3-2004

Measuring the Risk of Shortfalls in Air Force Capabilities

William E. Woodward

MEASURING THE RISK OF SHORTFALLS IN AIR FORCE CAPABILITIES

THESIS

William E. Woodward, Captain, USAF

AFIT/GOR/ENS/04-13

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright Patterson Air Force Base, Ohio

The views expressed in this thesis are those of the author and do not necessarily reflect the official policy or position of the United States Air Force, the Department of Defense or the United States Government.

AFIT/GOR/ENS/03-13

MEASURING THE RISK OF SHORTFALLS IN AIR FORCE CAPABILITIES

THESIS

Presented to the Faculty

Department of Operational Sciences

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

In Partial Fulfillment of the Requirements for the

Degree of Master of Science in Operations Research

William E. Woodward, BS, MPP

Captain, USAF

March 2004

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

AFIT/GOR/ENS/04-13

MEASURING THE RISK OF SHORTFALLS IN AIR FORCE CAPABILITIES

William E. Woodward, BS, MPP
Captain, USAF

Approved:

______//signed//________________________ _18 Mar 04__
Dr. Richard F. Deckro (Advisor) date

______//signed//________________________ _18 Mar 04__
Dr. Jeffrey P. Kharoufeh (Reader) date

______//signed//____ ___________________ _11 Mar 04__
Stephen P. Chambal, Maj, USAF (Reader) date

AFIT/GOR/ENS/04-13

Abstract

The U.S. Air Force seeks to measure and prioritize risk as part of its Capabilities Review and Risk Assessment (CRRA) process. The goal of the CRRA is to identify capability shortfalls, and the risks associated with those shortfalls, to influence future systems acquisition. Many fields, including engineering, medicine and finance, seek to model and measure risks. This research utilizes various risk measurement approaches to propose appropriate risk measures for a military context. Specifically, risk is modeled as a non-negative random variable of severity. Four measures are examined: simple expectation, a risk-value measure, tail conditional expectation, and distorted expectation. Risk measures are subsequently used to weight the objective function coefficients in a system acquisition knapsack problem.

Acknowledgments

This thesis would not have been possible without the help of many. My advisor, Dr. Dick Deckro, gave me latitude to explore and shared his immense store of knowledge. Dr. Jeff Kharoufeh kept me, probably, within proper mathematical bounds. Major Steve Chambal made contacts, set up meetings and generally boosted my reputation and my research in our nation's capital.

My sponsor, the Requirements Division of the Air and Space Operations Directorate on the Air Staff, provided funding, support and a continually interesting and relevant problem. In particular, Brigadier General (select) Mike Snodgrass and Colonel Kevin Martin enthusiastically encouraged me to challenge the Pentagon's conventional wisdom. Captain Rob Renfro at the Air Force Studies and Analyses Agency provided helpful comments and invaluable insight into the needs of the process.

I would like to thank my "wife," my "girlfriend" and the other eleven members of my class at AFIT, who have made the study of operations research interesting, fun and immensely quotable. Finally, I most gratefully acknowledge my Heavenly Father for providing me with yet another opportunity to study and analyze this amazing world He has created.

William E. Woodward

Ecclesiastes XII.xii

Table of Contents

Page

Page

List of Figures

Page

Page

List of Tables

MEASURING THE RISK OF SHORTFALLS IN AIR FORCE CAPABILITIES

I. Introduction

1.1. Issue Overview

In a continuing effort to prepare for future threats to United States security, the U.S. Air Force has implemented a new analytic planning tool, the Capabilities Review and Risk Assessment (CRRA). This process, a top-down analysis of Air Force capability, is designed to guide service planning, requirements development and system acquisition. The CRRA builds on six operational concepts to evaluate the value of specific Air Force programs to war-fighting effects. According to the Air Force Chief of Staff, the ultimate goal of the CRRA is "an operational, capabilities-based focus for acquisition program decision making" (Jumper, 2002).

There are six operational concepts that outline Air Force operations:

- Global strike: gain and maintain access to the battle space
- Space & C4ISR: integrate systems to provide information
- Global response: attack high-value targets within hours
- Homeland security: prevent, protect and respond to threats against U.S. territory
- Nuclear response: provide a deterrent and prepare to use
- Global mobility: project, employ and sustain U.S. power around the globe

Based on these six concepts, the CRRA process identifies an exhaustive list of desired Air Force capabilities known as the "master capabilities library". These capabilities must be systematically reviewed to identify where the Air Force falls short in its desired capability. Each capability shortfall can then be assessed for risk. Figure 1 shows the five, iterative steps of the CRRA process. The steps involving risk assessment are the focus of this thesis.

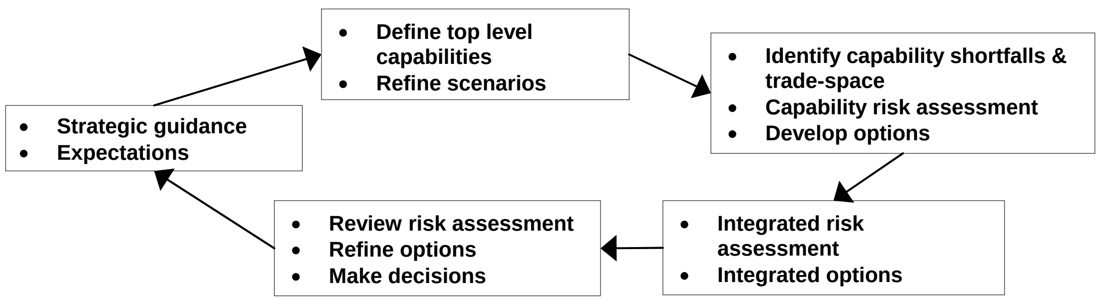

Figure 1. Capabilities Review and Risk Assessment Process

This research aims to provide a methodological basis for this risk assessment. The Presidential Commission on Critical Infrastructure Protection laid out an appropriate goal for risk assessment that may also be relevant to the CRRA process. For the quantification of risk the Air Force needs "methodologies, tools and organizational processes" to handle "uncertainties in, or incomplete knowledge of, threats, vulnerabilities, and protection measures; and for managing risks across multiple components and organizations" (PCCIP, 1997:90). The overarching question for this research, then, is how to prioritize risks when measuring Air Force capability shortfalls?

1.2. Background and Importance

Risk analysis is a diverse and growing field with a variety of opinions over its appropriate focus. Broadly speaking, risk analysis breaks into two areas: physical risk studied by engineers and the medical community, and financial risk in investment management and the insurance industry. There is little overlap between those who study risk measured in dollars and those who measure damage to equipment or loss of human life. Two authors, from the latter community, minimize insurance and portfolio management as risk fields, arguing that "within the professional communities on risk, most analysts would agree that damage to human health and the environment are at the fore of what we call risk analysis and risk management" (Klinke and Renn, 2002:1076). In addition, the study of risk has been largely separate from the study of choice within the academic research (Sarin and Weber, 1993:135).

The concept of risk can have multiple characteristics or qualities. Investors typically imply volatility when using the term risk (Survey of Risk, 2004:9). Depending on the situation, risk may refer to the possible outcomes or consequences, likelihood of occurrence of those outcomes, the significance, causes or affected population (Ayyub, 2003:36). The depth of risk assessment can vary greatly, depending on the available information and the level of detail required. With little data, qualitative risk assessment may be the only possible analysis. With more data available, a quantitative approach can be taken (Bennett *et al.*, 1996:468).

In an effort to provide structure to risk analyses in the public sector, the National Research Council provides four questions for validation of a risk assessment (National Research Council, 2000:5).

- Are the generated measurements complete (collectively exhaustive) and useful to decision makers?
- Are all relevant uncertainties accounted for?
- Are these uncertainties correctly specified?
- Are stochastic and statistical techniques properly implemented?

These questions highlight the two most important components of risk: outcomes and likelihood. The first of those components, the magnitude of the consequences, is a physical measure of severity in dollars lost, equipment damaged or human lives affected. The second is a mathematical construct, the probability that something goes wrong (Haimes, 1998:41). A common mathematical evaluation of risk is the product of these two factors, the likelihood of occurrence multiplied by the impact or severity of the consequence (Ayyub, 2003:37).

The current CRRA approach to risk involves two independent assessments for each identified capability. First, the process determines the current level of capability based on a combination of assessments of *proficiency* and *sufficiency*. The former is the *quality* of existing Air Force capability and the latter is the *quantity* of the existing capability. These two measures are combined to form a single measure of existing capability, which ranges from none (0% capable) to complete (100% capable). Second, Air Force subject matter experts are asked to identify the likely consequences if a scenario occurs that requires the capability, and no capability exists. The estimated severity assessment ranges from "minor" to "catastrophic". These independent assessments of capability and expected severity are combined, using a contour plot to determine a risk score as shown in Figure 2.

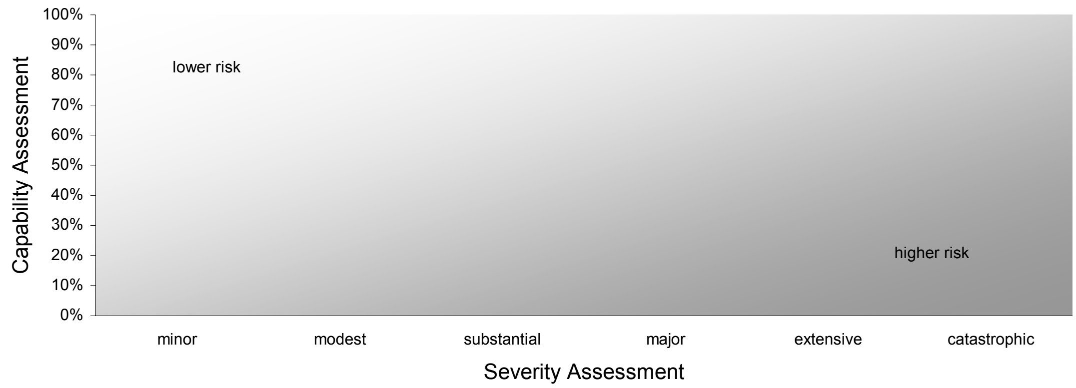

Figure 2. Existing CRRA risk methodology

There are three major factors this approach to risk assessment does not consider. First, the current level of a capability may have an effect on the outcome of a scenario. A higher capability with a mitigating effect, for example, would reduce the resulting severity. Second, the methodology does not allow for the possibility that a capability will never be needed. If a perceived threat does not materialize, an adverse event may never occur and no severity will be experienced, regardless of capability level. Finally, the existing approach does not include the range of possible severities. Estimating future severity involves both uncertain knowledge of threats and natural variability. Without accounting for the affect of existing capability on risk, the possibility that a capability will never be needed, and the variability in outcomes capability shortfalls and redundancies may be incorrectly identified and prioritized.

This research proposes several ways that risk can be handled mathematically to overcome these challenges. Borrowing from engineering, finance and actuarial science, this thesis models risk as a random variable with an associated probability distribution, rather than a single number. This captures the notion that the future severity of outcomes cannot be known with certainty. This distribution can be adjusted based on judgments of how changes in capability affect risk. Finally, using a variety of tools this risk distribution can be measured, summarized into a single number that allows risks to be ranked, prioritized or compared against each other.

