficiency in the second. This has important implications to the precedence one should apply to implementing the projects that underlie these alternatives.

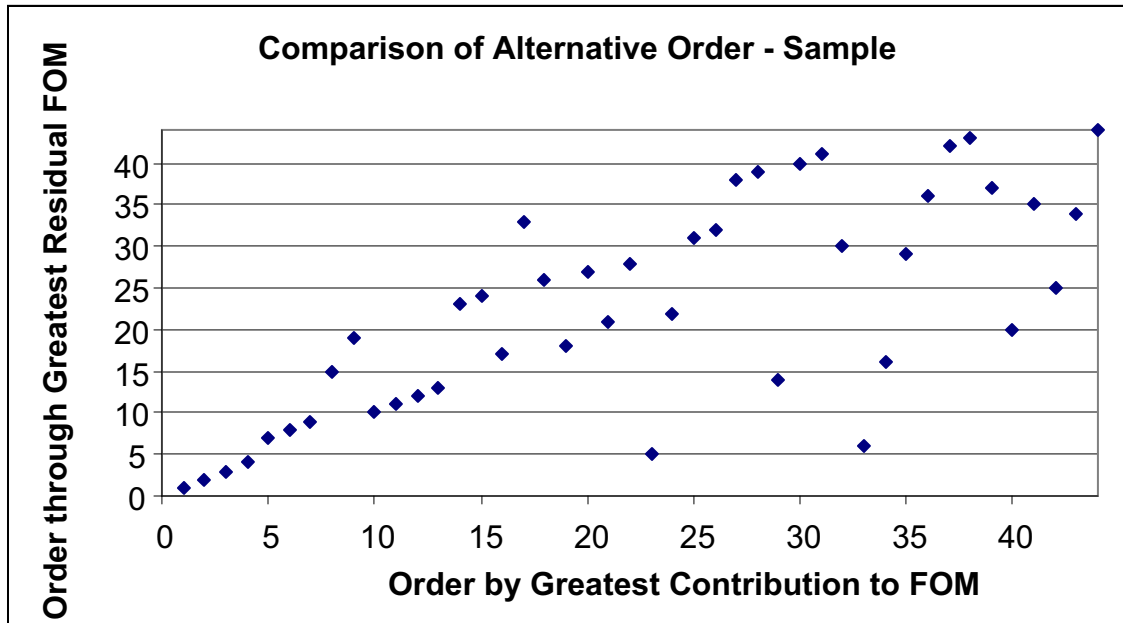


Figure 9: Stability of Alternative Orderings

3.7 Ordering Alternative Persistence

The next Table describes the persistence of alternatives with respect to the solutions in the vicinity of a solution point of interest. To allow this presentation to have as wide an audience as possible, the values shown the table and the issue labels bear no relation to actual outcomes of SCR 1.2 work. This is in contrast to the way results were presented in previous sections, in which preliminary SCR 1.2 information was used but with scale values suppressed.

The issues to be addressed are listed along the left hand column. The remaining columns correspond to alternatives available to addressing them. The percentage shown in each column is the frequency that the alternative is found among the solutions about a range of cost bounding a solution of interest. Each cell is given a colour background reflecting whether the alternative frequency exceeds 80% (Green), is between 50% and 80% (Orange) or lies below 50% (Yellow). This coding provides at a glance a convenient way of identifying persistent alternatives.

Against the Destroy 01 issue, it is noted that no alternative is selected among 58.5% of the solutions within the analysis range while the first alternative is selected 41.5%. This indicates that if the solution does use the first alternative, then it is not as robust an option choice as, for example, alternative 1 of Destroy 04 which appears 85.4% times in the analysis range.

Note as well that there is no guarantee that any alternative might be persistent about the cost range. For instance, none of the three alternatives in Detect 04 occurs with a frequency beyond 50%. Thus one might be indifferent to selecting any of these alternatives in a solution incorporating persistent choices and indeed select an alternative that might not be associated with optimality provided it improves the other performance metrics associated with the optimization.

A post analysis option is to sort the rows in this table in descending order of the maximum frequency found for alternatives. This imposes an order on the issues which reflects alternative persistence. Clearly one should attempt to implement to a greater degree those alternatives which show consistent application in the neighbourhood of a solution than those that don't. The decision whether to implement other than consistent alternatives should be made through a more careful analysis of solution characteristics. These may include whether the presence of the alternative in the solution provides better closure properties, or whether the resulting solution can be delivered within the problem's cost threshold with greater assurance.

Table 1: Illustration of Persistence of Alternatives in the Neighbourhood of a Solution

Issue	Alternatives					
	0	1	2	3	4	5
Destroy 01	58.5%	41.5%	0.0%	0.0%	0.0%	0.0%
Destroy 02	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Destroy 03	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%
Destroy 04	14.6%	85.4%	0.0%	0.0%	0.0%	0.0%
Destroy 05	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%
Decide 01	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Decide 02	70.7%	0.0%	12.2%	17.1%	0.0%	0.0%
Decide 03	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Decide 04	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%
Decide 05	0.0%	43.9%	56.1%	0.0%	0.0%	0.0%
Detect 01	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Detect 02	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%
Detect 03	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%
Detect 04	31.7%	46.3%	22.0%	0.0%	0.0%	0.0%
Detect 05	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Defend 01	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Defend 02	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Defend 03	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Defend 04	85.4%	14.6%	0.0%	0.0%	0.0%	0.0%
Defend 05	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Deliver 01	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%
Deliver 02	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Deliver 03	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%
Deliver 04	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%
Deliver 05	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%

3.8 Summary

Overall, this section has provided an overview of the some of the direct and post-analysis output made possible through the application developed to verify solution optimality. These do not exhaust the kinds of information that can be made available by exploiting available information. Moreover, there is much scope for the inclusion of further ways of analyzing solutions, many of which must be defined through dialogue with the sponsor about the kinds of metrics that need to be generated which help resolve the alternative precedence issue.

4 Discussion

4.1 Overview

The preceding application of the heuristic and its associated analytical features have demonstrated that either an @Risk mediated or even a native MSExcel simulation can provide a framework for rapid outline of the major characteristics of the efficient frontier for the kind of problems encountered in determining effective packages of alternatives for long-term capability planning. A trial applying the heuristic and the existing optimization procedure to the same problem was conducted in the summer of 2009. At the time the heuristic's grid did not cover the optimization range with the same degree of fidelity as the SCR 1.0 approach to the simulation. The set of efficient solutions defined by each simulation approach were then integrated with a view of producing a comprehensive efficient frontier for the problem. While the number of efficient solutions that came from the heuristic in the integrated solution was not as numerous as can be had using the previous optimization method – owing to the coarse grid used - a sufficient number of solutions uniquely determined by the heuristic, in the vicinity of 10%, appeared on the efficient frontier. If the number of grid elements covering the same range could have been increased, one could have expected a higher percentage of unique efficient solutions to be identified. Once again security of information precludes presentation of comparative results in other than the manner just described.

Insights gathered in tackling the optimization problem through the heuristic have been applied to reformulating the information architecture used to set up the problem using the SCR 1.0 approach. Preliminary evidence indicates that when these insights are applied run times in the other contexts, run times become comparable. This will be described in follow on work.

Unique to this application are representations of the technical risk associated with efficient solutions and costs associated with solution delivery at specified cost risk thresholds. Associated with specific solutions is a visualization of the degree to which issues in specific capabilities can be addressed. This is represented as a risk management plot in which the impact of the issue is rated against the achieved degree of closure. One can consider a solution effective if the higher valued issues are closed at a greater degree, while those issues not closed are of lesser significance.

The tool developed also can be used in other aspects of solution analysis. For example, the alternatives in a solution can be rated by the degree of contribution to the FOM achieved by each component. Another method of assigning precedence to the alternatives in a solution is through the persistence of the alternative along a range of neighbouring cost intervals. It is evident that an alternative that occurs in a greater percentage of the solutions in a range would likely be preferred over alternatives that occur less frequently.

The tool can support a rather sophisticated approach to determining cost associated with achieving the solution at specified risk. This is described in Annex D though it was not fully applied for SCR 1.2 work due to the nature of the solutions for cost range considered.

The heuristic has been linked to important post-analysis capabilities, the most significant of which involve the impact of cost sensitivity to the delivery of alternative packages. Through this application, the user can inspect the trade-off of mission capability with the need to deliver at cost or at specific personnel caps. Moreover, the application has taken the first steps to integrating cost risk as a constraining factor in solution delivery. This application could be considered the start of the development of the more comprehensive management system that Chief of Force Development requires to guide force development with due consideration of the multivalent risk environment he encounters.

The heuristic, originally reliant on @Risk but now fully implemented in native MSExcel, has thereby been made accessible to a wider community of analysts capable of using it. Thus the number of platforms that can conduct the sensitivity analysis that is so important for establishing credibility in Force Planning work has been expanded.

4.2 Directions For Future Development

4.2.1 Introduction

The described series of applications has achieved several analytical objectives of importance to SCR analysis. It has provided another optimization method for the Force Development alternative selection problem, one which has a broader base of user support within the Department. Instead of being restrained by the number of user licences for commercial software and the costs associated with specialized training in the product, SCR analysis can be conducted by any regularly trained military member familiar with MSExcel. This extends the capacity for conducting more extensive sensitivity analysis of the factors that contribute to overall FOM. A straightforward extension to the basic heuristic allowed the integration of personnel or cost assurance constraints as factors in the optimization process. Overall, the application has emulated analysis methods used in SCR 1.0 to almost the same degree of solution resolution but has extended them in order to develop a richer set of output products. Supplemental analyses and visualization are available to investigate relative persistence and importance of alternatives in solutions. Other analyses focus on solution deficiency closure or residual technological risk. Nonetheless there still exists room for improvement before the research invested into this series of applications can be considered complete – particularly with respect to cost and risk management and the scheduling of alternatives.

4.2.2 Enhancements for Contingency Management

With cost risk assessment now embedded in the optimization algorithm, the potential exists to provide insight into the establishment of solid financial contingencies to shepherd program management. In the discussion, a chart was shown that indicated that planning to utilize the complete funding to be made available for implementing a program of alternatives is fraught with a low likelihood of delivering the complete set of desired alternatives [Figure 5]. In contrast, working with a lower funding level for planning purposes can result in alternative sets that deliver over 90% of achievable FOM at over 95% assurance. The question arises: What becomes of the remainder of the planning space, the difference between the nominal cost and the desirable cost threshold?

Since the cost sensitivity analysis is strongly dependent on the cost uncertainty bounds, it is important for alternative developers to generate their cost estimates with as narrow a bound as is possible. This would certainly reduce any margin to be kept in reserves.

Another suggestion is to allocate the difference between nominal cost and expected cost back to the alternative developers/managers as an internal contingency along with a rather firm commitment on their part that they (together with the program management staff designated to deliver alternatives for the domain) not exceed the sum of the expected costs of the selected alternatives. Since, in the sample data from which the cost sensitivity analysis of Figure 4 is generated, the difference between nominal and expected values proves to be in the order of 30 - 40%, this difference would provide a generous contingency with which to assure that most alternatives could be delivered within the resources assigned to program and capability management. The remaining slack space should be kept as centrally managed reserve to cover those hopefully rare instances that critical elements of the program cannot be delivered. This should be considered the funding pool of last resort within the Department.

The cost sensitivity assumed that no unforeseen changes in cost occur during the planning period beyond what was captured by the sensitivity bounds. Should significant changes to cost occur that were outside the planning scope of the Department, it would appear justifiable to request remedial funding from Treasury Board or at least signify the need for a review of priorities.

The preceding discussion provides an outline of the kind of guidance that could be provided to Chief of Program and Chief of Force Development in the realm of risk management. There is a need for dialogue with these Principals to communicate the potential of the application to address their specific risk management requirements either in the form it has been developed or in an expanded version of the application to cater to a greater degree of program resolution.

4.2.3 Alternative Scheduling

A critical concern for the CF is the scheduling of project/program initiatives. Since these are linked to alternatives, the problem here is to select of alternatives and specify their respective implementation timeframes. By extending the problem range solved by the heuristic it may be possible to provide greater insight into resolving the former.

Particular features needing attention before such an extension of the heuristic is possible include:

- a) the knowledge of the earliest implementation date associated with alternatives;
- b) alternative lifetimes;
- c) a profile of expenditure constraints, which could be given explicitly in terms of cost limit per year, or multiplicative factors that shape end cost and
- d) a rating scale for the criticality of the deficiency over time.

With all this information in place, it may be possible to generalize the FOM calculation so that optimization is phrased in terms of the minimization of residual deficiency-years. On one hand, annual cost constraints would limit how many alternatives could be acquired, restricting them to

With such an extended optimization scheme in place, a natural ordering of affordable alternatives can be generated, supporting, or even superseding current manual and qualitative system for establishing precedence. This could eliminate some of the contention that could arise among capability management, programming and force generation communities in establishing principles for a schedule.

These two recommended development strands that have been outlined are seen to be complementary. For example, with a schedule in place, it would be possible to develop a program of phased contingency release so that it does not get expended too rapidly during the planning horizon. Conversely, if through risk management it is discovered that the program of alternative delivery is no longer tenable, then the scheduling analysis can be redone with a view of best delivery of deficiency reduction for a range of reduced budgets.

4.2.4 Longer Term Developments

In the longer term, there will be a need to extend the basic heuristic to support the analysis of alternatives for deficiencies in the Generate and Corporate Capability Domains as well as divestment alternatives in all capability domains so as to place these on the same analysis footing as the alternatives in the remaining capability domains. At first glance, the approach developed thus far could be extended to tackle both of these issues, with the proviso that the resulting FOM calculation, when supported by validated available data, will be decidedly more complex. Expanding the domain space to cater to these requirements is conceptually feasible but solving for efficient alternatives may take longer than what is currently the case.

The heuristic as presented caters to providing guidance for long-range capital requirements. As such, it may fulfil the CFD objective to provide this information to the Chief of Program (C Prog). However, there is also a need to provide C Prog with alternative selection/sequencing guidance for programming decisions to be implemented within the next five years.

What is envisioned here is a project-centred rather than an alternative-centred application. First, the costs associated with a project would be more finely defined and managed. Separate cost profiles corresponding to each of the development, acquisition, personnel, operations, training and maintenance activities associated with the project would have to be specified in time. Next implementation dependencies between projects would have to be recognized so that critical enabling projects are brought on line when they are needed for project implementation. The projects selected and the time at which they are implemented must also be linked to an operational measure of effectiveness so that performance can be rated between different programme/schedule configurations. Finally resource constraints in terms of cost caps per cost category have to be identified. The outcome of an optimization conducted at this level of detail would be a way of prioritizing projects in the short-term – with optimized schedules associated with ranges of costs within each budgetary category.

Recognizing that budget allocations and cost estimates come with some degree of uncertainty, one could imagine a situation when commitments to servicing short term requirements might impinge upon the funding needed in the long-term. If this should occur, one should note that this would affect the optimal solution for the long-term plan. Nevertheless, this does not mean that the long-term optimization needs to be redone, unless of course priorities have changed as well. The revised long-term plan can be identified by using the results of the previous long-term optimization at the long-term funding level implied by the short-term funding level adjustments.

As a consequence, a two phased optimization approach of the kind described would go a long way to provide a project and capability development management framework that can be both responsive to operational need and budgetary uncertainty.

In summary, the series of applications described in this document has the potential for delivering insight into important Force Development issues, many of which have not had their due attention in the past. It will no longer require specialty software or training to conduct either the initial or sensitivity analyses associated with generating a Strategic Capability Roadmap. Of course to do these analyses well requires the timely provision of complete and thoroughly validated source data – with this being more important to the validity of outcomes than optimization processing. SPORT seeks the support of CFD and CORA to continue proceeding along the outlined development path to realize the tool's potential in balance of investment decision making.

References

- [1] Vice Chief of Defence Staff, *Strategic Capability Planning for the Canadian Forces*, Department of National Defence, 13 June 2000.
- [2] TTCP Joint Systems Analysis Group Technical Panel 3, *Guide to Capability-Based Planning*, the Technical Cooperation Program, 2004.
- [3] Burton, R., G.L. Christopher, P. Chouinard, L. Kerzner and Maj K. Simonds, *Procedures, Processes and Tools of Capability-Based Planning: An Outline for a Canadian Approach*, DRDC CORA TR 2005-35, November 2005.
- [4] Vice Chief of Defence Staff, *Initiating Directive for the Production of the Strategic Capability Roadmap (SCR)*, Department of National Defence Letter, 195-0-1 (D Cap P), 25 October 2007.
- [5] Kerzner, L.F., *An Introduction to Capability Based Planning (U)*, DRDC CORA TM 2006-18, June 2006.
- [6] Solomon, Dr B., and L.F. Kerzner, *An Overview of the Strategic Costing Model*, presentation to DRDC CORA Annual Meeting, October 2006.
- [7] Chouinard, P and I. Wood, *The Department of National Defence Strategic Cost Model: Development*, DRDC CORA Technical Report 2007-14, September 2007.
- [8] Solomon, Binyam, Paul Chouinard, and Leonard Kerzner, *The Department of National Defence Strategic Cost Model Volume 2: Theory and Empirics*, DRDC CORA Technical Report 2008-03, October 2008
- [9] Blakeney, D.E., A.P. Billyard, L.F. Kerzner, B. Solomon, and P.L. Chouinard, *Operational Research Tools Supporting the Force Development Process for the Canadian Forces*, Information and Security Journal, Volume 23 No. 1, 2008.
- [10] Materiel Knowledge Network, *Project Management Body of Knowledge Guide*, Project Management Institute, 2004.
- [11] *Recommended Practice: Cost Estimate Classification System*, Cost Engineering Volume 39, Number 4, Association for the Advancement of Cost Engineering International, 1997.
- [12] Emond, E.J., *Cost Estimation using Three-Point Estimates*, ORAE Project Report PR303, May 1985.
- [13] Vincent, P., *A Methodology for Quick and Easy Risk Assessment and Cost Estimation*, Directorate of Logistics Analysis Research Note 93/3, April 1993.

- [14] Christopher, G., Blakeney, D., Petryk, Dr. B., Taylor, Dr. B., Kerzner, L., Fong, V., Ball, M., Beard, Dr. A., *Strategic Capability Roadmap Version 1.0 Analytic Framework (U)*, DRDC CORA Technical Report 2009 – 13, December 2009.
- [15] Emond, E. J., *Developments in the Analysis of Rankings in Operational Research (U)*, DRDC CORA Technical Report 2007 – 37, December 2006.

Annex A Capability Optimization in SCR 1.0

A.1 Optimization Problem

Identifying the best set of capability alternatives is fundamentally an optimization problem. In the simplest terms, the challenge is to identify the set of alternatives that provide the greatest resolution of the deficiencies (most capability) within the available funding limit or other constraining features as measured at the end of a specific planning horizon. This optimal set of capability alternatives would provide the best value for money.

In preparation for the optimization, all critical information must be assembled. All possible alternatives must be known. Each alternative proposed to remedy a deficiency must have its Equivalent Annual Cost (EAC) evaluated. EAC corresponds to the ratio of total cost to the Department, including prorated costs for such enabling activities as research and development, training, basing and equipment support, associated with the alternative, and as well as annualized direct costs incurred acquiring, staffing, using, maintaining and disposing the alternative over its expected life. Also required are the degree of deficiency closure and the impact of the alternative (as rated by its deficiency's mission value score). Finally information had to be provided to capture the interdependence between alternative potential in the presence of residual deficiencies across the problem space. Other information required includes ratings of cost, schedule and technical risk and Objective Force compliance. These were used in ancillary calculations that monitored secondary characteristics of found solutions.

The basic optimization problem was to find a set of alternatives that would deliver the highest aggregate capability at lowest risk with greatest Objective Force compliance for a given total EAC. For SCR 1.0, the solution could be a set of up to 87 alternatives (one alternative for every deficiency) chosen from among the 200-plus alternatives provided.

Consider first the scale of the optimization problem involved. As the problem ultimately became formulated, the potential solution space not considering constraints exceeded 2^{60} possibilities. This came about through certain specifications of alternatives for certain deficiencies that were guaranteed to be in the solution for policy reasons. This number of possibilities compares to the distance in inches that light can travel in a day! This can be reduced somewhat by applying further problem constraints. Nonetheless, a solution of this problem through any direct analytic approach is computationally difficult. The complexity in finding an optimal solution is also magnified by the formulation of the optimization objective. The presence or absence of an alternative in a solution affects the Figure of Merit (FOM) in two ways. First there is a direct contribution made by the alternative to closing its associated deficiency⁴. This is defined to be an initial closure of the deficiency. Next there is the indirect effect that is achieved in the potential to close other deficiencies by not having this alternative complete the closure of its deficiency. The combination of these factors in the development of the optimization FOM is represented as a non-linear relationship in the decision variables which excludes this problem from solution by

⁴ In SCR work completed so far, an alternative is provided to close in some degree only one deficiency. However through the interaction of deficiencies, the degree of closure of an enabling deficiency can affect the degree of closure possible by an alternative for another deficiency.

standard techniques. The combination of the initial closure and the dependency impact generates the final closure for the deficiency.

For example, a fighter aircraft initially closes a deficiency for North American air defence. However, the fighter's potential for the role is limited if there is no air to air refuelling capability provided. Therefore the final closure calculation involves the impact of the absence or presence of this capability in the solution,

The details of the FOM calculation follow.

Let C_{ij} be the closure of Deficiency j by Alternative i . As in previous SCR work [Ref 14], alternatives can be associated with one deficiency so that C_{ij} can be determined without reference to an alternative's impact on other deficiencies. The final closure of Deficiency j by Alternative i with dependence included, denoted by C_{ij}' , is defined as

$$C_{ij}' = C_{ij} \prod_{d=1, d \neq j}^{Nd} (1 - (D_{id} * (1 - \text{MAX}_{k=1}^A (C_{kd}))))$$

where and the product accounts for the potential degradation in closure when the Alternative i is dependent on the degree of closure of collateral Deficiency d . Nd is the total number of deficiencies. A is the number of alternatives in the current solution. D_{id} measured the strength of Alternative i 's dependency on the capability associated with Deficiency d and C_{kd} was the closure of Deficiency d provided the selection of one its alternatives, specifically Alternative k . C_{ij} and C_{kd} took on values in the range $[0,1]$ (0 when there was no alternative selected to address the deficiency, to 1 when an alternative was selected which closed the deficiency to the maximum extent possible), while the D_{id} was set to 0 if there was no dependency on the enabling capability, 0.5 when there was weak dependency, and 0.95 when the dependency on the enabler was strong. Qualitatively, this meant the potential closure of an alternative, C_{ij}' , was unaffected when there were no dependencies or when the enabling capabilities were fully provided. However, reductions to the potential closure accumulated multiplicatively with the extent to which deficiency remained unclosed in its enabling capabilities.

The optimization objective Figure of Merit (FOM) is a dimensionless variable that captures the overall value of a solution on a continuous numerical scale. It is calculated as the sum of the value of each deficiency, through its mission value score, adjusted by the degree it is closed by an alternative, as shown in the equation below. Another factor to note is that an alternative selected for a particular capability will never close more than the available magnitude of the deficiency when present.

$$F_s = \sum_{j=1}^A \sum_{i=1}^{Ni} S_j C'_{ij} A_{si}$$

where: F_s is the FOM of Solution S ;

S_j is the Mission Value Score of Deficiency j ⁵;

C'_{ij} is the final closure value of Deficiency j by Alternative i with dependence included;

A_{si} has the value of one if Alternative i is in Solution s , otherwise A_{si} equals zero;

Nd is the total number of deficiencies; and

Ni is the total number of alternatives.