1.3. Scope and Limitations

The primary goal of this research is to determine a methodology that will assist Air Force decision makers to order or prioritize risks associated with shortfalls in capabilities. Accurately ordering these risks will point senior Air Force leadership to the areas that require the most focus of future system acquisition or tactics development. A secondary research goal is to explain approaches to risk from different fields, providing military analysts with an expanded toolbox for modeling and measuring risk. Quantifying and measuring the downside risk of capability shortfalls requires projections of future needs and threats; this research suggests ways to add mathematical rigor to that process. The final research goal is to determine an appropriate risk measure and apply it to a system acquisition problem for the optimal allocation of scarce resources.

There are several assumptions that form the foundation of this thesis. First, while this research provides methodological recommendations to the CRRA, it uses only notional numerical data and does not provide any programmatic recommendations. Second, the proposed methodologies add mathematical rigor to the risk assessment portion of the CRRA, but still require subjective estimates of probabilities and severities of future events. Third, this thesis considers only downside risk. All outcomes of capability shortfalls are assumed to be undesirable severities. Upside or positive risk—the possibility that capability exceeds need—is not considered.

Using the proposed methodology, risks can be assessed at any level of the CRRA hierarchy of operational concepts, desired effects, general capabilities and specific tasks. Risk modeling at a higher level probably means a less complicated result, but may mean additional analytic challenges. Risk modeling lower in the hierarchy means a more complicated final product, but may be easier to assess.

This research focuses on the assessment of the risk associated with previously identified and quantified capability shortfalls. It does not propose ways to quantify the current level of a capability or consider whether all capabilities have been correctly specified, emphasizing instead the prioritizing of risks associated with shortfalls already identified.

A general risk analysis process suggested by Haimes involves five distinct steps. First, *risk identification* involves specifying all the imaginable things that could go wrong, particularly noting possible failures in hardware, software, organization or humans. Second, *risk quantification and measurement* requires objective or subjective assessment of the likelihood that the identified events will occur, including interactive and causal relationships. Third, *risk evaluation* develops alternate courses of action with associated costs or tradeoffs. Fourth, *risk acceptance and avoidance* means choosing between alternatives. Finally, *risk management* implements the decision and provides feedback (Haimes, 1998:55-56).

This research focuses on the second and fourth steps in risk analysis: quantification, measurement and evaluation. An overall methodology for the prioritization of Air Force capability shortfall risks involves four steps, shown in Figure 3. The first step, identifying capability level, the *capabilities review* portion of the CRRA process, takes place outside the scope of this research. Quantifying likelihood and severity of adverse effects forms the second step. This may involve an objective or subjective approach or some combination of the two. Presumably intelligence will play a role in estimates of the likelihood of undesirable events and the severity of the consequences. The identified level of friendly force capability should be considered in these quantitative estimates; a higher level of capability may make an event less likely to occur (prevention) or lessen the severity of the outcome (mitigation).

The third step of risk prioritization involves taking the distribution identified in step two and translating it into an appropriate measure or measures of risk. The measure may use the expected or average severity, the variance or dispersion of the amount of severity or other relevant mathematical quantities. This measure, a number rather than a probability distribution, can then be ordered with other measures in step four.

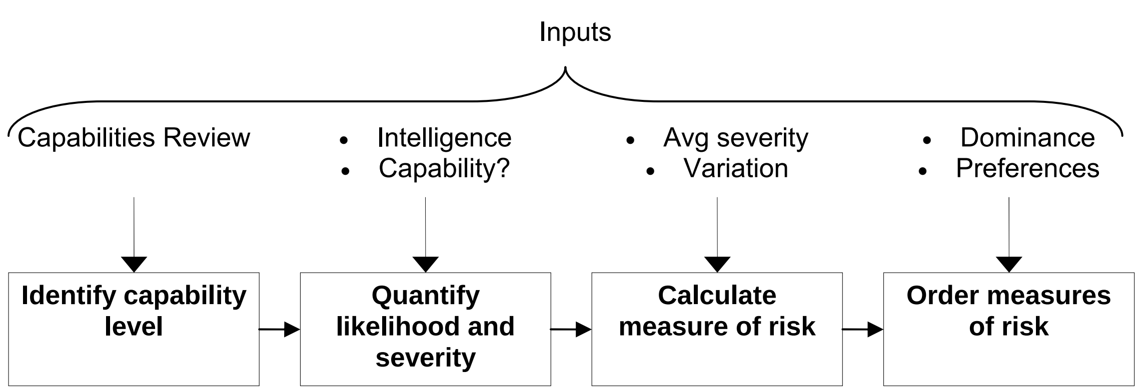

Figure 3. Capability Shortfall Risk Assessment

1.4. Thesis Organization

The remainder of this thesis describes methodologies for the mathematical modeling and measurement of risk in the risk assessment phase of the Capabilities Review and Risk Assessment process. Chapter II provides an overview of the academic literature, including a discussion of the causes and remedies for uncertainty, and an explanation of a variety of risk analysis techniques that may be useful in the assessment risk in a military context. Chapter III explains the factors involved in modeling capability shortfall risk and offers four mathematical risk measures for summarizing risks in a single quantity. Chapter IV examines a set of notional risks based on nine high-level capabilities and shows how risk measures can be used to guide system acquisition decisions. Finally, Chapter V summarizes the results of this research and suggests questions for future study.

II. Literature Review

2.1. Overview of Risk Analysis

Risk analysis plays a prominent role in a number of disciplines, including engineering, decision analysis, statistics, medicine, financial management and actuarial science. While the exact approaches applied to risk vary, some common themes emerge. In general, risk includes some aspect of uncertainty and some aspect of negative consequences. The first goal of risk analysis is to understand—and perhaps reduce—the uncertainty. The second is to understand—and perhaps prevent—the negative outcomes.

There is some disagreement in the academic literature over both aspects of risk. Some analysts claim that a deterministic situation, one with complete certainty, cannot be considered risky. The past, for example, has no risk because all of its uncertainties have been resolved, and risk can only belong to the future (Ayyub, 2003:35). Others describe any situation with a downside or negative outcome as a risk, even if that negative outcome is certain (Fishburn, 1984:397). The Defense Department defines risk as the "probability and severity of loss linked to hazards" (Department of Defense, 2003:459).

While risk commonly implies negative outcomes, some analysts also use the term risk to include positive outcomes as well. This is particularly true in the financial management field, where an investment can have a positive or negative return (Jia and Dyer, 1996:1692). In other contexts, risk is only used to describe negative outcomes and does not refer to success (Ayyub, 2003:35). The Capabilities

Review and Risk Assessment process focuses on the negative side of risk only and has defined risk as "the impact on combat operations if … capability is not available to provide the required effects" (AFSAA, 2003:35). The CRRA definition of risk does not include any reference to probability, but does not explicitly exclude probability either.

For this research, risk will be considered to include any situation with negative consequences, with an emphasis—but not a restriction—on the uncertainty associated with those consequences.

Risk analysts break the process of studying risk into two phases: risk assessment and risk management. Risk assessment seeks to gain an understanding of the factors, outcomes and parameters of the search for answers to three questions (Haimes, 1998:55).

- What can go wrong?
- What is the likelihood of it going wrong?
- What are the consequences?

Risk management seeks to reduce or control risk. Like risk assessment, it has three broad questions (Haimes, 1998:55).

- What options are available?
- What are the costs and benefits?
- What is the future impact?

2.2. Uncertainty

This section describes the causes of uncertainty and some existing approaches to describe uncertainty in verbal and mathematical terms. Risk, though related to uncertainty, is not quite the same thing. Where uncertainty can include any absence of knowledge, risk is more an "educated gamble" (Survey of Risk, 2004:4).

2.2.1. Causes of Uncertainty

Uncertainty can occur for an array of reasons. This section describes some of these reasons, and explains the distinction between uncertainty caused by a lack of knowledge and uncertainty due to natural variability. It then includes a brief overview of game theory, a mathematical approach for understanding uncertainty caused by intelligent opposition.

2.2.1.1. Categorizing Uncertainty

Historically, the term "uncertainty" was used to describe situations for which probability distributions could not be used because of insufficient data. The National Research Council no longer finds this an acceptable definition in the analysis of risk, favoring uncertainty as a more general word to describe any situation in which outcomes are not fully known (National Research Council, 2000:41).

Uncertainty can take many forms, but in general it can be broken into three broad categories: natural variability, knowledge uncertainty and decision model uncertainty (National Research Council, 2000:48). Natural variability (also called aleatory, external, objective, random or stochastic uncertainty) refers to the inherent instability in the physical and human world, the understanding that the same process will not play out the same way every time. Knowledge uncertainty (also called epistemic, functional, internal or subjective uncertainty) refers to the imprecision of our understanding of a system (National Research Council, 2000:42).

Knowledge uncertainty affects calculations in a different way from natural variability (National Research Council, 2000:6). For example, soliciting expert opinion may measure variability but still leaves uncertainty (Kelly and Taylor, 2003:495). Uncertainty about data contributes to knowledge uncertainty. Data uncertainty comes from measurement or transcription errors, sampling that is not representative of the entire population, or a system that is inconsistent or heterogeneous in time or space (National Research Council, 2000:44).

The final category of uncertainty is the decision model. The decision maker may have poorly defined or continuously changing objectives or values which prevent consistent decisions (National Research Council, 2000:42). When the model is uncertain, even complete knowledge and zero natural variability are insufficient for correct insight into the system in question.

2.2.1.2. Uncertainty from an Intelligent Opponent

In a traditional analysis of reliability, engineers assume negative effects follow some probability distribution based solely on the design specifications of the system. Building a more robust system, with stronger parts or redundant components, improves reliability, the probability that the system will continue to function through some time period. In a military context, where damage may occur because of enemy attack instead of random accident, new analyses are necessary. Game theory, which requires decisions against an intelligent opponent, can help to bridge the gap between classical probability theory and a world that faces threats from enemies intent on destruction.

When a threat is natural, the analyst can build a probability distribution of risk on the design of the system in question. When a threat comes from an intelligent source such as an enemy military, however, the probability distribution associated with risk can change over time.

Non-state enemies, such as terrorists, add additional complexity. In some sense, terrorists threaten in illogical and unpredictable ways, because no obvious procedure exists to test for the appropriate safety factor (Smith, 2002:40). However, the management of risk requires the same kinds of tradeoff between cost and productivity whether the system faces an intelligent threat or a random one (Smith, 2002:41).

Two papers from the journal *Military Operations Research* describe ways to incorporate a game theoretic model into a risk analysis of military systems. In a 2002 paper, "Risk Management and the Value of Information in a Defense Computer System," Hamill *et al.* provide a model of threats and protections to an information system. Paté-Cornell and Guikema (2002), in "Probabilistic Modeling of Terrorist Threats: A Systems Analysis to Setting Priorities Among Countermeasures," explain how to separate beliefs from actual capabilities in a model and how to handle learning by both terrorists and those defending against them.

Hamill *et al.* (2002) define risk assessment as the linkage among three factors: threat, vulnerability and impact. Natural or accidental human threats can be modeled with a classical probability approach. That leaves threats that are not accidental, but intended attacks (Hamill *et al.*, 2002:64). These intentional human threats can be modeled with a game-theoretic method to account for possibility that the threat can change—in a rational manner—depending on the defenses set up.

Two approaches allow the defender to identify the payouts and probabilities in the game. The Red Team "hacker approach" to vulnerability assessment involves putting together a team that attempts to break the system. This is equivalent to playing the game multiple times to see if equilibrium can be reached. The advantage of this approach is that it most closely models reality, with actual human decision makers seeking their optimal strategy (Hamill *et al.*, 2002:65). At each iteration the damage to the system (whether sensitive information acquired by the attacker or data destroyed) can be measured, along with the ease or speed with which the attacker gained access. These attacks can be paired with the defensive measures employed to build the two strategy vectors and associated payoff matrix for insight into the risk of damage to the system.