The FOM shown above describes a calculation that depends on a single set of mission value scores. For SCR work two sets of these scores, closure rates and dependency values were retained corresponding to International and Domestic/Continental operations respectively. The final FOM for the analysis is a weighted average of the FOMs associated with these roles.

A.2 Optimization Approach

For SCR 1.0, an optimization tool was built using the Phoenix ModelCenter 7.1 programming software suite from Phoenix Integration Inc. The Phoenix ModelCenter suite (hereafter simply referred to as Phoenix) provides a proprietary optimization engine, Darwin, that conducts its search for an optimum based upon a genetic algorithm. The optimization engine can be linked to other software tools to provide data or evaluate options. In the case of the optimization for the SCR, Microsoft Excel[®] (MSEExcel) spreadsheets were utilized to provide the input data on the capability alternatives and evaluate the value of potential solutions.

For the SCR work using Phoenix an initial set of 120 potential solutions was generated to seed the model. Each solution contains an entry for each deficiency. These entries identify which alternative, if any, has been selected for its associated deficiency. Each deficiency can have at most one non-nil alternative selected and none should it be decided that the deficiency remain entirely unresolved. The overall value, referred to as the Figure of Merit (FOM), of each solution is calculated as the sum of the value of each deficiency adjusted by the degree it is closed by an alternative. The final FOM for the solution is a weighted sum of the adjusted deficiency values for International and Domestic/Continental operations⁶. As well as the FOM, other solution attributes such as cost, total personnel, risk and Objective Force compliance are also determined for each solution. Comparing all these solution attributes against ordinal scales and thresholds allows each solution to be ranked. The next generation of solutions is formed by cloning solutions, evolving solutions and mutating solutions. In cloning solutions a number of the best

⁵ Mission value score associated with a deficiency is derived measure that is taken from scores captured by the CATCAM model. Specifics about the CATCAM model and the score calculation are found in Reference 14.

⁶ Recall that in the data collected for each capability alternative separate closure values were specified for each of Domestic/Continental and International operations.

solutions (16 in the case of the SCR) are carried over into the next generation. To evolve new solutions, the genetic algorithm then takes the top-ranked (best) solutions and randomly selects alternatives identified in these solutions to create new solutions for the next generation. Finally in mutating solutions, some existing solutions are chosen and a number of the identified alternatives are randomly replaced with ones selected from the broader pool. In this way a new generation of solutions is populated from the best individual solutions of the previous generations with some random variability, mimicking the biological, evolutionary process of nature.

This process is repeated multiple times to produce many new generations of solutions. With each generation a measure of improvement over the previous generation is calculated. Propagation of new generations of solutions continues until the degree of improvement drops below a set threshold, whereupon the process stops. For the SCR 1.0, the search stopped if fewer than three better solutions were produced in 25 generations. This process results in an intensive, directed search of the solution space in an attempt to locate the best (highest FOM (deficiency closure), highest Objective Force compliance, lowest risk) solution within the set constraints of cost and personnel. Because this procedure uses a heuristic algorithm and terminates when very little improvement is being obtained, there is no guarantee that the best solutions have been located. However, given the intensity of the search in the vicinity of the best solution, there is a very high probability that very good solutions (possibly including the best solution) will be found. For SCR 1.0, 35 to 55 thousand solutions were generated for each combination of constraints.

While it did provide the required high quality solutions, a noted drawback of the Phoenix approach was the limited ability then available to ordinarily trained military staff officers to perform the optimization. Since SCR 1.0 was an innovative activity, there was little partner concern with the degree of Operational Research support dedicated to the effort, both in terms of personnel and the specialty software required. However, future SCR activity will be expected to be supported by regular trained CFD Military personnel with minimal OR involvement.

Two issues required immediate resolution in order to address this drawback: the limited number of licensed copies of the Phoenix software available to DND; and the processing time to generate a set of optimal alternative packages for the cost range of interest. The first limited the number of platforms that could be dedicated to conduct the optimization. The second also constrained the number of runs that could be performed during the extremely limited timeframe that was made available for the analysis. Taken together these issues limited the scope for the sensitivity analysis that could have been performed for SCR 1.0. Furthermore, there had been no independent mechanism for verifying the optimality properties of proposed sets of solutions. The conduct of a parallel analysis using a different optimization approach would have enhanced the credibility of the optimal solution found by the SCR analysis team.

Annex B About @Risk

B.1 Risk Analysis

Traditionally, analyses combine single "point" estimates of a model's variables to predict a single result. This is the standard MSeExcel model — a spreadsheet with a single estimate of results derived from cell contents that contain a single value. Estimates of model variables must be used because the values which actually will occur are not known with certainty. In reality, however, many things just don't turn out the way that as planned. In the modelling for FOM and cost assessment, these include cost estimates, deficiency closure, mission value scores and project implementation dates. Perhaps one might be too conservative with some estimates and too optimistic with others. The combined errors in each estimate often lead to a real-life result that is significantly different from the estimated result. The decisions made based on "expected" result might be the wrong decision, and this might have been avoided if a more complete picture of all possible outcomes were available. Business decisions, technical decisions, scientific decisions ... all use estimates and assumptions.

With software such as @RISK (pronounced "at risk"), Crystal Views and parts of the Phoenix Integration tools one can explicitly include the uncertainty present in estimates to generate results that show all possible outcomes. The focus of the following discussion is @RISK, with the other cited tools applied the same way and generating similar outcomes.

@RISK is a system which brings risk analysis techniques to the Microsoft Excel spreadsheet package. System modellers are the best judges how to represent the interaction of components of their systems. They do this by building point models for the systems under investigating using MSeExcel. They are also aware of the uncertainty associated with their input data or of risk associations tied to intermediate values. @RISK facilitates the design of risk analysis models as a straightforward extension of MSeExcel models.

Applying @RISK to a point model transforms it into a simulation which combines all the uncertainties identified in the modelling situation into output uncertainties. The user includes what is known about the uncertainty associated with a variable, including its full range of possible values and some measure of likelihood of occurrence for each possible value. @RISK uses all this information, along with the Excel model, to generate a wide variety of possible outcomes. This has the same effect of running hundreds or thousands of "what-if" cases all at once! Then @RISK visualization features are applied to identify the range of what could happen across simulated cases. All this added information might seemingly complicate decision making, but in fact, one of simulation's greatest strengths as a modeling paradigm is its power of communication. For the most applications @RISK generates results in an easily understood graphical format that is easily explained to others, although the user could access individual simulation results as the need arises.

As an "add-in" to Microsoft Excel, @RISK links directly to Excel to add Risk Analysis capabilities. The @RISK system provides all the necessary tools for setting up, executing and viewing the results of Risk Analyses. @RISK works in the familiar Excel style menus and functions. @RISK allows the user to define uncertain cell values in Excel as probability

distributions using functions. @RISK adds over thirty new functions to the Excel function set, each of which allows users to specify a different distribution type for cell values. Distribution functions can be added to any number of cells and formulas throughout model worksheets and can include arguments which are cell references and expressions — allowing extremely sophisticated specification of uncertainty. To help assign distributions to uncertain values, @RISK includes a graphical pop-up window where distributions can be previewed and added to formulas.

@RISK has sophisticated capabilities for specifying and executing simulations of Excel models. Both Monte Carlo and Latin Hypercube sampling techniques are supported, and distributions of possible results may be generated for any cell or range of cells in a spreadsheet model. Both simulation options and the selection of model outputs are entered with Windows style menus, dialog boxes and use of the mouse. High resolution graphics are used to present the output distributions from @RISK simulations. Histograms, cumulative curves, and summary graphs for cell ranges all lead to a powerful presentation of results. And all graphs may be displayed in Excel for further enhancement and hard copy. An essentially unlimited number of output distributions may be generated from a single simulation — allowing for the analysis of even the largest and most complex spreadsheets!

The product comes with extensive analysis tools to determine what parameters are the most important for the simulation and to assess their impact on model processes and outcomes. These tools strongly resemble similar functionality found in Phoenix Integration.

Like MSeExcel Solver and some Phoenix Integration functionality, one can perform goal seek analysis using @Risk although this would not be efficient for this purpose alone; it has been included for this purpose to permit conduct of stochastic optimization, a facility not available in either of the previously mentioned products. It is still an open question whether an Integration tool can control @Risk enhanced models to give a richer framework to conduct stochastic optimization.

B.2 Previous Use of @Risk in SCR 1.0 Analysis⁷

A key consideration and major concern of DND senior management was the financial risk associated with the recommended optimal set of capability alternatives for the SCR. Recall that each alternative has a defined level of cost uncertainty identified with it. In the optimization process conducted for SCR 1.0 cost uncertainty was not included directly as part of the optimization. The allowance for cost uncertainty was established in a committee setting. The final set of optimal solutions, mapped against FOM and cost, were presented to the Joint Capability Requirements Board with a request for direction on the size of the financial buffer that should be reserved to deal with potential cost over-runs due to the uncertainty in the cost estimates of the

⁷ The majority of this section has been taken from a similar section in the recently published Reference 14. While the author did not write that section or perform the end analyses that were presented, he did suggest the tool, conducted several of the early cost sensitivity analyses and prototyped several of the visualizations presented therein.

alternatives. To assist with this decision, analytic results in the form of probability estimates of the likelihood of exceeding the program funding limit were presented to the Board.

Cost uncertainty for each capability alternative was identified using the three-level scale shown in the following. The level of cost uncertainty was determined primarily as a function of the phase the planning was at in the project development process. The cost uncertainty ranges were defined in accordance with accepted project management practices described in the Project Management Body of Knowledge [9] and the Cost Estimate Classification System[11]. Greatest uncertainty was associated with alternatives that were basically new ideas for which no detailed definition work had been completed. Here the cost of the capability alternative was estimated as possibly reaching as high as three times the cost estimate provided and could be as low as half the estimated cost. At the other end of the scale, where planning for the alternative was at the project identification stage, alternative cost could be 75% greater than the estimate to as low as 25% less than the estimate. These cost boundaries established defined ranges for the costs of each capability alternative.

Table 2: Cost Uncertainty Scale

	Project	Concept/Feasibility	Minimal
Level of Definition	Project exists at identification phase or beyond	Detailed feasibility or concept studies have been conducted	No detailed or specific studies have been conducted.
Expected Cost Uncertainty	Project ROM Cost accuracy: +75%/-25% or better dependent on project phase.	Concept ROM Cost accuracy: ~+100%/-50%	Placeholder ROM Cost accuracy: +200% / – 50%.

To assess the probability of an SCR solution exceeding the funding limits, the cost estimates for the alternatives were treated as probability distributions. The provided cost estimate was taken to be the most likely cost, while the cost uncertainty defined the minimum and maximum values of the cost range. To ensure a robust perspective of the financial risk, three separate analyses were performed using different probability models to represent the alternatives' cost distribution. In one iteration, alternative cost was modelled as a standard triangular distribution, shown at the top of Figure 10. Anecdotal evidence suggested that projects rarely complete under budget. To investigate this perspective, a second version of the triangular distribution was used, with the minimum and most likely costs being the same (i.e. the cost estimate of the alternative). This version of the cost distribution was referred to as the pessimistic triangular distribution and is shown in the middle chart of Figure 10. The last distribution examined was one where alternatives' cost was modelled as a uniform distribution, i.e. all possible costs within the defined range are equally likely. This distribution is shown at the bottom of Figure 10.