An alternative approach to vulnerability assessment is the "algorithmic approach," which is a "methodical and systematic evaluation" of the system. The advantage of the algorithmic approach is that it may identify threats that the unsystematic hacker approach does not happen to explore (Hamill *et al.*, 2002:65). This is equivalent to attempting to completely identify strategies and payoffs and solve the game theoretically. In practice, a combination of both hacker and algorithmic approaches will generally lead to the greatest understanding of the game parameters.

In another application of game theory to risk analysis, the Paté-Cornell article focuses on building an "overarching model," focused on model structure rather than numbers, to collect information from different sources on threats, potential enemies, possible damage and targets. The game theory aspect of terrorism and counterterrorism comes from its dynamic nature as each side updates priorities with the other side's changes (Paté-Cornell and Guikema, 2002:5-7).

One of the greatest difficulties in analyzing risk in a game theory setting is the sheer number of possibilities. The places and ways an enemy can attack are enormous. When combined with the number of ways to deter or mitigate damage, the problem—at least at the strategic level—is unmanageable. The authors propose a model that attempts to cut through some of the problems with size by combining possible outcomes. They suggest that every event or severity random variable (risk) in their model can be analyzed at a more detailed level if desired (Paté-Cornell and Guikema, 2002:5).

As with Hamill's approach, Paté-Cornell assumes that the model of enemy strategy requires separate assessments of capability and motive. When modeling multiple enemies (for example, different terrorist groups), each enemy may have a different combination of these two factors (Paté-Cornell and Guikema, 2002:7). This does not necessarily mean that *n*-person game solution methodologies are required, however. Because the defender is not (presumably) forming coalitions with some terrorist groups against others, these are a set of two-person games rather than a single *n*-person one. Either all attackers can be lumped together as a single opponent, accounting for any synergies the various attackers gain from each other, or defending against each opponent can be considered a separate game.

Clearly, the actual employment of game theory in the assessment of risk is difficult. The number of strategies available to a potential attacker is immense, and the number of strategies to deter or mitigate risk is also large.

2.2.2. Quantification of Uncertainty

In understanding and modeling risk, uncertainty must be translated into probability. In well-defined, well-understood situations, probabilities can be determined directly. For example, it is clear and widely-understood that a fair coin has probability 0.5 of landing heads and probability 0.5 of landing tails. In other, less intuitive situations, probabilities can be estimated based on empirical data. When historical data is available for risks, objective probabilities can be estimated. Typically, however, sparse historical databases lead away from objective probabilities in risk assessment to subjective probabilities based on expert judgment (Haimes, 1998:138). This section describes approaches and methods to determine these subjective probability estimates.

In the context of risk, there can be uncertainty in both outcomes and probabilities of those outcomes. Decision makers may find it helpful to break their problem into four classes: probabilities and outcomes known, probabilities uncertain and outcomes known, probabilities known and outcomes uncertain, or both probabilities and outcomes uncertain. (Langewisch and Choobineh, 1996:140)

A linear mathematical program involves an objective function to maximize or minimize subject to a set of linear constraints defining a set of feasible solutions. In the standard form of this model all parameters must be known. Eum, Park and Kim (2001) provide a set of linear programming tools to handle simultaneous uncertainty about weighting and value scoring in a multi-criteria decision analysis. The authors suggest several conditions, besides exact estimates, that can be used to define weighting and value scoring (Eum *et al.*, 2001:399). Based on whether weighting, value scoring or both weights and values are uncertain, the authors show how a linear programming model can identify dominated and potentially optimal alternatives (Eum *et al.*, 2001:405). This tool could be used for prioritizing risks when relatively little is known about the likelihood of various outcomes.

It is possible to distinguish between decision making "under partial information" and the sensitivity of a decision. The former involves imprecisely specified weights. The latter includes exact weights but a decision maker uncertain about which factors are important and interested in refining those original "exact" estimates. Even if a decision is robust in its parameters, sensitivity analysis is invaluable in helping the decision maker understand the problem (Rios Insua and French, 1991:177). Some other authors suggest a Bayesian approach to handle uncertain parameters. However, when parameters are considered as random variables with probability distributions, there may be more imprecision from the new distributions than additional benefit to the model. An iterative process of the decision maker making judgments and the analyst performing sensitivity analysis may be a more appropriate approach (Rios Insua and French, 1991:180).

Choobineh and Behrens (1992) caution against assuming too much about the underlying probability distribution of a random variable. One alternative to fitting a theoretical probability distribution is to use an interval distribution. An interval distribution makes no assumption about the probabilities of any outcome other than to place upper and lower bounds. An alternative to the interval distribution is a possibility distribution. Possibility distributions essentially take multiple intervals, rather than a single interval, to allow for a gradual decrease in possibility (Choobineh and Behrens, 1992:910). Figure 4 shows an interval distribution, where a parameter is equally likely to take on any value within the range, and a possibility distribution, where the parameter has the same expected value as the interval distribution, but is less likely to take on values at the extremes.

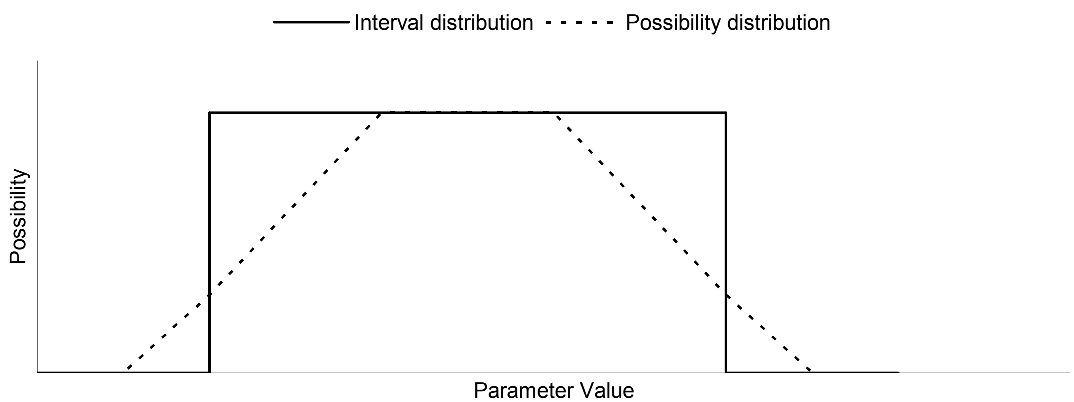

Figure 4. Interval and Possibility Distribution (Choobineh and Behrens, 1992:910)

A final approach to the quantification of uncertainty is the translation of subjective, verbal expressions of likelihood into numerical probabilities. There is large disagreement over what is meant by, for example, an infrequently occurring event. For any verbal to numeric translation, the only consistency is that "unlikely" means less than 0.5 probability and "likely" means greater than 0.5. The context of the verbal description has a large effect on the numerical translations. In a situation in which negative consequences occur very infrequently, a high probability may still be much closer to zero than to one.

Table 1 shows three verbal-numeric translation schemes. In the first, sixteen risk managers were surveyed for their interpretation of different phrases describing likelihood (Jablonowski, 1994:52). The average response of the sixteen individuals is shown, along with the range of their responses. The ranges overlap, except for the gap between "somewhat likely" and "likely".

The second set of verbal-numeric translations comes from an engineering setting, with a failure rate describing the occurrence of system failure per cycle or per unit of time (Ayyub, 2003:61). Because system failure is a rare event, these probabilities are much closer to 0 than to 1, with "high" occurrence, for example, equivalent to probabilities of 0.025 to 0.05.

The third set of translations comes from a military regulation (MIL-STD-1629A, 1980:section 3.1). These translations are also intended for engineers, but are expressed independent of time, as simply the probability of system failure.

Table 1. Verbal to Numeric Probability Translations

Description	Average	Range	Occurrence	Failure Rate	Description	Prob of Failure
Rare	.05	.01-.15	Minor	<1 in 1,000,000	Extremely unlikely	0.001
Very unlikely	.10	.03-.25	Low	1 in 20,000 to 1 in 4000	Remote	0.001-0.01
Unlikely	.19	.09-.30	Moderate	1 in 1000 to 1 in 80	Occasional	0.01-0.10
Somewhat unlikely	.26	.09-.45	High	1 in 40 to 1 in 20	Probably	0.10-0.20
Likely	.77	.52-.98	Extreme	1 in 8 to 1 in 2	Frequent	>0.20
Frequent	.78	.60-.90				
Extremely likely	.93	.85-.99				

2.2.3. Probability Distribution Tails

Sparse data in the extreme values of a probability distribution can make fitting a correct distribution a difficult task. In some cases having a small amount of data can be particularly dangerous since it results in too confidently fitting a distribution that does not accurately represent the true distribution. Even when a distribution is fit using a significant amount of data it should be subjected to sensitivity analysis of its parameters (Bratley *et al.*, 1987:125).

When a lack of data in the tail of a distribution does not allow a theoretical distribution to be fit, an exponential tail is a reasonable approximation. This can be adjusted in sensitivity analysis with various Weibull distributions (Bratley *et al.*, 1987:133).

Alternatively, the distribution tail can be fit using the statistics of extremes, which is the mathematical study of the largest (or smallest) values a random variable can assume. The statistics of extremes identifies three forms of probability distribution tails, depending on the type of data. A Gumbel distribution, with cumulative distribution function $H(x) = \exp(-e^{-x})$, allows tails in both the positive and negative domains. The exponential, lognormal and normal distributions all fall into the Gumbel family. A Weibull form only works when the domain of the random variable is negative and has cumulative distribution function of the form $H(x) = \exp[-(-x)^{\gamma}]$. Uniform and triangular distributions follow this Weibull form. The final tail distribution is the Frechet approximation with cumulative distribution function $H(x) = \exp(-x^{-\gamma})$. The Frechet form can only be used when the domain is positive. The Pareto distribution is an example of a distribution that falls in the Frechet family (Lambert *et al.*, 1994:734).

2.3. Existing Approaches to Risk Modeling

A great deal of the risk analysis literature deals with specific techniques for risk assessment and mitigation in mechanical or biological systems. Some of the approaches are more general, however, and may be useful in a military context. This section overviews these risk approaches, providing a basic definition of the technique, the context in which it has been used, the inputs required to implement as well as the outputs generated, and some of the advantages and disadvantages relative to other tools.

2.3.1. Engineering Approaches to Risk

Engineering risk analysis focuses broadly on breaking complex systems into more easily understood parts. The most general of these approaches is reliability assessment. Other tools or techniques used by engineers to assess risk include hierarchical holographic modeling, which emphasizes the different perspectives experts bring to an analysis, the partitioned multiobjective risk method, which simplifies a risk distribution into multiple risk measures, and impact intensity, which multiplies different risk factors into a single number.

2.3.1.1. Reliability Assessment

In the engineering community, reliability is a major field of risk analysis. The study of reliability involves the analysis of complex systems to identify their chance of failure over time. In general, reliability analysis focuses on breaking a system into smaller components which are more easily understood.

Several concepts are available to express system reliability quantitatively. Reliability itself is generally modeled as a function of time. The function value is the probability a system continues to work, under specified conditions, for a specified period of time (Ebeling, 1997:5). This function, the complement to a cumulative distribution function, is called the survival function denoted by $S(t)$.

Reliability is often expressed as the mean time to failure (MTTF) or mean time between failures (MTBF), numbers which are calculated as the average of the survival function. For instance, the mean time to failure is

$$\text{MTTF} = \int_0^\infty S(t)dt\,. \tag{1}$$

Other reliability measures include *maintainability*, where systems are analyzed for both their time to failure and the subsequent time for repair, and *availability*, the proportion of time a system is working in the long run (Ebeling, 1997:6).

Several tools are available to help an engineer identify and quantify the possible failure of a complex system. Preliminary hazards analysis is a first step in reliability assessment. This is a non-mathematical approach to identify the elements of a system or events in a process where something could go wrong (Henley and Kumamoto, 1981:21). A more detailed, systematic approach is failure mode and effect analysis (FMEA). This is an iterative, bottom-up process to identify the ways (modes) a system can fail, explaining the causes and quantifying the probabilities of occurrence (Ebeling, 1997:167). FMEA is widely-used and well-accepted in the engineering community. The primary disadvantage of this approach is its tendency to ignore combinations of problems that together lead to failure, even though independently they are not dangerous (Henley and Kumamoto, 1981:40).

Fault tree analysis is a graphical view of the ways and causes of system failure. Figure 5 shows an example fault tree. Where FMEA starts at the lowest possible component level of a system to analyze reliability, fault tree analysis is a top-down approach that focuses on events rather than system components. The top event in the tree is the event of the overall failure of the system. The tree then breaks down this overall failure into all the possible resultant events that cause the overall failure. A series of logical AND and OR "gates" are used to show when all resultant events are required for a top event or if a single resultant event is sufficient. At the bottom level of the tree are "basic events" which are not analyzed in further detail. When these basic events have probabilities attached to them, the overall system failure probability can be calculated.

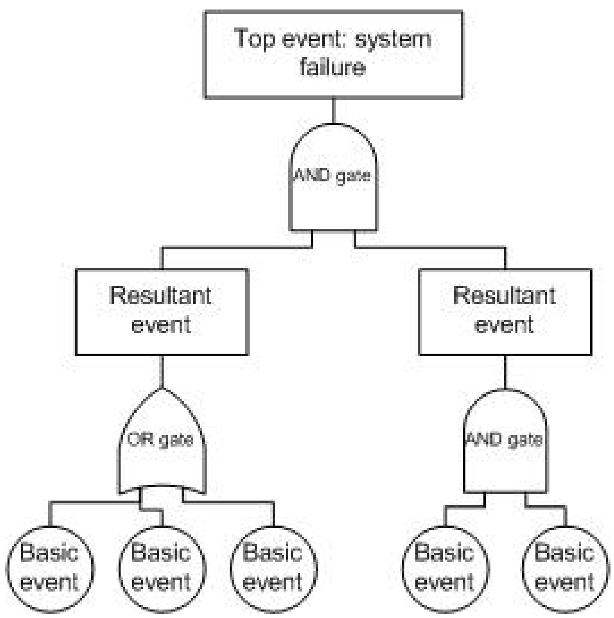

Figure 5. Fault Tree

The advantage of a fault tree approach is its flexibility in level of detail. Component or system failures may be decomposed into extremely precise, detailed events, or kept at a simple level. Unfortunately, this means that fault trees can grow large and complex very quickly. Because they do not (necessarily) visually match the system or process, even individuals familiar with the system may have difficulty following them (Henley and Kumamoto, 1981:40).

Criticality analysis is a quantitative tool to prioritize system components based on their relative importance to the overall system. After a failure mode and effect analysis has identified all of the ways (modes) a system can fail, a criticality index number can be calculated for each component as the product of three factors: the conditional probability of damage given a particular failure mode occurs, the rate of occurrence of the particular failure mode and the time period being analyzed. Summing over all failure modes affecting the component in question results in an index number for ranking component criticality (Henley and Kumamoto, 1981:34; Ebeling, 1997:170).

2.3.1.2. Hierarchical Holographic Modeling

Hierarchical holographic modeling is a tool to identify risks in large-scale, complex systems. The goal of the method is to take advantage of multiple expert views of the system in order to provide different perspectives on the vulnerabilities and hazards in the system. The approach requires examining the overall system from different, overlapping perspectives: time, economics, geographical, legal, and environmental, for example (Haimes, 1998:98). Hierarchical holographic modeling has been used to identify risks in energy utilities, water resource systems, sustainable development projects and system acquisition (Haimes, 1998:99-108).

To identify risks, a hierarchical structure of factors based on the different perspectives under consideration must be built. Experts can then provide subjective input at the different levels of the hierarchy where they have expertise. Sparse historical databases lead away from objective probabilities in risk assessment to subjective probabilities based on expert judgment (Haimes, 1998:138). The primary advantage of this modeling technique is that it allows expert opinions to overlap; the elements of the hierarchy do not have to be mutually exclusive (Haimes, 1998:95).

2.3.1.3. Partitioned Multiobjective Risk Method

The partitioned multiobjective risk method (PMRM) splits the risk distribution into two or more sections and calculates the conditional expectation, given that severity falls with each of these sections. In addition, the overall expected severity is calculated. This results in at least three numbers, which serve as measures of the risk. The method is used for multiobjective risk analysis problems, and each of the conditional expectations, plus the overall expectation, are used as quantities in a multiobjective decision framework (Haimes, 1998:312).

Partitioning is a subjective exercise and there is no general rule for selecting the points at which to split. Partitioning can be done on the severity axis or the probability axis. That is, the n partitions can be defined by severities β_i such that $0 < \beta_1 \leq \beta_2 \leq \ldots \leq \beta_{n-1} \leq \infty$. Alternatively, the n partitions can be defined by probabilities α_i such that $0 < \alpha_1 \leq \alpha_2 \leq \ldots \leq \alpha_{n-1} \leq 1$ (Haimes, 1998:315).

PMRM requires the entire risk distribution to be known, but takes advantage of that knowledge by calculating measures over the entire distribution. By calculating multiple measures, the method retains information lost in other techniques that do more simplification. The disadvantage of this is that multiple measures do not allow risks to be easily ranked.

2.3.1.4. Impact Intensity

The basic concept of impact intensity is to identify a number of risk factors such as likelihood of occurrence, expected severity, chance of detection or expense of mitigation, and assign an index value to each of these factors. This models risk as an n-dimensional vector. An "impact intensity" or "risk prioritization number" can then be calculated in several ways using the values in this vector.

The first formulation is the linear multi-attribute value function where each factor receives a score between 0 (low risk) and 1 (high risk) and a relative weight (Cho *et al.*, 1997:26).

$$\text{Impact Intensity} = \sum_{i=1}^{n} weight_i value_i \qquad (2)$$

Alternatively, impact intensity can be calculated as a multiplicative function that results in a maximum score when any single factor is at its maximum, similar to the calculation of reliability in a parallel components system (Cho *et al.*, 1997:27).

$$\text{Impact Intensity} = 1 - \prod_{i=1}^{n} \left(1 - value_i\right)^{weight_i} \qquad (3)$$

A simpler impact intensity function involves multiplying the factor scores together, a calculation like the system reliability of components in series (Ayyub, 2003:62).

$$\text{Impact Intensity} = \prod_{i=1}^{n} value_i \qquad (4)$$

Table 2 shows an example of the three impact intensity functions with four risk factors scored at low (0.0), medium (0.5) and high (1.0) and all factors are equally weighted. Under Equation (3) when any factor scores a one, the impact intensity is one. With Equation (4), when any factor scores a zero, the impact intensity is zero.

Table 2. Example of Impact Intensities

Factor Scores	Eq (2)	Eq (3)	Eq (4)
(0.0, 0.0, 0.0, 0.0)	0.00	0.00	0.00
(0.0, 0.0, 0.0, 0.5)	0.25	0.50	0.00
(0.0, 0.0, 0.5, 0.5)	0.35	0.75	0.00
(0.0, 0.5, 0.5, 0.5)	0.43	0.88	0.00
(0.0, 0.0, 0.0, 1.0)	0.50	1.00	0.00
(0.5, 0.5, 0.5, 0.5)	0.50	0.94	0.06
(0.0, 0.0, 0.5, 1.0)	0.56	1.00	0.00
(0.0, 0.5, 0.5, 1.0)	0.61	1.00	0.00
(0.5, 0.5, 0.5, 1.0)	0.66	1.00	0.13
(0.0, 0.0, 1.0, 1.0)	0.71	1.00	0.00
(0.0, 0.5, 1.0, 1.0)	0.75	1.00	0.00
(0.5, 0.5, 1.0, 1.0)	0.79	1.00	0.25
(0.0, 1.0, 1.0, 1.0)	0.87	1.00	0.00
(0.5, 1.0, 1.0, 1.0)	0.90	1.00	0.50
(1.0, 1.0, 1.0, 1.0)	1.00	1.00	1.00

Impact intensity offers the advantage of summarizing risk in a single number, but allowing components of that risk to be easily highlighted. A risk can be considered critical if it has a high overall score or if any single component score is above some threshold. Risks with impact intensities are easily ranked, because the technique reduces the complexity of multi-dimensional risk to a single number.

The method does not, however, allow or account for any variability. All scores are deterministic. Depending on the equation selected to calculate intensity, the result might inappropriately focus on a risk with a high single component score that is not really dangerous overall, or ignore a truly important risk with moderate component scores that combine to a low score overall.

2.3.1.5. Farmer Curve

The Farmer curve is a graphical tool to display tradeoffs in risk and indicate risk acceptance. Risk acceptance is an acknowledgement of the existence of the possibility of adverse effects and a willingness to live with the situation. It was originally employed to explain the risk of radioactive release from nuclear power plants (Henley and Kumamoto, 1981:13).

The curve, shown in Figure 6, plots frequency versus severity. The curve is the maximum acceptable level of risk. Scenarios that are more likely or more severe than the curve are deemed unacceptable risks. That is, points above or to the right of the curve are unacceptable. Points below or left of the curve, representing less likely or less severe scenarios, are classified as acceptable risks and do not require mitigating resources.

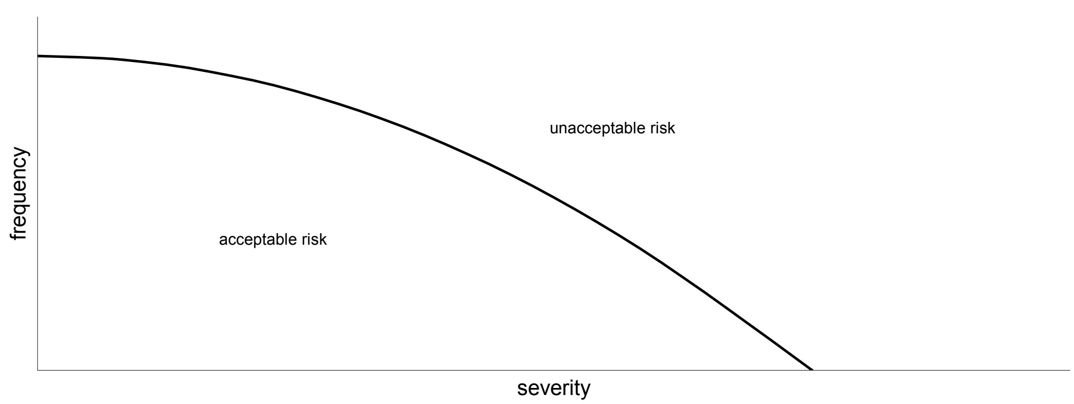

Figure 6. Farmer Curve

2.3.1.6. Precautionary Principle

The precautionary principle comes out of the field of toxicology and essentially states that no level of risk is acceptable. If there is any chance that a substance will cause damage to human beings or the environment, efforts should be taken to completely eliminate its release (Klinke and Renn, 2002:1071).