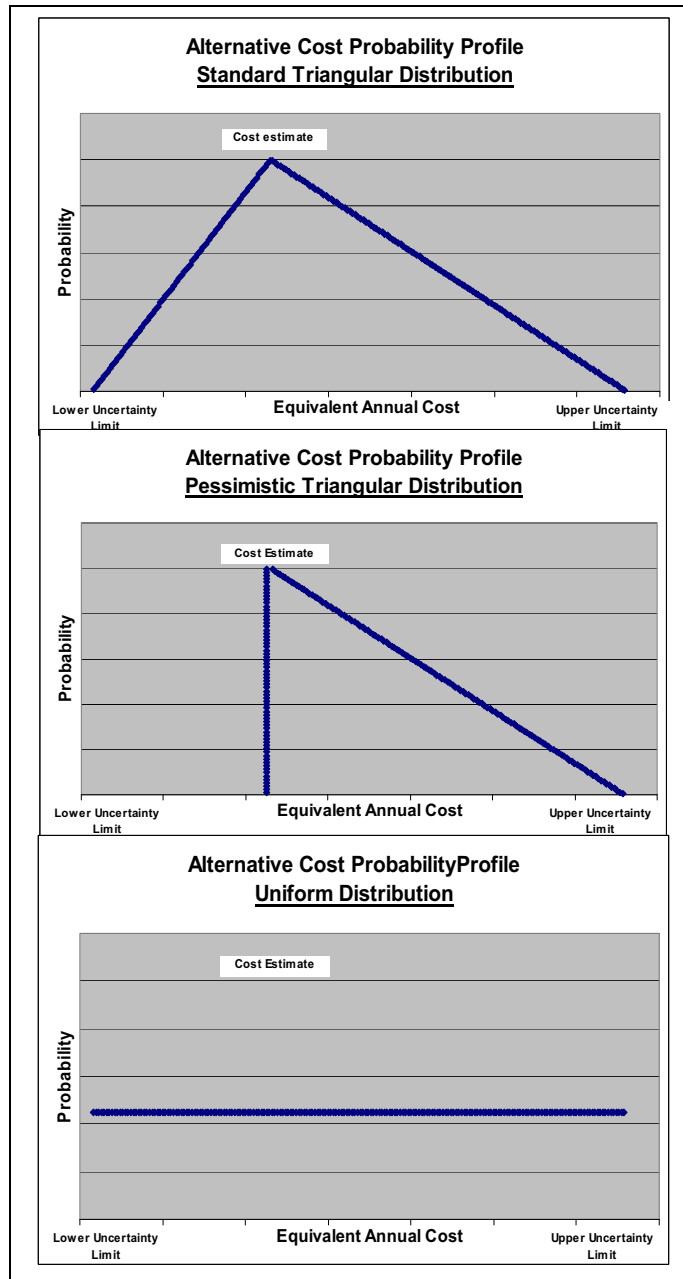


Figure 10: Alternative Cost Probability Models

To perform the financial risk assessment, a program-cost model was developed using @RISK. In this model, the cost of each capability alternative in the solution is determined by sampling from the associated cost distribution. The total cost of the entire program is then calculated by summing all the individual alternative costs. Performing this procedure 1000 times generates a probability distribution of total program cost, which could be compared to the SCR funding limit. Figure 11 presents some indicative results of the financial risk assessment at the program level, based on the general triangular distribution as illustrated in the top chart in Figure 10. From the Figure, solution 25778 has a financial risk level of 80 percent, which means there is an 80 percent probability of the entire program (set of alternatives) exceeding the funding limit. The solutions shown are optimal solutions in sequential order moving down the efficiency frontier away from the funding (cost) constraint. Because the level of cost uncertainty is unique to each capability alternative, a solution further away from the funding limit does not necessarily have less financial risk than a closer solution. This situation occurs for the first two solutions (25778 and 22652) shown in Figure 11.

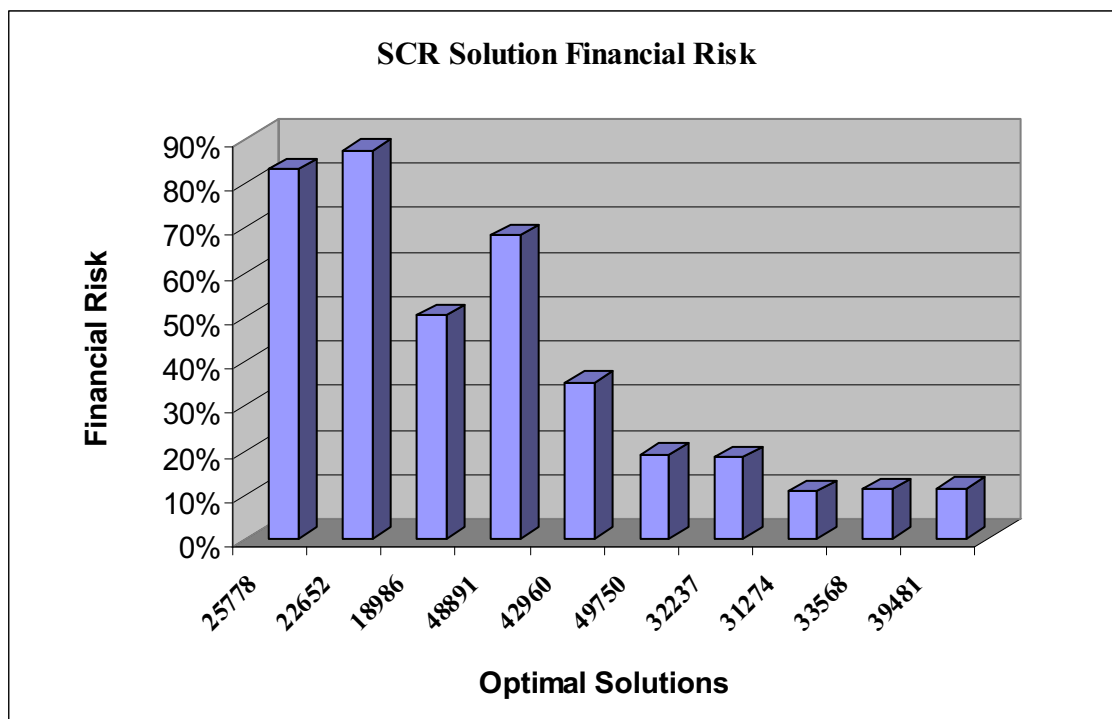


Figure 11: SCR Solution Financial Risk

To better understand the degree to which a program could be at risk from a costing perspective, a cumulative risk profile was prepared based on the individual cost contributions of the capability alternatives in the solution. As the cumulative risk profile is a function of the order in which alternatives are added, the alternatives in the solutions were first placed in a sequence based on best value for money (mission-value score (see Footnote 5) divided by component cost), allowing the most cost effective components to be scheduled first.. The assumption here is that one would want to minimize the risk for the alternatives that offer the best cost-benefit values. Figure 12 presents the cumulative risk profile for solution 25778 for the three different alternative cost

distributions. The profile shown in the Figure actually displays the inverse of the risk profile, namely the cumulative probability that the program can be achieved within the defined funding limit. For this solution there is no risk for the vast majority of alternatives. Irrespective of the cost distribution, only the last eight alternatives are at risk from a financial perspective. The degree of risk varies according to the cost distribution used.

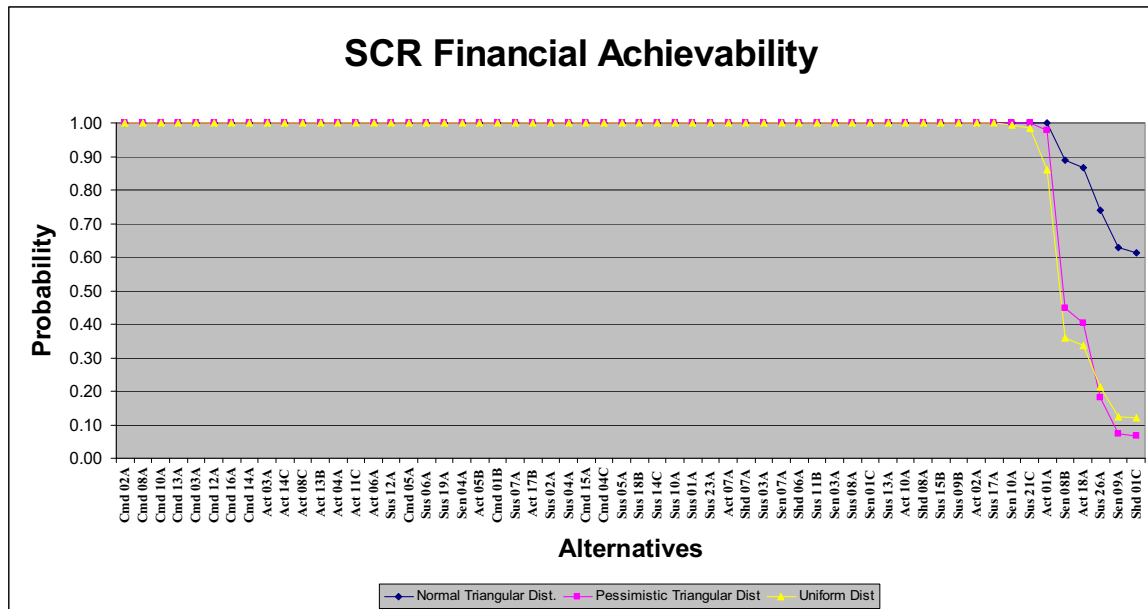


Figure 12: SCR Solution Cumulative Risk Profile

B.3 Use of @Risk in the Optimization Heuristic

The optimization heuristic described in this paper was originally designed to use minimal @Risk functionality. The purpose of the heuristic was not the collection of summary statistics but rather to build a systematic way of identifying solutions with optimal properties within the context of a random search. In effect all that @Risk provides are simulation management activities and a few built in functions to simplify decision-making about trial candidate solution selection, rates of crossing, span of crossing and mutation rates. Indeed, it was recognized early on that it may be possible to develop the heuristic so that @Risk would no longer be required. All that was done was to use existing MSeExcel functionality to emulate the outcome of random selection from discrete, uniform and binomial distributions. The original heuristic used the RiskMean function to calculate means of monitoring variables for simulations conducted in batches. Instead, the required calculation was reconfigured to perform inside an MSeExcel macro. This macro also applies simulation controls that were now embedded into the application that had been formerly part the @Risk overlay. The removal of specific @Risk functionality seems to have compromised processing time. Indeed, demonstration examples indicate a reduction of 20% in processing speed when the change to a pure MSeExcel model was made. This can be considered acceptable since the newer version of heuristic could be run by anyone in the Force Development community to

generate optimization results for cases of interest with only a small degree of developer-provided training.

As for program risk associated with solutions within the efficient frontier discussed in the previous section, a very different approach to obtaining these estimates was considered. Early on in the heuristic design, it was envisioned that cost assurance could be one of the constraints that could affect the development of optimal solutions. Rather than building a two-tier risk simulation, that for each solution that is found feasible and improves on an interim solution there is of necessity another @Risk simulation that must be run within the first to assess its cost risk characteristics, it was found to be far more computationally efficient to apply an approximate analytical means of calculating this risk. This approach is described in much greater detail in Annex D. As shown there, for solutions with a greater number of alternatives, this approximation was of sufficient accuracy to reproduce the characteristic features shown in Figures 11 and 12 with the exception that only the standard triangular case was evaluated. Future work on the heuristic may incorporate assessments made based on uniform distributions of cost as well as the triangular distribution already considered. Indeed, much of the theoretical framework remains the same regardless of distribution but the evaluation of the moments about zero for the uniform distribution takes a much simpler form.

This page intentionally left blank.

Annex C Technical Details

C.1 Introduction

The discussion in the main body provides an overview of the processes that underlie the set of applications that have been constructed to facilitate Capability Integration. It has avoided some of the internal technical issues that may be of interest to future users and developers of the application. This Annex provides some of this necessary background.

C.2 Grid Element Selection

The application supports a very rich array of methods to specify the analysis range and the degree of granularity within that range. The user is given the option to specify all or part of the complete grid element set as limiters to selection range for candidate generators. This can prove useful when, after a certain large number of iterations, it is discovered that a certain range of costs do not have a large number of efficient solutions. In this case, one can limit the selection process to this range so that its update is greatly favoured in contrast with the remaining analysis range.

C.3 Grid Element Selection Bias

The heuristic has been constructed so that grid element selection can be biased in favour of those grid elements with solutions not on the efficient frontier. This is accomplished by adding an increment to the pre-existing grid element weight. The increment is set to be large, say 5 times the default weight, when the number of efficient solutions covers a small fraction of the cost span and decreases to zero as this fraction increases. In practice, once the number of solutions exceeds 50%, the bias is practically zero. This is also subject to user control between simulation batches.

C.4 Mutation

Mutation is considered at each point having alternatives through a random decision driven by a binomial distribution of sample size 1 with mean ρ . This mean is chosen simultaneously with the grid element defining the candidate progenitor, implying that the mutation rate for candidate definition is constant across all mutation loci. Specifically, ρ is taken from a uniform random distribution ranging from 0 to 0.5. It should be understood that that a ρ value of 0 means that no mutation occurs whereas a ρ value of 0.5 means that approximately half the elements of the solution will be altered. On the average, once a mutation path is taken, approximately 25% of the contents of mutation loci will be altered.

Mutation rate can be altered by the application user between simulation batches.

Alternative selection during mutation is taken from a modified discrete distribution. During optimization set-up, the user has to identify for each deficiency which alternatives will be considered. There are six spaces per deficiency row in an intermediate table in which the user can

assign an indicator value in which the user can define which deficiencies are active. The first space corresponds to the null alternative, that is the one associated with not selecting an alternative. The user is free to choose any combination. Therefore, if a specific alternative or even the null alternative must be delivered, then that fact can be dealt with within this framework. Another table shows the modification of this assignment when the current alternative for each efficiency is considered. In this case, the possibility of reassignment of the starting alternative is turned off. The replacement alternative is selected from this adjusted table.

C.5 Crossover

Once an interim solution is selected for possible modification a preliminary decision has to be taken as whether use a mutation or crossover procedure to effect it. The choice of whether to mutate or cross it is governed by a binomial distribution at a rate of that can be altered between simulation batches.

Crosses for the interim solution are made by randomly selecting alternatives from a pair of neighbouring parent solutions for each deficiency using a binomial distribution of sample size one with mean 0.5. Effectively each parent has an equal opportunity to contribute to the hybrid.

Neighbouring parents are identified using the following procedure. The crossing span identifies how far above and below respectively from an interim solution's designated grid element to select the parent interim solutions. A crossing span of 1 means that the solutions one grid element above and below the interim are to be used as the source of crossover alternatives.

As implemented in the heuristic, the crossing span is a randomized integer ranging from 1 to 4. The procedure for selecting the crossing span is biased on a declining scale so that the value 1 is selected five times more frequently than the value 4. This means that solutions closest to the interim are five times more likely to be used as parents when compared to those that are three units away from the solution. The user can define both the range of distances and the associated weight between simulation batches.

C.6 Stopping Criteria

Rather than predetermine the number of iterations required to get at the efficient frontier, it has been found useful to conduct the simulation in batches. Each batch initializes several monitoring variables that provide indications of progress towards convergence. These include:

- a) the percentage of cells in the cost grid that have changed during the batch;
- b) the percentage change between the FOM total summed across the current solution set against that batch start;
- c) the percentage of solutions on the current efficient frontier;
- d) the average FOM of the five most costly (and therefore most productive) candidates; and
- e) a surrogate measure for the number of iterations between solution change.

One can reasonably be assured of near termination when:

- a) the percentage of cells changed as recorded at the end of batch is zero,
- b) the percentage change is below a ten thousandth of a percentage point,
- c) the number of points on the efficient frontier has peaked over several batches,
- d) the average FOM of the five most costly solutions is in the vicinity of the efficiency frontier established in any preceding work, and
- e) the average number of iterations between jumps is in the order of several ten thousands.

A reasonable batch size should contain enough simulation iterations to cover 100 generations. That is for a grid with one hundred points, 10,000, for a grid of 400, 40,000 and so on.

C.7 Improving Run Times

Over the extensive number of runs conducted to develop the heuristic, several factors seem to influence the time required to identify good solutions. The first is the number of range elements associated with the problem. Where there are a greater number of interim solutions that can be replaced, there are more FOM comparisons that need to be made in the model before a good effective frontier is identified. For the limited number of examples examined to date, the time to achieve a solution appears to increase linearly with the grid. It seems much less sensitive to the number of alternatives available for mutation.

Another factor affecting run time is the proportion of iterations that generate feasible though not necessarily optimal solutions. Clearly a large proportion of infeasible candidates only delays the time required to effectively define the problem's efficient frontier. The following recommendation is offered in an attempt to resolve this issue.

The optimization should be conducted at least initially over a coarse grid covering a very broad cost range. This should produce a grid with high percentage of efficient solutions. Then one should conduct a second optimization using a finer grid over a more restricted range by "seeding" the heuristic with those efficient solutions taken from the preceding optimization which would now correspond to grid element ranges in the second optimization. Overall, it has been found the total time required to define the efficient frontier for this second optimization in this two stage optimization is often faster than starting it from a zero base. It is also needless to remark that the efficient frontier that results from this is usually "better" than that generated by starting from a zero base.

The reason for the run time improvement can largely be attributed to the effectiveness of the crossing procedure to generate good candidate solutions for unfilled range elements from a pair of parent solutions, one of which producing a FOM lying near the efficient frontier for the second optimization.

The last factor affecting run time is the complexity of underlying cost, constraint, objective function and ancillary information calculation. It behoves the user to organize these so they are conducted as efficiently as possible.

C.8 Managing FOM Sensitivity

The method used here presupposes that the FOM relates to deficiencies and alternatives in a manner that involves no uncertainties. In actuality, the FOM is the result of an assessment process whose translation of qualitative ratings to numeric values may be fraught with stochastic uncertainty. To properly address this would require the comparison of solutions at specific nominal costs for an extensive range of FOM's calculated over the spectrum of potential values that their elemental components span. This means that for each selected combination of FOM components, one would calculate a set of solutions for the large scale alternative selection problem. If enough of these combinations could be assessed during the analysis, one would focus on particular nominal costs and define a solution at each cost examined through a search for persistent combinations of alternatives. It should be noted this approach to dealing with FOM uncertainty is considered outside the scope of the technology available to the Department at this time.

Should sensitivity analysis with respect to the FOM calculation become an issue, the best that could be done at the moment is to use one or more common statistics of distribution location in place of the nominal values provided. In particular, the mode of the distributions should be used if the thrust of the analysis are robust solutions, while the mean of distributions should be used if achievable effectiveness is the goal. The mode statistic is the most probable value of a particular random variable. If we were to perform the required simulations, the most common input configurations would be clustered near the modes of associated FOM distributions. Accordingly a mode based optimization would be considered typical of the kind of solutions to be encountered. Mean value optimization is more typically used if best achievable FOM is required. This follows from standard principles of stochastic optimization. However, distribution skewness of FOM is expected to be significant so that solutions so derived are likely to be outliers in the space of all solutions. In the SCR work at hand, nominal values, rather than either of these measures of location, were provided. Accordingly, it is suggested that this deficiency in the provided information be remedied in the work-up to any future SCR analysis.

C.9 Identifying the Pareto Frontier

The Mission Value Score, the key element that drives the calculation of FOM, is an aggregated function of scenario scores (see Reference 14). Suppose that the within scenario optimization has been resolved along the lines suggested above. Instead of choosing these aggregated values for the set of deficiencies as has been done for SCR 1.0, it may be more feasible to conduct the optimization by scenario. In this case, for specified cost, scenario specific optimal solutions could be defined – one for each scenario so examined.

Subsequently, each such optimal solution can be evaluated in the context of the remaining scenarios so that the degree that this solution generates FOM in each can be assessed. Taking the ratio of this FOM to the scenario maximum produces a dimension free rating of how the solution measures against its scenario optimum solution. Doing this for all these solutions against all

scenarios produces a scorecard similar to that depicted in Table 3 below, where the off-diagonal terms are the relative FOM delivery scores for each solution in a row and a column, which will typically range between 0 and 1.

Table 3: Solution Scenario Scorecard

	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5	Scenario 6
Solution 1	1	a_{12}	a_{13}	a_{14}	a_{15}	a_{16}
Solution 2	a_{21}	1	a_{23}	a_{24}	a_{25}	a_{26}
Solution 3	a_{31}	a_{32}	1	a_{34}	a_{35}	a_{36}
Solution 4	a_{41}	a_{42}	a_{43}	1	a_{45}	a_{46}
Solution 5	a_{51}	a_{52}	a_{53}	a_{54}	1	a_{56}
Solution 6	a_{61}	a_{62}	a_{63}	a_{64}	a_{65}	1

Using this scorecard, one can apply a number of techniques to decide upon which of these solutions comes closest to satisfying all.

The most intuitive approach is to combine all of the scenario objective functions into a single functional form. A well-known combination is the weighted linear sum of the objectives. One specifies scalar weights for each objective to be optimized, and then combine them into a single function that can be solved by any single-objective optimizer. Clearly, the solution obtained will depend on the relative values of the weights specified. Thus, it may be noticed that the weighted sum method is essentially subjective, in that a decision maker needs to supply the weights. Moreover, this approach cannot identify all non-dominated solutions. (Only solutions located on the convex part of the Pareto front can be found).

More objective ways of solving this problem might involve selecting the solution that maximizes the minimum relative scenario closure across the scenarios, maximizing the number of scenarios whose score lies beyond a specified threshold, minimizing the distance in some metric from complete closure or even greatest consistency ranking methods such as the DRDC developed MARCUS system [Ref 15].

To expand the analysis further, most of the measures cited above can serve as fitness functions for subsequent application of a genetic algorithm that creates hybrid solutions from the given set that may outperform all of the solutions associated a row in that table.

Another approach in dealing with feasible solution optimization within a multiple scenario construct is to use a generalization of the grid approach that was used to develop the heuristic in the first place. In this case one would consider an n dimension cubic lattice where n corresponds to the number of objectives, with an appropriate grid spacing whose proximal vertex, the one closest to the zero vector in n space, occurs at the minimal $a_{i,j}$ found in the analogue of Table 2 and whose furthest vertex occurs at the vector whose coordinates is 1. Each lattice point can be associated with combinations of FOMs that can be treated as bounds for a single objective function optimization problem to be described below.

Choose, for example, one of the bottom faces of this cubic lattice, one for which a minimal $a_{i,j}$ is attained. For each lattice point in this face, solve the single objective optimization problem for the objective function associated the free edge, the remaining parameter of the lattice, using the lattice point values of the lower face to define upper and lower bounds for the remaining objectives. It is assumed in passing that the cost upper bounds associated with solutions are retained as a constraint for all solutions. Doing this for a fine enough grid will should provide a reasonable outline of the Pareto response surface. However, approaching the optimization this way leads to the need to solve a growing set of intermediate problems that increases exponentially with the number of scenarios to be examined.

A more sophisticated approach to doing this in a single MSeExcel model would be to store interim solutions in a dynamically managed indexed table. The index is constructed to uniquely identify cost and scenario FOM ranges for as many scenarios as are needed for the analysis. To develop a solution candidate for possible insertion into this table, one first identifies an index to start the process of generating a solution. If an interim solution exists in the table, the associated set of alternatives is extracted in order to apply mutation or crossing - with the latter process extended to use its nearest neighbour information in a multi-dimensional setting. If an interim solution does not exist for the index, a zero default is assumed, with only mutation applied to modify it.