Because no risk is deemed acceptable under the precautionary principle, only two factors need to be considered: the most catastrophic possible outcome and the cost of risk management. Decision makers should seek the biggest bang for the buck in risk mitigation. Two principles used in practice are "as low as reasonable" and "best available control technology" (Klinke and Renn, 2002:1071).

The precautionary principle has an advantage of simplicity, since it does not require detailed assessment of possible outcomes or likelihoods attached to particular severities. Unfortunately, it is often unrealistic to completely eliminate risk and ignoring the probability distribution associated with various outcomes may result in a poor allocation of resources.

2.3.2. Decision Analysis Tools

Decision analysis is concerned with selection between multiple competing alternatives. Multiple criteria, as well as multiple alternatives may be part of the problem. In a risk assessment situation, the different risks can be considered the alternatives, and the decision tools could help in the ranking or prioritization of these risks.

2.3.2.1. Non-Parametric Decision Rules

When probabilities are completely inestimable, Fleischer suggests that five different decision rules are available to rank alternatives. Although these rules do not require quantification of likelihood, there still must be analysis of all possible alternatives and outcomes. Each of these rules will not necessarily give the same answer, but they will help frame the decision (Fleischer, 1984:292).

The *minimax rule* is the extreme pessimistic approach that assumes the worst possible outcome will happen. Among all alternatives, select the one with the best (minimum) of the worst (maximum) possible costs. If the problem is concerned with gains rather than losses, the equivalent rule is *maximin*, that is, selection of the alternative with the best (maximum) of the worst (minimum) possible profit (Fleischer, 1984:286).

The *minimin rule* is the opposite approach to minimax, taking instead an extreme optimistic approach that assumes the best possible outcome will happen. Among all alternatives, select the one with the best (minimum) of the best (minimum) possible costs. Again, if the problem is measured in gains instead of losses, the rule is *maximax*, selection of the alternative with the highest possible profit (Fleischer, 1984:287).

The *Hurwicz rule* takes a middle ground between extreme optimism and extreme pessimism. This rule, named after econometrician Leonid Hurwicz, involves a linear combination of the best and worst possible outcomes for each alternative. Multiply the worst possible outcome by the "index of optimism," a number α between 0 and 1, and multiply the best possible outcome by $(1 - \alpha)$. The sum of these two numbers can be compared across alternatives. If α=0, the Hurwicz rule is equivalent to minimax and if α=1 it is equivalent to minimin (Fleischer, 1984:288).

The *Laplace rule*, named after mathematician Pierre Simon de Laplace, assumes that all outcomes are equally likely. The rule says to calculate the expected value of these equally likely outcomes, and select the alternative with the best expected value (Fleischer, 1984:288).

The *Savage rule*, also known as the principle of *minimax regret*, seeks to minimize the difference between the actual outcome and the outcome if the future had been correctly forecasted. This difference is the decision maker's "regret." In order to apply the Savage rule, named after statistician L.J. Savage, calculate a *regret matrix*, where each row is a different alternative and each column is a different "state of nature." Each entry in the matrix is the difference between the outcome of that combination of alternative and state and the best possible outcome in that state of nature. Select the alternative with the smallest maximum regret value. The most significant disadvantage of the Savage rule is that adding an additional alternative can shift the answer, even if the new alternative is not preferred (Fleischer, 1984:291).

2.3.2.2. Lexicographic Method

The lexicographic method is a technique to rank different alternatives under multiple criteria. In the context of risk, criteria could be the worst possible outcome, a chance of any adverse event occurring or the most likely outcome. The decision maker first ranks all of the attributes from most important to least important. Each alternative is then scored for the most important attribute. Alternatives that meet some acceptability threshold according to the most important attribute are then scored for the next most important attribute. This process continues until only one alternative remains, or the alternatives have been scored for every attribute and a set of possible solutions remains (Chankong and Haimes, 1983:200).

The primary advantage of this approach is its cognitive simplicity. It does not require that every alternative be scored for every criterion and it does not require a precise score for each alternative, only a decision on whether the acceptability threshold has been met. In addition, the lexicographic method closely relates to the way individuals make decisions in practice, focusing on the single most important attribute to screen alternatives rather than examining all alternatives with all attributes simultaneously (Chankong and Haimes, 1983:200). The primary disadvantage of the lexicographic method is its emphasis on ranking the attributes. An alternative that scores low in the single most important attribute but is superior in every other category may be eliminated even though it is important (Chankong and Haimes, 1983:205).

2.3.2.3. ELECTRE Method

The ELECTRE method is a tool for multiobjective decisions where the number of alternatives is relatively small and the value of each alternative is known with certainty. According to Chankong and Haimes, the method was first proposed by Bernard Roy in 1968. The method can result in a preferred alternative, or a preferred class of alternatives. To implement the technique each alternative is compared to the others and assigned an "outranking" relation, specifying that one alternative is preferred to another. These relationships can be displayed in a directed graph (Chankong and Haimes, 1983:205-6).

ELECTRE does not require that every set of alternatives be comparable, though every comparison and outranking relation adds strength to the assessment (Chankong and Haimes, 1983:208). The primary disadvantage of the ELECTRE method is its requirement of certainty in outcomes.

2.3.3. Risk Measurement

The purpose of quantifying risk and simplifying that quantification in a risk measure is to order different risks and, ultimately, to choose between them. The quantification requires the probability distributions of the risk and the risk measure requires a preference function for those distributions (Landsman and Sherris, 2001:103). Increasing risk can mean one of two things: that bad outcomes are becoming more likely or that likely outcomes are getting worse (Fishburn, 1984:397). Risk measurement seeks to combine both of these aspects into a single number.

2.3.3.1. Need for Risk Measures

In the simplest case one risk stochastically dominates another and specific probability distributions of risks need not be known in order to rank risks. Under the risk-return dominance property, a gamble with a higher (expected) value *and* a lower risk will always be preferred (Sarin and Weber, 1993:136).

The simplest measure of risk (to understand) is the expected severity. There is significant danger of conflating events with high probability of occurrence and low cost with events of low probability of occurrence and high cost through simple expected value comparisons because the catastrophic outcomes that could occur may be too high to bear no matter how small the probability (Haimes, 1998:17).

2.3.3.2. Properties of Risk Measures

Several authors identify a list of properties or attributes that should be considered when selecting a risk measure. Sarin and Weber argue that a risk measure should increase when the range or variance of severity increases, when a constant is added to all severity outcomes, when outcomes are multiplied by a constant greater than 1 or when a gamble is repeated multiple times (Sarin and Weber, 1993:138).

Landsman and Sherris proposes four properties of risk measures: risk aversion, diversification, additivity and consistency. The risk aversion property means that a risk measure will be greater than or equal to the expected value. Under risk-neutrality the risk measure is the expectation. The diversification property means that multiple small risks should be preferred to a single large risk. The additive property means that a risk measure of the sum of risks is equal to the sum of risk measures. Finally, consistency applies to risks with positive and negative outcomes and implies that if one risk (of loss) is preferred to another, equivalent gains should have the same preference ordering (Landsman and Sherris, 2001:105).

In a 1999 paper regularly cited in the literature, Artzner, Delbaen, Eber and Heath, outline four desirable characteristics of a risk measure. When a risk measure meets these four axioms, it is considered a "coherent" risk measure (Artzner *et al.*, 1999:210).

The first property of coherence is *translation invariance*, which means that any constant added to a risk changes the risk measure by a corresponding amount (Artzner *et al.*, 1999:209). Expectation, for example, has translation invariance since for any random variable X and constant α.

$$E[X+\alpha] = E[X] + \alpha. \quad (5)$$

Variance, on the other hand, does not have the translation invariance property since

$$\mathrm{Var}[X+\alpha] = \mathrm{Var}[X]. \quad (6)$$

Any risk measure then that includes variance, or a function of variance like standard deviation, cannot be a coherent risk measure. Figure 7 shows the probability density functions of two risks, identical except for a constant shifting one to the right. Under translation invariance the shifted risk should have a higher risk measure.

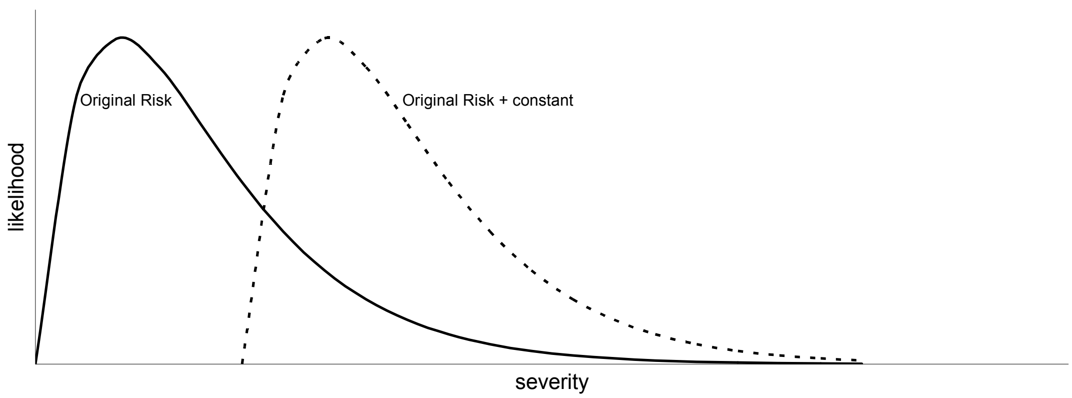

Figure 7. Translation Invariance Axiom Illustration

The second property is *subadditivity*, which means that the measure of any two risks together must be less than or equal to the sum of the measures of the two risks. This property ensures that a single large, unacceptable risk cannot be separated into two smaller, acceptable ones (Artzner *et al.*, 1999:209). The potential problem with this property is that it does not allow for the possibility that putting two acceptable risks together may create a situation with unacceptably large risk.

Figure 8 shows the probability density functions of two independent risks, X and Y, and a third, $Z = X + Y$, which is the sum of the first two. Under subadditivity, the risk measure of Z must be at least as large as the sum of the risk measures of X and Y.

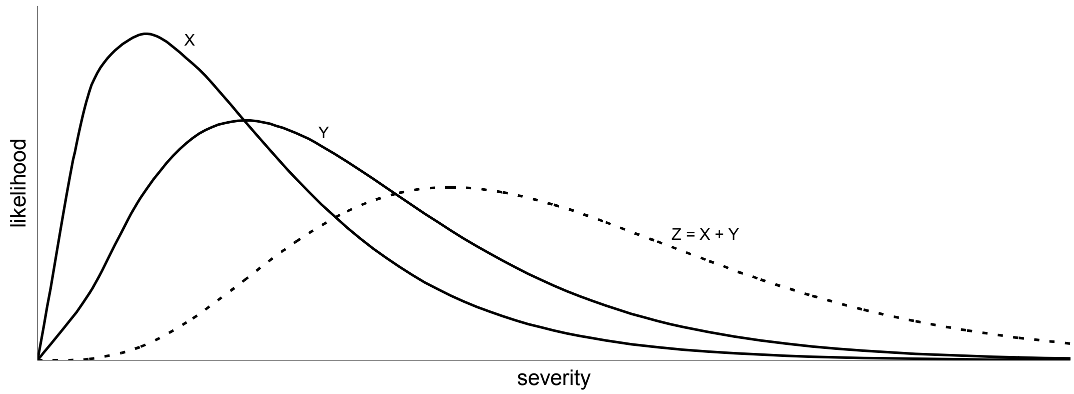

Figure 8. Subadditivity Axiom Illustration

The third coherence property is *positive homogeneity*, which states that if a risk is multiplied by a positive constant, its risk measure must also be multiplied by the positive constant. This property guarantees that the measure of risk increases proportionally to the risk.

Figure 9 shows the probability density functions of a single risk, X, and two other risks, $Y = 2X$, and $Z = 3X$. Under positive homogeneity, the risk measure of the Y must be exactly twice the risk measure of X and the risk measure of Z must be exactly three times the risk measure X.

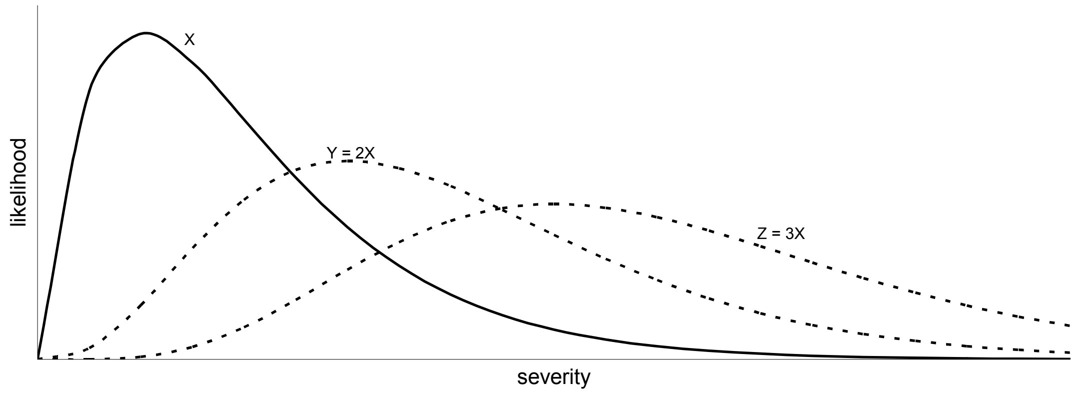

Figure 9. Positive Homogeneity Axiom Illustration

The final coherence axiom is *monotonicity*, which states that larger severities should result in larger risk measures. Under monotonicity, then, a larger risk measure implies a riskier situation (Artzner *et al.*, 1999:210). This axiom implies that if one risk stochastically dominates another risk, it will have a larger risk measure.

2.3.4. Finance and Actuarial Science Approaches to Risk

A second broad area of risk modeling comes from financial management and actuarial science. These fields have the advantage of dealing with dollars, so severities tend to be more easily quantifiable. In finance, the general approach to risk-modeling is risk-value theory. Actuarial science uses two different models, the individual model which sums policy claims in a single time period and the collective model which traces claims over multiple time periods.

2.3.4.1. Risk-Value Modeling

In every application of risk, risk modeling attempts to capture two characteristics of the system in question. First, there is an aspect of variation. The less certainty in the outcome of a situation, the more risky it is considered. Second, there is an aspect of "badness" in which more severe outcomes are more risky than less severe ones (Sarin and Weber, 1993:139). Trading off these two aspects against each other is the basic premise of risk-value modeling.

In risk-value modeling risks are considered gambles, with each risk a random variable that can have both positive and negative outcomes. In a financial application where the risk is associated with the final wealth of some investment, the literature proposes variance as one possible measure of risk. This measure explicitly ignores expectation, with risk measured purely on the spread of the outcomes. The problem with variance as a sole measure of risk is that an investment with an increasing variance in the direction of *increasing* wealth will intuitively be less risky (Mitchell and Gelles, 2002:109).

The first significant work in risk-value was done by Markowitz in the 1950s. He proposed "semi-variance" as a measure of risk. Semi-variance is the variance of the risk random variable in the downside or worst outcome tail (Markowitz, 1959:189). Generally the expected value is used to define the start of the tail, but any arbitrary point can be used (Estrada, 2003:10). The semi-variance of a risk X (where larger values correspond with worse outcomes) with probability density function $f(x)$ and expected value μ is calculated as

$$\mathrm{E}[\min(X\text{-}\mu,0)^2] = \int_{\mu}^{\infty}(x-\mu)^2 f(x)dx\,. \qquad (7)$$

"Downside standard deviation" can be calculated by taking the square root of Equation (7) (Estrada, 2003:10).

Alternatively, variance can be combined with expected return in a linear combination (Sarin and Weber, 1993:137). This allows for the intuitively pleasing property that "mean preserving spreads" or lower expected returns correspond to higher risk. A mean preserving spread is a change in a random variable that has no effect on the expectation, but increases the variance (Mitchell and Gelles, 2002:110). The linear combination measure, as a tradeoff between the expectation and the variation, allows for the fact that an improvement in expected value may result in a willingness to accept more uncertainty.

Some authors propose going beyond expectation and variance to use higher order moments of the risk distribution in order to measure risk (Sarin and Weber, 1993:138). While these measures, such as $E[X^{\theta}]$ where θ is a parameter to be varied by the decision maker, add complexity to the analysis, it is not clear what additional insight they provide into the riskiness of a situation.

Another risk measurement from finance is the risk premium, the difference between the expected payoff and the amount an individual is willing to accept with certainty. Calculation of the risk premium requires knowledge of not only the distribution of the outcomes of the gamble, but also the utility associated with different outcomes (Sarin and Weber, 1993:139).

Recent work in risk-value theory goes beyond distribution moments to incorporate utility theory. This allows risk judgment and preferences to be incorporated into the modeling of risk (Sarin and Weber, 1993:135). Using utility in a risk-value model accounts for two different factors that may influence individual ordering of risks. First, individual preferences may result in different individuals ranking identical risks in different ways. Second, the number of times that a risk will be experienced affects uncertainty. In one formulation of risk-value using utility, the "standard risk", X', is the original risk random variable adjusted so that the expected value is 0, that is, $X'=X-E[X]$. The utility that the decision maker attaches to this new random variable is a measure of risk independent of expected return (Jia and Dyer, 1996:1692).

2.3.4.2. Value at Risk

Similar in name but not otherwise related to risk-value modeling, value at risk is the most common shorthand description of risk in financial applications. Value at risk focuses on risks over time and is an estimate of the maximum amount of loss possible for a given investment (Sarma *et al.*, 2003:339). Various forecasting techniques are used to calculate value at risk. The "delta-gamma" model, for example, can be used when risks in an investment portfolio are quadratic and normally distributed (Castellacci and Siclari, 2003:530). Another approach uses the "autoregressive conditional heteroskedastic" (ARCH) model (Cabelo Semper and Clemente, 2003:516). Regardless of the particular forecasting method used, once an estimate of the value at risk has been calculated, a likelihood is attached to this figure (Cabelo Semper and Clemente, 2003:517).

2.3.4.3. Individual Actuarial Model

The individual risk model considers each insurance policy as a unique random variable. In a specified time period, some of the policies will have no claims and the others that do have claims will vary in size according to some probability distribution.

The total dollar amount of claims is a random variable $C = X_1 + X_2 + X_3 + \ldots$ where X_i are independent random variables of the claim sizes of different insurance policies (including the possibility of no claim). Time is not included in the model (Kaas *et al.*, 2001:28).

The model requires two different inputs. First, the probability that each policy will have a claim in the time period in question must be known. Second, the distribution of the size of claims must be known. With these two inputs, the distribution of C can be calculated using convolution or numerical approximation techniques (Kaas *et al.*, 2001:20).

The individual actuarial model intuitively matches the real world, since the claim size of each policy is represented as its own random variable. It has the disadvantage of only considering one time period. The assumption of independence between the different policies can be inappropriate when, for example, a fire in an apartment building results in multiple claims from several different policies (Kaas *et al.*, 2001:28).

2.3.4.4. Collective Actuarial Model

An alternative actuarial approach is the collective risk model, in which an insurance portfolio is viewed as a stochastic process. In this model claims occur at random time intervals and the size of each claim follows some probability distribution. The total dollar amount of claims, $C(t)$, is a function of the number of claims through time t, $N(t)$, and the size of each claim.

$$C(t) = X_1 + X_2 + X_3 + \ldots + X_{N(t)} \tag{8}$$

where the X_is are independent, identically distributed random variables of claim size over time. The number of claims over a time interval is often modeled according to a Poisson or negative binomial distribution (Kaas *et al.*, 2001:45).

Implementation of the collective model requires the frequency with which claims are filed and the distribution of claim sizes. The model assumption of independence between the number of claims and the size of each claim may conflict with reality, where a single catastrophic event can result in a large number of claims of large size, but in practice the collective model seems to work well (Kaas *et al.*, 2001:46).

The canonical risk model in actuarial science assumes that insurance claims arrive according to a Poisson process and that claim sizes are independent and identically distributed. The resulting stochastic process is modeled as a surplus process, $U(t)$, representing the wealth of the insurer at time t, given by

$$U(t) = u + pt - \sum_{i=1}^{N(t)} Y_i \,. \tag{9}$$

The initial wealth of the insurer is represented by u. Premium payments arrive at a constant rate, p, and the claims, Y_i, follow some general distribution with mean size, β. The number of claims in time t, $N(t)$, is a Poisson random variable with rate λt. "Ruin" occurs when the total claims paid-to-date, $C(t)$, is greater than the sum of initial wealth and premiums paid-to-date. Figure 10 shows an example surplus process (Vázquez-Abad and LeQuoc, 2001:71).

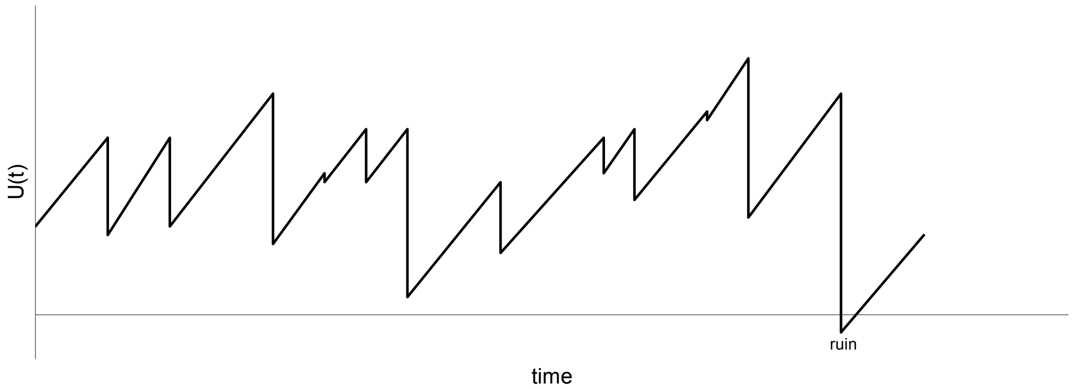

Figure 10. Surplus Process

"Ruin probability," ψ, is the chance that ruin ever occurs. Ruin probability is a measure of the credit risk of the firm or the riskiness of some portfolio of policies. The stability condition for the surplus process is $p > \lambda\beta$, which implies that, in the long run, premium payments exceed claims (Vázquez-Abad and LeQuoc, 2001:71).

The probability of ruin in the canonical risk model is determined by four factors: the arrival rate of claims, the rate at which premiums are paid, the initial wealth of the firm, and the distribution of claim size. Ruin probability increases with a higher claim rate, lower premium rate, lower initial wealth or larger claim sizes. Except in rare cases, ruin probability cannot be solved analytically and must be simulated. When claim size is exponentially distributed, however, ruin probability can be calculated directly according to the following formula:

$$\psi = \frac{\lambda\beta}{c}\exp\left[\left(\frac{\lambda\beta}{c} - 1\right)\frac{u}{\beta}\right] \tag{10}$$

For all distributions of claim size, an upper bound on the probability of ruin can be calculated if the moment generating function of the claim size distribution exists.