Next, the resulting solution candidate is evaluated and tested for constraint feasibility. If infeasible, the process for generating solution candidates is restarted. Otherwise, the evaluation results of the solution candidate are used to identify an index against whose interim solution's FOM (if an interim solution has been identified with the index) will be compared to the candidate's.. As before the interim solution for that index will be replaced by candidate solution if the candidate solution outperforms interim. In the event, that no interim solution is associated with the candidate's associated index, then the indexed table is expanded to accommodate the candidate solution and its characteristics.

Subsequently, the process just described for generating and evaluating candidate solutions is repeated until optimization convergence criteria are met. The indices of the table containing the last set of interim solutions can be parsed out to provide "coordinates" for their corresponding solutions in n dimensional space. These coordinates can then be used to visualize the required Pareto frontier for this analysis setting.

Annex D Estimating Cost Risk from Information about Skewed Distributions

D.1 Introduction

Distributions outlining cost risk for activities are notoriously skewed – the peak of these distributions, the figure most often presented to managers are most often lower than the expected value or the cost at which the activity ultimately delivers. For a large numbers of activities, the law of large numbers applies, so that regardless of distribution the cost of the sum of activities associated with statistical uncertainty is normally distributed about the sum of the expected values of each and with a variance equal to the sum of the individual variances. With this result, all the analytic functionality for estimating package cost risk can be had without resorting to simulation and only using basic facts about the normal distribution. For the analysis in the main body of the report, solutions often provided non-trivial alternatives for over forty issues so that risk estimated were generated by this means.

However, when this condition is not be met, such as for packages with thirty or fewer alternatives, the analysis might involve the approximation of the cost risk distribution through the Edgeworth, Gram-Charlier or Cornish – Fisher expansions. These require estimates of higher order moments of the sum distribution which then depend on the calculation of higher order moments of the cost risk model used for alternatives. The particular advantage of using these expansions is that the influence of the skewness of the cost distributions can be explicitly considered. In general, ignoring skewness and higher order moments generates an estimate for probability that the cost of the selected package of alternatives exceeds a given threshold that is below what it should be.

The following application of the Cornish-Fisher expansion will be considered. The Cornish-Fisher expansion is a formula for approximating quantiles of a random variable based only on its first few cumulants. The impact of carrying these higher terms for risk calculation is shown in the worked example at the end of this Annex.

D.2 Cumulants

D.2.1 Definition

The cumulants of a random variable X are related to its moments. They are defined, as those values κ_r such that the identity.

$$\exp\left(\sum_{r=1}^{\infty} \frac{\kappa_r t^r}{r!}\right) = \sum_{r=0}^{\infty} \frac{E(X^r) t^r}{r!}$$

holds for all t .

D.2.2 Cumulants and Moments

The cumulants are related to the raw moments of a distribution by the following recursion formula where the μ'_n corresponds to the E_n

$$\kappa_n = \mu'_n - \sum_{k=1}^{n-1} \binom{n-1}{k-1} \kappa_k \mu'_{n-k}.$$

The derivation of this formula is as follows. Take the first order derivative of both sides shown of the definition the cumulant. One notes on the left hand side that this derivative can be expressed as the product of the moment generating function and the derivative of the sum shown in definition in parenthesis in the definition of the cumulant. Evaluating both sides at $t=0$ leads to $\mu'_1 = \kappa_1$. Now take the $n-1$ th order derivative of this product and evaluate it at $t=0$. On the right hand side, we note that the product is μ'_n while we have the sum of κ_n and the sum shown in the above formula with a positive sign on the left. The formula is proven when the abovementioned sum is transposed to the other side of the equation and the sides of the equation ordered with κ_n first.

By inverting the formula and substituting every lower order instance of μ'_n by its expression in cumulants, the n th moment μ'_n is accordingly seen to be an n th-degree polynomial in the first n cumulants.

D.2.3 Properties of Cumulants

D.2.3.1 Invariance

The first cumulant is shift-equivariant; all of the others are shift-invariant. To state this less tersely, denote by $\kappa_n(X)$ the n th cumulant of the probability distribution of the random variable X . The statement is that if c is constant then $\kappa_1(X+c) = \kappa_1(X) + c$ and $\kappa_n(X+c) = \kappa_n(X)$ for $n \geq 2$, i.e., c is added to the first cumulant, but all higher cumulants are unchanged.

A consequence of this principle is the following: the cumulants of a distribution are closely related to distribution's central moments. If a random variable X admits an expected value $\mu = E(X)$ and a variance $\sigma^2 = E((X-\mu)^2)$, then these are the first two cumulants: $\mu = \kappa_1$ and $\sigma^2 = \kappa_2$. More generally, the following proves true.

$$\begin{aligned}\kappa_1 &= \mu \\ \kappa_2 &= \mu_2 \\ \kappa_3 &= \mu_3 \\ \kappa_4 &= \mu_4 - 3\mu_2^2 \\ \kappa_5 &= \mu_5 - 10\mu_3\mu_2\end{aligned}$$

This is proved by applying the recursive definition of cumulants cited above to a translation of the given distribution so that the mean of the latter is zero. The original and translated distributions have the same central means but now the terms in the recursion corresponding to mean value

vanish. The appearances of κ 's on the right hand side of the resulting expression, for example, those that appear in the equations for κ_4 and κ_5 , are eliminated through use the equations that resulted for them through the application of this procedure for lower order terms.

Conversely by successive substitution, the first five central moments in terms of cumulants are

$$\begin{aligned}\mu_1 &= \kappa_1 \\ \mu_2 &= \kappa_2 \\ \mu_3 &= \kappa_3 \\ \mu_4 &= \kappa_4 + 3\kappa_2^2 \\ \mu_5 &= \kappa_5 + 10\kappa_2 \kappa_3\end{aligned}$$

shown below.

D.2.3.2 Homogeneity

The n th cumulant is homogeneous of degree n , i.e. if c is any constant, then

$$\kappa_n(cX) = c^n \kappa_n(X).$$

This relation is used in two places. First, it is the relation from which the cumulants of a triangular distribution whose peak value is its nominal cost are calculated from those of a triangular distribution with peak at 1. The scale factor in this case is obviously the nominal cost. Later on in this section it is used to relate the cumulants of a normalized distribution to those of its parent. In that case, the scale factor is the standard deviation of the parent or equivalently the square root of κ_2

D.2.3.3 Additivity

If X and Y are independent random variables then $\kappa_n(X + Y) = \kappa_n(X) + \kappa_n(Y)$, so that each cumulant of a sum is the sum of the corresponding cumulants of the addends. In particular, for the analysis at hand, the cumulants of the sum of the component costs can be estimated as the sum of the cumulants of each component. It is this property that motivates the suggested extension.

It should be first noted that the distribution of the sum of independent random variables is specified by the repeated convolution of its component distributions. In this circumstance, it can be shown that the moment generating function of the sum of independent random variables is the product of the moment generating function of each. This is proved in a manner similar to that found in the proof of the fact that the Laplace transform of a convolution is the product of the transforms of the convolution components. This translates to the sum of the cumulants through the taking of logarithms of the moment generating functions.

D.3 Application

D.3.1 Theoretical Approach

The following table provides the cumulants related to the lower and upper bounds that were associated the cost risk rankings provided with alternative description information provided by the Capability Management Community [Reference 11]. This is the basic information needed to motivate the application the method. The values for the cumulants were derived by first calculating the raw moments of the triangular distribution associated with α and β as in Section E.2 below and then applying cumulant- raw moment relation of Section D.2.3.1 recursively.

Table 4: Cumulants Associated With Selected Cost Uncertainty Bounds

	α	β	κ_1	κ_2	κ_3	κ_4	κ_5
Low	0.75	1.75	1.166667	0.045139	0.004051	-0.00122	-0.00078
Medium	0.50	2.00	1.166667	0.097222	0.009259	-0.00567	-0.00386
High	0.50	3.00	1.500000	0.291667	0.075000	-0.05104	-0.09375

All alternatives come with a cost uncertainty rating of either Low, Medium or High cost uncertainty. These have been assigned increments or decrements of the nominal cost to define upper and lower bounds for cost. These adjustments define the α and β shown in the table. With these pairs given, the cumulants associated with the triangular distribution with mode 1 are shown across the remaining five columns of the table. Using the homogeneity property, the cumulants for the cost of the selected alternative can be determined accordingly. Assuming independence of the triangular distributions associated with alternative costs, the cumulant for the sum of costs for the selected solution is the sum of the cumulants of the component costs through the additivity property.

By way of example, for a normal distribution, the number of standard deviations from the mean for which the cumulative distribution achieves its 95 percentile is 1.64. In the following, let q be .95 and $\Phi^{-1}_Z(q) = 1.64$

Suppose X has mean 0 and standard deviation 1. Cornish and Fisher (1937)⁸ provide an expansion for approximating the q -quantile, $\Phi^{-1}_X(q)$, of X based upon its cumulants. Using the first five cumulants, the expansion is

⁸ Cornish, E. A. and Fisher, R. A. "Moments and Cumulants in the Specification of Distributions." *Extrait de la Revue de l'Institut International de Statistique* 4, 1-14, 1937. Reprinted in Fisher, R. A. *Contributions to Mathematical Statistics*. New York: Wiley, 1950.

$$\begin{aligned}\Phi_X^{-1}(q) \approx & \Phi_Z^{-1}(q) + \frac{\Phi_Z^{-1}(q)^2 - 1}{6} \kappa_3 + \frac{\Phi_Z^{-1}(q)^3 - 3\Phi_Z^{-1}(q)}{24} \kappa_4 \\ & - \frac{2\Phi_Z^{-1}(q)^3 - 5\Phi_Z^{-1}(q)}{36} \kappa_3^2 + \frac{\Phi_Z^{-1}(q)^4 - 6\Phi_Z^{-1}(q)^2 + 3}{120} \kappa_5 \\ & - \frac{\Phi_Z^{-1}(q)^4 - 5\Phi_Z^{-1}(q)^2 + 2}{24} \kappa_3 \kappa_4 + \frac{12\Phi_Z^{-1}(q)^4 - 53\Phi_Z^{-1}(q)^2 + 17}{324} \kappa_3^3\end{aligned}$$

where $\Phi_Z^{-1}(q)$ is the q -quantile of a standard normal random variable Z . Although the above formula applies only if X has mean 0 and standard deviation 1, we can still use it to approximate quantiles if X has some other mean and standard deviation. Simply define the normalization of X as

$$X^* = \frac{X - \mu}{\sigma}$$

which has mean 0 and standard deviation 1. Central moments of X^* can be calculated from central moments of X with

$$\mu_r^* = \frac{\mu_r}{\sigma^r}$$

and the central moments of X can be obtained from its cumulants as shown above. Using the transformed central moments, calculate their associated cumulants through formulas 2 to 6. Apply the Cornish-Fisher expansion to obtain the q -quantile x^* of X^* . The corresponding q -quantile x of X is then

$$x = x^* \sigma + \mu$$

D.3.2 Numerical Example

To apply these formulas for the example quantile, consider that the cumulants for a particular combination of alternatives turn out to be as follows:

Table 5: Sample Cumulants

κ_1	κ_2	κ_3	κ_4	κ_5
2.978	0.05	0.001	-1E-04	-3E-05

Transform cumulants into central moments:

Table 6: Central Moments Corresponding to Sample Cumulants

μ_1	μ_2	μ_3	μ_4	μ_5
2.978	0.05	0.001	0.0074	0.00072

Transform central moments into normalized central moments:

Table 7: Normalized Central Moments Corresponding to Central Moments

μ_1	μ_2	μ_3	μ_4	μ_5
0	1	0.134	2.9462	1.27787

Transform normalized central moments into normalized cumulants.