Lundberg's upper bound of ruin probability, $\psi < e^{-Ru}$, is a function of the initial wealth and the adjustment coefficient, R, which is the unique solution to the equation

$$\lambda + cR = \lambda M_Y(R) \tag{11}$$

where $M_Y(\cdot)$ is the moment generating function of the claim size distribution (Vázquez-Abad and LeQuoc, 2001:71).

When claim sizes do not follow an exponential distribution, ruin probability cannot be directly calculated. Because ruin is a rare event in many cases, direct simulation can be highly inefficient. The surplus process can be simulated over a long period of time without ruin ever occurring, even though the actual probability of ruin is greater than zero. This means that there is no natural stopping condition for a simulation. There are several techniques available to handle this problem and estimate the probability of ruin. Importance sampling uses a change in probability to measure to make ruin certain. The process can then be simulated until ruin occurs, and ruin probability calculated based on the simulated time (Vázquez-Abad and LeQuoc, 2001:72). The storage process technique (also known as the buffer content) measures the amount of time the process spends above some level of wealth. Finally, the convolution formula technique uses the sequence of losses in the process, where each loss is defined as a new low in the value of the process (Vázquez-Abad and LeQuoc, 2001:73).

Calculating ruin probability for the collective actuarial model might apply in a military context to risk in weapons inventory or troop levels. Inventories slowly building over time and depleting quickly in wartime may possibly be modeled with a surplus process. The probability of ruin is the probability that the inventory stock empties completely or falls below some critical threshold level.

2.4. Summary of Risk Literature

Because risk appears in many academic disciplines, the approaches vary significantly. While much of the literature focuses on the specifics of gathering empirical data or taking actions to mitigate risk in a particular field, some techniques are general to many kinds of risk. This chapter has summarized issues and techniques in the more mathematical approaches to risk.

Although a situation with a deterministic negative outcome can be considered a risk, in most cases risk implies uncertainty about future events. This uncertainty can come from insufficient knowledge, natural variability or vagueness in model specification. When uncertainty comes from an intelligent opponent, a game-theoretic framework may be helpful in understanding the relevant factors. When parameters are uncertain, interval or possibility distributions may be useful for modeling.

Risk in engineering applications, known as reliability analysis, focuses on breaking possible failures into their component parts. After risk has been quantified, tools like the Farmer curve and precautionary principle can point decision makers to the mitigating actions necessary to reduce or eliminate risk.

Generic decision analysis techniques may be relevant to the study of risk as well. Ranking tools like non-parametric decision rules and the ELECTRE and lexicographic methodologies allow the analyst to use limited information about risks to order or prioritize them.

The most mathematically oriented risk approaches come from the financial management and actuarial science fields. Risk-value models use moments of the probability distribution of risks to create single-number risk measures. Actuarial science employs two models of risk, the individual and collective models, to combine multiple risks into a single portfolio.

Table 3 provides a summary of all of the approaches explained in this chapter.

Table 3. Summary of Academic Approaches to Risk

Approach	Features	Application	Advantages/Disadvantages	Reference
Individual actuarial model	total risk as sum of separate independent risks; ignores time	insurance	simple concept; requires good estimate of CDF, especially tail; requires convolution or numerical methods	Kaas et al. 19
Collective actuarial model	total risk as a sum of claims over time; claims often Poisson distributed	insurance	computationally efficient; requires distribution of number of claims and size of each claim	Kaas et al. 46
Non-parametric decision tools	select best possible outcome, least worst outcome or variation of these	many applications	does not require probability distributions; different rules can give different result; requires all alternatives to be identified	Fleisher 292
ELECTRE method	sequential elimination; series of pairwise comparisons	many applications	requires certainty does not require completeness or that each pair of alternatives are comparable	Chankong 207
Farmer curve	plot of frequency versus severity with line indicating acceptability threshold	nuclear radiation release levels	simple in concept but difficult to determine where to draw curve	Henley 13
Reliability	bottom-up approach breaks system into components or identifies all possible failure causes (FMEA, criticality analysis, fault tree analysis)	engineering systems	generally easier to estimate probabilities at component level; analysis quickly becomes very large	Henley 40
partitioned multi-objective risk method	breaks severity axis into pieces and calculates conditional expected values	flooding	complex	Haimes
Hierarchical holographic modeling	layered multiple models examining system from different perspectives	water supply	draws from expertise in management, technology, law, etc.; more useful for identifying than quantifying risks	Haimes working paper 15

Approach	Features	Application	Advantages/Disadvantages	Reference
Impact intensity	measures component level (cost, schedule, technical, etc) potential negative impact and then system impact using additive or multiplicative function	engineering acquisition	flags a risk when only one aspect of many is a problem; relies on subjective probability estimates	Cho et al. 25, Ayyub 61
Risk-value theory	measures of risk based on expected value and variance	finance	measure of risk (number) easier to understand and rank than a distribution; hard to determine the best measure for a particular application	Sarin and Weber 137
MCDA linear programming	takes ranking or other conditions specifying weighting and value scoring and converts to linear program to identify dominated or potentially optimal alternatives	general	allows for uncertainty in both value scoring and weighting; number of required LP problems can grow quickly	Eum et al. 397
Lexicographic method	sequential elimination; ranks alternatives one criterion at a time, starting with most important	general	simple implementation, requires only ordinal scoring	Chankong 200
Precautionary principle	assume worst case scenario	environmental protection	no probability estimates necessary; requires identification of worst possible outcome	Klinke and Renn 1071

III. Methodology

3.1. Methodology Overview

There are two key steps in the prioritization of risks. First, the proposed methodology models each risk as a random variable with an associated probability distribution. For risks associated with shortfalls in Air Force capabilities, each of these risks can be conditioned on existing capability, with adjustments based on the prevention and mitigation effect of any capability change. The second step in risk prioritization is the development of an appropriate risk measure that translates each distribution into a single quantity.

3.2. Modeling Risk

This thesis borrows from the actuarial science definition of risk as a non-negative random variable of severity (Kaas *et al.*, 2001:223). This differs from the CRRA definition by its inclusion of likelihood. Each shortfall in capability has an associated risk, a chance that undesirable consequences will occur. This methodology expands the actuarial definition by making a subtle distinction between severity and risk. Conceptually, severity refers to any possible undesirable outcome. A distribution of severity describes the likelihood of occurrence of any of these outcomes. Risk includes the distribution of severity, but also includes the possibility that no severity will occur. When the occurrence of an adverse event is certain, then, risk and severity are equivalent concepts. When some probability exists, however, that an adverse event will *not* occur, the distribution of risk and distribution of severity are different. The following sections describe a method for mathematically constructing these severity and risk distributions.

3.2.1. Visualizing Risks

Graphically, severity can be shown with a probability density function (for continuous risks) or probability mass function (for discrete risks). A density function shows the likelihood of taking on any severity level. Figure 11 shows three example severity density functions. A probability density function for a risk could have a point mass at zero to account for the probability of no adverse event occurrence.

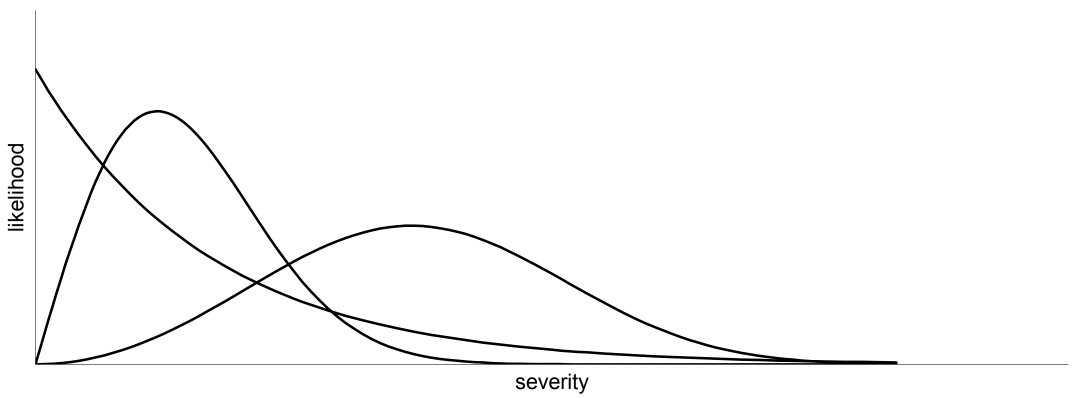

Figure 11. Example Probability Densities of Severity

Alternatively, severity and risk may be visualized by a distribution function. This function, describing the probability that severity will be greater than some value, is known as the complementary cumulative distribution function, the decumulative distribution function or the survival function. This thesis will use the term severity distribution function and the notation $S(x)$, or the term risk distribution function and the notation $R(x)$ for these functions describing the probability that severity exceeds the value *x*. Figure 12 shows example severity distribution functions. Note that for any severity distribution function, the probability that severity is greater than zero is unity. This does not have to be true for a risk distribution function.

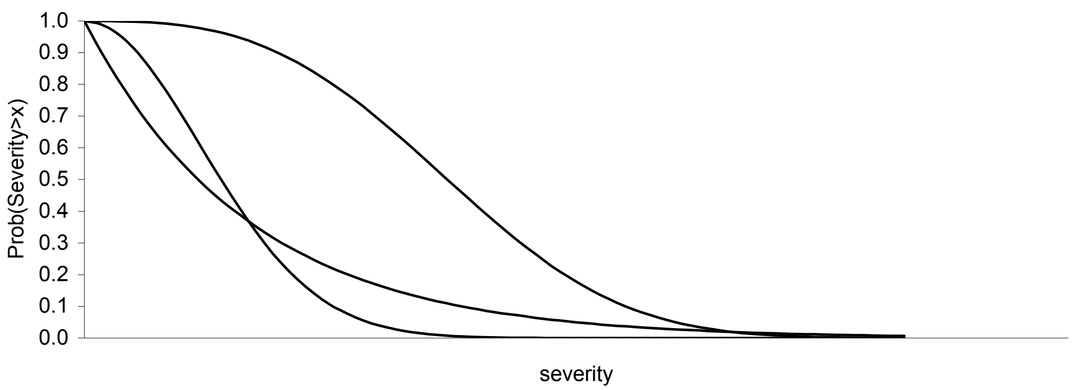

Figure 12. Example Severity Distributions

3.2.2. Modeling Severity

The CRRA team has specified severity using eight risk factors, reflecting the fact that negative consequences in a military context come in several forms. The eight risk factors are *achievement of objectives*, *friendly casualties*, *friendly capability*, *friendly infrastructure*, *collateral damage*, *enemy escalation/weapons of mass destruction*, *U.S. national integrity*, and *U.S. government function*. Each of these risk factors has verbal descriptions identifying the level of risk with one of six severity categories. Severity categories range from *minor* to *catastrophic* (AFSAA, 2003:19-20). Table 4 shows the severity descriptions for the *friendly casualties* factor. The descriptions and severity categories of all eight factors are provided in Appendix A.

Table 4. Severity Categories and Descriptions for Friendly Casualties

Minor	Modest	Substantial	Major	Extensive	Catastrophic
Few citizens/troops killed/ injured. Citizens overseas threatened.	Tens of citizens/troops killed/ injured. Citizens overseas attacked/ injured.	Hundreds of citizens/troops killed/ injured. Citizens overseas attacked/ taken hostage.	Hundreds to thousands of citizens/troops killed/ injured. Citizens overseas killed/ taken hostage.	Thousands to tens of thousands of citizens/troops killed/ injured. Citizens overseas killed/ taken hostage.	Hundreds of thousands of citizens/troops killed/ injured. Many citizens overseas killed/ taken hostage.