Table 8: Normalized Cumulants Corresponding to Normalized Central Moments

κ_1	κ_2	κ_3	κ_4	κ_5
0	1	0.134	-0.054	-0.0606

Apply Cornish - Fisher formula.

$$1.69 = 1.64 + 0.038 + 0.00109 - 0.00034 + 0.002988 - 0.00126 - 0.00029$$

Transform modified normal quantile into original variables.

$$3.354 = 2.978 + 1.69 * \sqrt{0.05}$$

Accordingly the actual cost of the solution will exceed this identified cost with a probability no more than the 5% rather than the 3.344 had correction not been made. This indicates that the assertion made in the main body of the text that the normal model should provide reasonable thresholds for this problem setting appears to be justified in this instance.

The final step in the analytic evaluation of a solution with respect to cost risk is to assess the chance that a solution exceeds a fixed cost threshold. This is the inverse problem of the one we have investigated. At this time, there appears to be no other alternative other running the above expansion for a suitably comprehensive grid of quantiles and then interpolating results. Insofar as the heuristic is concerned, a normal distribution was presupposed to apply to the solution total so that a Z variate can be determined and its probability could be read off a normal table, which in our case is a built in function in MSExcel.

Observe once again that the specification of a distribution through its cumulants is not specific to triangular distributions. The steps described in this section apply as well to any distribution family compatible any finite ranged continuous cost risk specification. Examples of these are the uniform and displaced Beta families of distributions. Moreover, representatives of any these distributions can be applied to any single cost in this framework with the consequence that for example some alternatives have their risk specified by a triangular distribution while others are specified as uniform or displaced Beta. The only recommendation to be considered is the rapid calculation of the selected distribution's cumulants to the required degree or at least its raw or centralized moments which accordingly can be used to calculate its cumulants as described earlier in this Annex. The only reason that a triangular family for cost risk is assumed was to remain consistent with the SCR 1.0 formulation as described in Annex B.

This page intentionally left blank

Annex E Some Properties Of Skewed Triangular Distributions

E.1 Introduction

In working with our defence partners to specify the degree of confidence they can give to cost estimates, one frequently encounters statements of the sort:

“The most likely cost is to be C but it could be as low as $\alpha\%$ of C and as high as $\beta\%$ of C .”

What the partner is saying is that the cost involved is specified by a probability distribution with these characteristics, and which for the analyst means that it could be safely represented as the product of the modal cost, which we in this report have called nominal cost, and a skewed triangular distribution of mode 1.

This Annex is devoted to presenting and proving several statistical facts about these distributions for application in contingency planning. In particular, the exact formulation of the triangular distribution will be given, its expected value will be determined and finally percentile points will be identified. These outcomes will be applied elsewhere in this document to determine mean contingency for a spending program.

Note that the information given by the partner is also compatible with other families of probability distributions – most notably an affine transform of the Beta distribution. The line of argument shown in Annex D about using distribution cumulants to generate approximate values of risk applies as well to costs specified by this latter family of distributions. Furthermore there are well – known closed formulae for the cumulants of the Beta distribution that depend solely on the two distribution shape parameters that specify a unique distribution from this family. However, there are multiple Beta distributions compatible with the information shown in quotes above. In particular, there are parameter combinations that can guarantee the location of the distribution peak, that is the most likely value, but each generates a different concentration of probability about it. The given information provides no guidance about what probability concentration is most appropriate.

It has also been suggested that cost risk may be better described through log-normal distributions compatible with the given information. As stated above, there was no guidance offered to use the information given above to provide an unambiguous specification of the distribution or its moments. Are the tails of the log-normal 5 and 95 percentile points or some other limitation of the potential range of cost values? Once this is resolved, it should be rather straightforward to calculate appropriate cumulants and so apply the methods described in Annex D of this report.

After all this discussion, a minimalist approach of using the triangular family of statistical distributions was adopted for cost risk specification for this development. This furthermore ensured compatibility with the risk calculation approach of SCR 1.0. See Annex B. For other CORA publications applying three-point estimation in uncertainty analysis, including an application using a triangular distribution model for uncertainty, see References 12 to 14.

The only issue remaining was that the cumulants and other important statistical properties of these triangular distributions are not readily found in recognized references. As a result, these facts were derived from first principles and presented here for completeness.

E.2 Elementary Properties

Let α be a fraction between zero and one representing the lower bound of the triangular distribution, β its upper bound. Here α and β represent the fractions outlined in quotes in the previous section and where C, the mode, has been taken as 1. One can always reconstruct the properties to be outlined below for the fully specified triangular distribution through the scaling properties of moments and cumulants outlined in Annex D. We also assume that α and β are distinct although either of which can be 1. These assumptions cover the triangular cases for cost uncertainty modelling discussed in Annex B. Since a triangular distribution is normalized, the triangle specified by these upper and lower bound with mode 1 must have a height of

$$h = 2 / (\beta - \alpha),$$

which follows from the elementary formula that the area of a triangle is product of $\frac{1}{2}$ its base times height. With h known, the formula for the distribution can be stated explicitly as

$$y = f(x) = h (x - \alpha) / (1 - \alpha),$$

when $\alpha \leq x \leq 1$ and

$$y = f(x) = h (x - \beta) / (1 - \beta),$$

when $1 \leq x \leq \beta$.

To calculate the expected value for the distribution, the standard formula is applied to the triangular distribution just derived, that is, if E is the expected value then

$$E = \int_{\alpha}^{\beta} x f(x) dx,$$

where y is given above. When the integral of the two components are evaluated, the one between α and 1, and the other between 1 and β , the following formula for E arises:

$$E = h * [\beta (\beta - 1)/2 - (\beta - 1)^2/3 + \alpha (1 - \alpha)/2 + (1 - \alpha)^2/3],$$

or more simply following the derivation of moments in general given later in this section,

$$E = (\alpha + \beta + 1)/3$$

While this formula does not appear too intuitive at this stage, it does collapse nicely to an obvious result when the triangular distribution is symmetric. In this case,

$$d = 1 - \alpha = \beta - 1.$$

This implies that

$$\beta - \alpha = 2d.$$

Substituting into the second formula for E gives

$$\begin{aligned} E &= [(1+d) + (1-d) + 1]/3 \\ &= 1 \end{aligned}$$

In terms of practical application, the following table gives the degree to which the mean exceeds the mode for a range of typical values of α and β encountered in cost analysis. The implication of this when applied to the selection of future projects should be that the nominal value shown for costs should be adjusted by the appropriate E factor to account for inherent cost uncertainty associated with capability alternative components.

Table 9: Expected Values Associated With Specified Risk Ranges

α	0.50	0.60	0.70	0.80	0.90
β	2.00	1.67	1.43	1.25	1.11
E	1.17	1.09	1.04	1.02	1.00

Similarly, higher order moments for this distribution can be calculated. Of particular interest is the second moment, which proves useful for the subsequent calculation of the distribution variance. This moment, proven explicitly later in the presentation, is given by

$$E_2 = \int_{\alpha}^{\beta} x^2 f(x) dx = [\beta^2 + \beta\alpha + \alpha^2 + \beta + \alpha + 1] / 6.$$

Now the variance of this triangular distribution, Var, is calculated as

$$\text{Var} = E_2 - E^2$$

or explicitly as

$$\text{Var} = (\alpha^2 - \alpha\beta + \beta^2 - \alpha - \beta + 1) / 18.$$

Table 10 below gives the variance associated with the range of upper and lower bound found in Table 9

Table 10: Variance Associated With Specified Risk Ranges

α	0.50	0.60	0.70	0.80	0.90
β	2.00	1.67	1.43	1.25	1.11
Var	0.0972	0.0487	0.0224	0.0085	0.0018

In working with large numbers of uncertain costs, each of which is independently distributed, the characteristics of the cost sum can be identified through the expectation and the variance of the individual costs. Using the common statistical practice of treating the sum as a normal distribution about the sum of the individual cost expectation and whose variance is given by the sum of the individual variances, one can thereby estimate the probability that the given spending program exceeds assigned cost thresholds.

When correlations exist between the items being cost estimated, a modification of the preceding argument can be applied to correctly calculate the variance of the sum. The limitation in applying the more general case is the generation of estimates of inter-cost correlation.

In general, calculate the n'th moment about the origin of a triangular distribution given above through

$$\begin{aligned}
 E_n &= \int_{\alpha}^{\beta} x^n f(x) dx \\
 &= \int_{\alpha}^1 x^n f(x) dx + \int_1^{\beta} x^n f(x) dx \\
 &= (2/(\beta - \alpha))(1 - \alpha) \left(\int_{\alpha}^1 s^n (s - \alpha) ds \right) + (2/(\beta - \alpha))(\beta - 1) \left(\int_1^{\beta} s^n (\beta - s) dx \right), \\
 &= 2/(\beta - \alpha)(1 - \alpha) \left[s^{n+2}/(n+2) - \alpha s^{n+1}/(n+1) \right]_{\alpha}^1 \\
 &\quad + 2/(\beta - \alpha)(\beta - 1) \left[\beta s^{n+1}/(n+1) - s^{n+2}/(n+2) \right]_1^{\beta},
 \end{aligned}$$

$$\begin{aligned}
&= 2/(\beta - \alpha) ((n+1)) - \alpha(n+2) + \alpha^{n+2}) /[(1-\alpha)(n+1)(n+2)] \\
&\quad + 2/(\beta - \alpha) ((n+1)) - \beta(n+2) + \beta^{n+2}) /[(\beta-1)(n+1)(n+2)], \\
&= 2/(\beta - \alpha) ((n+2)) - \left[\sum_{j=0}^{n+1} \alpha^j \right]) /[(n+1)(n+2)] \\
&\quad + 2/(\beta - \alpha) (\left[\sum_{j=0}^{n+1} \beta^j \right] - (n+2)) /[(n+1)(n+2)], \\
&= 2 \left[\sum_{j=0}^{n+1} (\beta^j - \alpha^j) \right] /[(\beta - \alpha)(n+1)(n+2)].
\end{aligned}$$

.

Since

$$(\beta^j - \alpha^j) / (\beta - \alpha) = \sum_{k=0}^{j-1} \beta^{j-1-k} \alpha^k,$$

the above formula for E_n can be written as

$$\begin{aligned}
E_n &= \left[\sum_{j=0}^n \sum_{k=0}^j \beta^{j-k} \alpha^k \right] / [(n+1)(n+2)/2] \\
&= \left(\sum_{i+j+k=n} 1^k \beta^j \alpha^i \right) / \binom{n+2}{2}
\end{aligned}$$

where it should be noted that

$$\begin{aligned}
&[(n+1)(n+2)/2] \\
&= [n!(n+1)(n+2)/((2!)(n!))], \\
&= [(n+2)!]/[(2!)(n!)], \\
&= \binom{n+2}{2}.
\end{aligned}$$

with i, j and k are integers ranging from zero to n . Despite the ugly formulation, the formula for E_n has some elegant symmetry properties. First, the roles of α and β are symmetric, that is any permutation of these variable in the formula generate the same value. Next note the number of terms in the sum is equal to the value of the denominator. This fact can be proved through mathematical induction.

One verifies immediately that

$$E = [1 + \alpha + \beta] / 3 \text{ and}$$

$$E_2 = [1 + \alpha + \beta + \alpha^2 + \alpha\beta + \beta^2] / 6.$$

For possible future application, we note that

$$E_3 = [1 + \alpha + \beta + \alpha^2 + \alpha\beta + \beta^2 + \alpha^3 + \alpha^2\beta + \alpha\beta^2 + \beta^3] / 10,$$

$$E_4 = [1 + \alpha + \beta + \alpha^2 + \alpha\beta + \beta^2 + \alpha^3 + \alpha^2\beta + \alpha\beta^2 + \beta^3 + \alpha^4 + \alpha^3\beta + \alpha^2\beta^2 + \alpha\beta^3 + \beta^4] / 15,$$

$$E_5 = [1 + \alpha + \beta + \alpha^2 + \alpha\beta + \beta^2 + \alpha^3 + \alpha^2\beta + \alpha\beta^2 + \beta^3 + \alpha^4 + \alpha^3\beta + \alpha^2\beta^2 + \alpha\beta^3 + \beta^4 + \alpha^5 + \alpha^4\beta + \alpha^3\beta^2 + \alpha^2\beta^3 + \alpha\beta^4 + \beta^5] / 21,$$

The n th moment about the mean, μ_n , can be calculated from the E_n and all preceding E_n using the binomial theorem as follows:

$$\begin{aligned}
\mu_n &= \int_{\alpha}^{\beta} (x - E)^n f(x) dx, \\
&= \int_{\alpha}^{\beta} \sum_{j=0}^n \binom{n}{j} (x)^{n-j} (-1)^j E^j f(x) dx, \\
&= \sum_{j=0}^n \binom{n}{j} \left(\int_{\alpha}^{\beta} (x)^{n-j} f(x) dx \right) (-1)^j E^j, \\
&= \sum_{j=0}^n \binom{n}{j} E_{n-j} (-1)^j E^j.
\end{aligned}$$

For third and fourth central moments, the corresponding values reduce to

$$\mu_3 = -(\beta + \alpha - 2)(1 + \beta - 2\alpha)(1 + \alpha - 2\beta) / 270, \text{ and}$$

$$\mu_4 = (\alpha^2 - \alpha\beta + \beta^2 - \alpha - \beta + 1)^2 / 135$$

$$= 2.4 * \mu_2^2$$

E.3 Percentiles

Also in the current application it would be useful to determine the probability of exceeding an assigned value of X between α and β , given the distribution cited above.

This probability is just the converse of the probability that the distributed value is exceeded by the assigned threshold. This latter probability is easily determined through the structure of the distribution:

Using the explicit formula for the triangular distribution given earlier, the probability P is given as

$$P = \int_{\alpha}^X f(x) dx = (X - \alpha)^2 / [(\beta - \alpha)(1 - \alpha)]$$

if $X \leq 1$, and

$$P = 1 - \int_X^{\beta} f(x) dx = 1 - (\beta - X)^2 / [(\beta - \alpha)(\beta - 1)]$$

otherwise. Note that for X at 1,

$$P_I = (1 - \alpha) / (\beta - \alpha),$$

which is the ratio that the distance to 1 from α makes against the base of the triangle. Note that $(1 - \alpha) = (\beta - \alpha) P_I$. Due to the nature of skewness for this application P_I is almost always below .5.

Accordingly with a given value of cumulative probability P known to be below P_I , the value of X associated with it is

$$X = \alpha + (\beta - \alpha) \sqrt{PP_I}.$$

Similarly, if $P > P_I$,

$$X = \beta - (\beta - \alpha) \sqrt{(1 - P)(1 - P_I)}.$$

This reverse specification can be used in the following context. Suppose that the package of alternatives has to have a high assurance of delivery for each item. Say of the one hundred items chosen, no more than 5 percent of these is to exceed an item specific cap on costs. To get this at a reasonable degree of assurance say 75% of the time might require that project cost should never exceed the 97th percentile, The cap associated with this threshold can be determined by applying the second formula for $P=0.97$. The associated X value is then the cost to be applied in program optimization.

An often important statistic involving percentiles is the median – the 50th percentile point. Due to the nature of skewness for this application, the median is then prescribed as

$$\begin{aligned} X &= \beta - (\beta - \alpha) \sqrt{(1 - P)(1 - P_I)}, \\ &= \beta - (\beta - \alpha) \sqrt{(1 - P_I)/2}, \\ &= \beta - (\beta - \alpha) \sqrt{(1 - [(1 - \alpha) / (\beta - \alpha)]) / 2}, \\ &= \beta - (\beta - \alpha) \sqrt{[\beta - 1] / (2[\beta - \alpha])}. \end{aligned}$$

This follows from the preceding formula for X with P set to $\frac{1}{2}$ and P_I defined as above. If the distribution is skewed to the left, that is if P_I greater than .5,

$$X = \alpha + (\beta - \alpha) \sqrt{[1 - \alpha] / (2[\beta - \alpha])}.$$

These last few results enable a resolution of the following extension to the question that was asked at the beginning of this Annex.

“The most likely cost is to be C but it could be as low as $\alpha_l\%$ of C with a probability of P_a of being below and as high as $\beta_l\%$ of C with a probability of exceeding it of P_b .”

Provided that P_a and P_b are below 15 – 20 %, it is possible to use this information to fit a triangular distribution that is compatible with this information. Note that the role of P_b is the inverse of that the use of the generic probability P earlier in the section.. Using

$$P = (X - \alpha)^2 / [(\beta - \alpha)(1 - \alpha)],$$

with P replaced by P_a and X by α_l , it is seen that a compatible pair of α and β must satisfy the following :

$$\alpha = \alpha_l - \sqrt{(P_a (\beta - \alpha)(1 - \alpha))}.$$

Similarly, noting the changed role of P_b ,

$$\beta = \beta_l + \sqrt{(P_b (\beta - \alpha)(\beta - 1))}.$$

Thus we have a pair of implicit equations in two unknowns. Under very general conditions, that are satisfied for these equations, a pair of α and β exists satisfying them provided P_a and P_b are sufficiently small. This requirement follows from the basic existence theorem for implicit functions of multiple variables and is also necessary so that standard convergence criteria for the successive iteration procedure to be described below remain satisfied. Accordingly, the following iteration procedure is suggested to determine what these are from the given information. First, set α and β to 1 on the right hand side of this pair of equations. This results in α_l and β_l for α and β respectively.

For the $n+1$ 'th step, with α_n and β_n as specified from the preceding step, then, α_{n+1} and β_{n+1} are given by

$$\alpha_{n+1} = \alpha_l - \sqrt{(P_a (\beta_n - \alpha_n)(1 - \alpha_n))},$$

and

$$\beta_{n+1} = \beta_l + \sqrt{(P_b (\beta_n - \alpha_n)(\beta_n - 1))}.$$

For the number of example cases examined, it was found that this iteration converged for small values of P_a and P_b in about 20 steps. One should be cautioned that iteration outcomes for α that converge below zero indicate that the problem information is fundamentally incompatible with a triangular distribution model. In that event, the analyst should ask the information provider for greater clarification. In practice, the risk information provided by our military partner was of the type discussed earlier in the Annex.

List of symbols/abbreviations/acronyms/initialisms

C Prog	Chief of Program
CBP	Capability-Based Planning
CDS	Chief of Defence Staff
CF	Canadian Forces
CFD	Chief Force Development
DND	Department of National Defence
DRDC	Defence Research & Development Canada
DRDKIM	Director Research and Development Knowledge and Information Management
EAC	Equivalent Annual Cost
FD	Force Development
FOM	Figure of Merit
R&D	Research & Development
SCR	Strategic Capability Roadmap
TTCP	The Technical Cooperation Program

DOCUMENT CONTROL DATA		
(Security classification of title, body of abstract and indexing annotation must be entered when the overall document is classified)		
1. ORIGINATOR (The name and address of the organization preparing the document. Organizations for whom the document was prepared, e.g. Centre sponsoring a contractor's report, or tasking agency, are entered in section 8.) DRDC Centre for Operational Research & Analysis National Defence Headquarters 101 Colonel By Drive Ottawa, ON K1A 0K2	2. SECURITY CLASSIFICATION (Overall security classification of the document including special warning terms if applicable.) UNCLASSIFIED	
3. TITLE (The complete document title as indicated on the title page. Its classification should be indicated by the appropriate abbreviation (S, C or U) in parentheses after the title.) Unifying Capability Integration Analysis – Initial Insights (U)		
4. AUTHORS (last name, followed by initials – ranks, titles, etc. not to be used) Kerzner, L		
5. DATE OF PUBLICATION (Month and year of publication of document.) September 2011	6a. NO. OF PAGES (Total containing information, including Annexes, Appendices, etc.) 84	6b. NO. OF REFS (Total cited in document.) 15
7. DESCRIPTIVE NOTES (The category of the document, e.g. technical report, technical note or memorandum. If appropriate, enter the type of report, e.g. interim, progress, summary, annual or final. Give the inclusive dates when a specific reporting period is covered.) Technical Memorandum		
8. SPONSORING ACTIVITY (The name of the department project office or laboratory sponsoring the research and development – include address.) DRDC CORA		
9a. PROJECT OR GRANT NO. (If appropriate, the applicable research and development project or grant number under which the document was written. Please specify whether project or grant.)	9b. CONTRACT NO. (If appropriate, the applicable number under which the document was written.)	
10a. ORIGINATOR'S DOCUMENT NUMBER (The official document number by which the document is identified by the originating activity. This number must be unique to this document.) DRDC CORA TM 2011-022	10b. OTHER DOCUMENT NO(s). (Any other numbers which may be assigned this document either by the originator or by the sponsor.)	
11. DOCUMENT AVAILABILITY (Any limitations on further dissemination of the document, other than those imposed by security classification.) Unlimited		
12. DOCUMENT ANNOUNCEMENT (Any limitation to the bibliographic announcement of this document. This will normally correspond to the Document Availability (11). However, where further distribution (beyond the audience specified in (11) is possible, a wider announcement audience may be selected.) Unlimited		

13. **ABSTRACT** (A brief and factual summary of the document. It may also appear elsewhere in the body of the document itself. It is highly desirable that the abstract of classified documents be unclassified. Each paragraph of the abstract shall begin with an indication of the security classification of the information in the paragraph (unless the document itself is unclassified) represented as (S), (C), (R), or (U). It is not necessary to include here abstracts in both official languages unless the text is bilingual.)

In 2005, the Chief of Defence Staff (CDS) of the Canadian Forces (CF) mandated that Capability-Based Planning (CBP) be institutionalized as a part of a centrally driven, top-down approach to Force Development (FD) within the Department of National Defence (DND). For the last four years, military and defence analyst staffs have developed and implemented the first iteration of an end-to-end capability-based FD process. This process ultimately produces a list of programs necessary to meet the demands of the future security environment over the next 20 years. As part of the Capability Integration phase of FD, distinct methods are employed to determine the set of alternatives which maximizes CF capability that is affordable within the available Defence budget, financial risk, alternative priority and the implementation schedule for the selected alternatives. This report describes the results of a research initiative conducted in 2008 whose purpose was to construct a different optimization scheme which overcomes the processing and validation deficiencies that were observed during SCR work. At the same time, several other Capability Integration components have been incorporated into this revised optimization framework so that much more of the integration related activity, as well as the sensitivity analysis work so necessary to establish solution credibility, can be conducted in an affordable and timely fashion.

14. **KEYWORDS, DESCRIPTORS or IDENTIFIERS** (Technically meaningful terms or short phrases that characterize a document and could be helpful in cataloguing the document. They should be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location may also be included. If possible keywords should be selected from a published thesaurus, e.g. Thesaurus of Engineering and Scientific Terms (TEST) and that thesaurus identified. If it is not possible to select indexing terms which are Unclassified, the classification of each should be indicated as with the title.)

Capability-Based Planning, Force Development, Strategic Capability Roadmap, Analytic Framework, Capability Management, Capability Integration, Cost Sensitivity, Capability Deficiency, Mission Value Score, Capability Alternative, Phoenix Integration, Optimization, Genetic Algorithm, Objective Force, Risk, Simulations

Defence R&D Canada

Canada's Leader in Defence
and National Security
Science and Technology

R & D pour la défense Canada

Chef de file au Canada en matière
de science et de technologie pour
la défense et la sécurité nationale



www.drdc-rddc.gc.ca