In order to mathematically model risk, these qualitative categories of severity must be translated into numerical values. An index linking each category to a number explicitly states not just *that* some outcomes are worse than others, but by *how much* they are worse. This changes the existing ordinal ranking into a ratio scale, with zero equivalent to no severity.

Under the existing quantification system of the categories, all categorical step increases in severity are equal, with *minor* indexed to one, *modest* to two, up to *catastrophic* indexed to six (AFSAA, 2003:37). With this index, a shift from minor severity to modest severity is equivalent in magnitude of change to an increase from major severity to extensive severity. Multiplying each of the values by 100 and dividing by 6 results in the normalized index (rounded to the nearest tenth) in Table 5, with zero equivalent to no adverse event and one hundred equivalent to catastrophe. This index implies that, for example, seven minor events are worse than a single catastrophic event because 7 times 16.5 is greater than 100.

Table 5. Linear and Discrete Severity Index

Category	Severity Index
minor	16.7
modest	33.3
substantial	50.0
major	66.7
extensive	83.3
catastrophic	100.0

This discrete index leaves unclear, for example, whether severity 40 would be classified as *modest* or *substantial*. To avoid these gaps in the index, we can employ intervals as shown in Table 6 (Choobineh and Behrens, 1992:909). Unlike the discrete index, the range zero to one hundred is divided into five equal intervals. This index does not restrict catastrophic severity, which can grow infinitely large as any value greater than one hundred. The interval severity index allows translations from category to number or from number to category. The linear trend still holds, since, for example, a step from the worst minor severity to the worst modest severity is the same as a step from the "best" substantial severity to the "best" major severity.

Table 6. Linear Interval Severity Index

Category	Severity Index
minor	$0 < x \le 20$
modest	$20 < x \le 40$
substantial	$40 < x \le 60$
major	$60 < x \le 80$
extensive	$80 < x \le 100$
catastrophic	$x > 100$

Categorical step increases in severity do not have to be equal. If a step in categorical severity grows multiplicatively, for example, a multiplicative index must be used. Table 7 shows indices where an increase of one category implies a doubling, tripling or ten times the severity. To calculate these indices start with the top value, 100, and divide by the appropriate multiplicative factor. Continue dividing until all five sub-catastrophic categories are specified.

Table 7. Logarithmic Interval Severity Indices

Category	Log_2 Severity Index	Log_3 Severity Index	Log_{10} Severity Index
minor	0 < x ≤ 6.3	0 < x ≤ 1.2	0 < x ≤ .01
modest	6.3 < x ≤ 13	1.2 < x ≤ 3.7	.01 < x ≤ .10
substantial	13 < x ≤ 25	3.7 < x ≤ 11	.10 < x ≤ 1.0
major	25 < x ≤ 50	11 < x ≤ 33	1.0 < x ≤ 10
extensive	50 < x ≤ 100	33 < x ≤ 100	10 < x ≤ 100
catastrophic	x > 100	x > 100	x > 100

Regardless of the precise index used to convert categorical severity ratings to numerical scores, care must be taken to correctly express the true relationships. It is possible to mix indices with some steps increasing by half, others doubling or tripling. However, if a major severity event is considered equivalent to five modest events, for example, the lower limit of the major score must be five times the lower limit of the modest score and the upper limit of the major score must be five times the upper limit of the modest score.

3.2.3. Eliciting Probabilities

After severity has been appropriately quantified, a probability distribution of severity can be estimated by eliciting probabilities from subject matter experts. A number of existing parametric probability distributions may be appropriate. Actuarial science most often uses the exponential, Weibull and Pareto distributions to model the size of insurance claims (Vázquez-Abad and LeQuoc, 2001:78). Each of these three distributions generally follows a form in which less severe outcomes are more likely than more severe ones, which may make these distributions appropriate for modeling severity outcomes in a military context.

The great advantage of the exponential distribution is its simplicity. If severity is exponentially distributed, only one number must be elicited from the decision maker (or other subject matter experts) in order to determine the value of the single parameter, λ. There are two possible approaches. First, the decision maker may provide the average or mean severity, in which case $\lambda = 1 / (\text{mean severity})$ (Wackerly *et al.*, 2002:178). Alternatively, the decision maker may provide the probability that severity *exceeds* some value, x. The λ parameter follows directly as

$$\lambda = -\,ln(\mathrm{P}\{\text{Severity}>x\}) \,/\, x \qquad (12)$$

where

$$S(x) \equiv \mathrm{P}\{\text{Severity}>x\} = e^{-\lambda x}. \qquad (13)$$

If severity cannot be assumed to be exponentially distributed, other parameters must be elicited. One alternative is to specify a discrete distribution for the first five severity levels and use an exponential distribution for the catastrophic severity tail (Bratley *et al.*, 1987:125). An example of such a distribution is shown in Figure 13. This distribution specifies a probability of minor severity, a probability of modest severity, and so forth. The sum of the probabilities of severities less than catastrophic is subtracted from one, and this result is used in Equation (12) to determine the exponential tail of catastrophic outcomes.

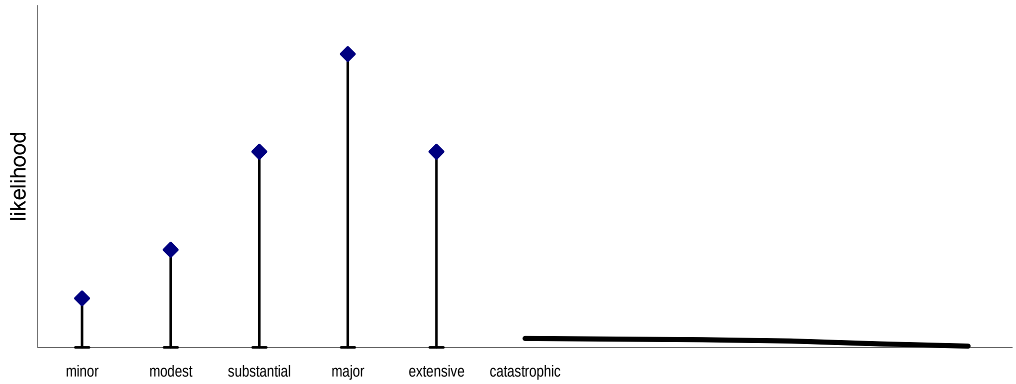

Figure 13. Probability Mass Function with Continuous Tail

Alternatively, the decision maker may provide the probability that severity is greater than some (or all) of the severity categories and a continuous probability distribution may be fitted to the data as shown in Figure 14. (See Section 4.4 for an example of how to fit a Weibull distribution given three exceedance probabilities.)

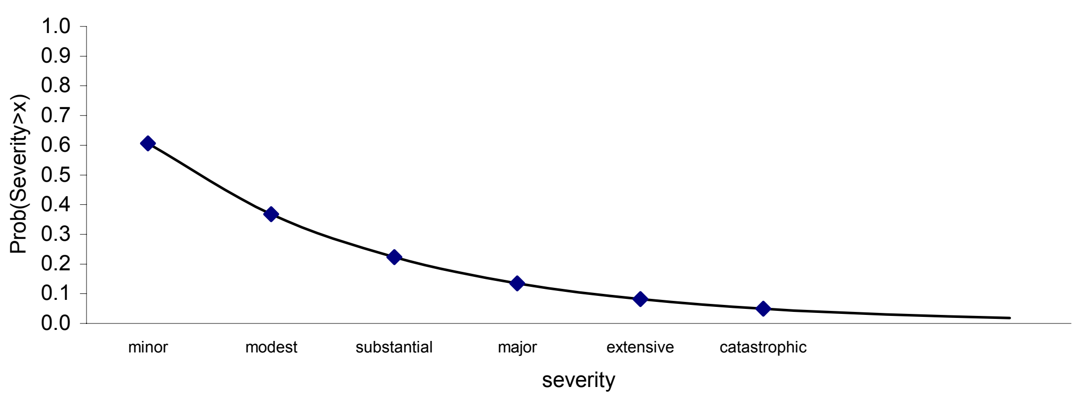

Figure 14. Fitted Continuous Distribution

To move from a distribution of severity to a distribution of risk, subject matter experts must provide one additional input, the probability that an adverse event occurs, that is, the probability that severity is greater than zero. The risk distribution function, $R(x)$, then, is the severity distribution function, $S(x)$, multiplied by this probability of occurrence, $(1\text{-}p)$.

3.2.4. Conditioning on Capability

Severity and risk distributions may not be fixed over time. As Air Force capabilities change, the probability of adverse event occurrence may increase or decrease, or the shape of the severity distribution may change in some way. In general, a change in capability can have two effects. First, it can *prevent*, reducing the probability that *any* adverse event will occur. This is an effect on the risk distribution and not on the severity distribution. Second, it can *mitigate*, reducing the severity of the effects of an adverse event. This is an effect on the severity distribution. The academic literature does not always distinguish between these two effects, preferring the term "mitigation" to refer to both a reduction in probability of occurrence and resulting severity (Ayyub, 2003:107). These two effects make distinctly different changes in a distribution, however, and the next sections describe the ways in which these two effects of capability act on the distribution of risk.

3.2.4.1. Prevention

Prevention is a reduction in the chance that any severity will occur. Under complete prevention, there is no chance of any severity occurring, a riskless situation. A preventive action could make it physically impossible for an adverse event to occur or merely discourage or deter an enemy from creating that adverse event. In either case, the risk is reduced by a decrease in the probability of any negative outcome.

Prevention may be a monotonically increasing function of capability, defined between zero and one, as shown in Figure 15. That is, any increase in capability decreases the likelihood of an adverse event. In some cases, however, increasing capability could encourage preemptive action from an enemy; in such a case an increase in capability would increase the likelihood of an adverse event. This would result in prevention as a *decreasing* function of capability. While this latter function is possible, this research assumes increasing capability increases prevention over the analysis time horizon.

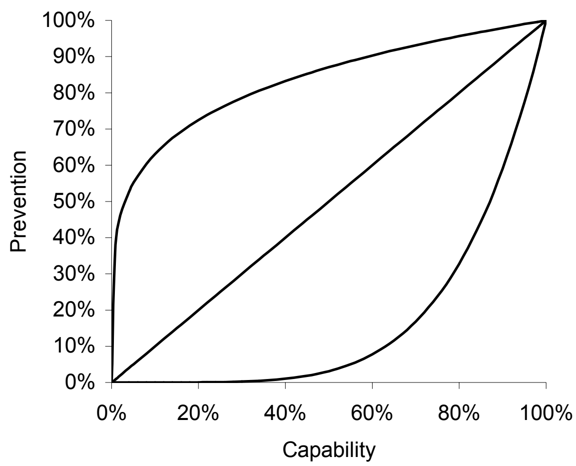

Figure 15. Examples of Prevention Functions

The prevention function focuses particularly on the capability being considered, but potentially also includes any number of other capabilities with interactive effects. A simple model assumes that the other capabilities provide negligible preventive effect.

Minimum prevention occurs when capability is zero and the prevention from other capabilities is negligible. This case does not imply that an adverse event is guaranteed to occur, because the initial risk distribution could have a nonzero probability that the adverse event will not occur even without any capability.

Under prevention, the distribution function of the risk, rather than starting at zero, starts at the prevention level (Kaas *et al.*, 2001:27). Figure 16 shows three risks with different levels of prevention. With low (0.1) prevention, the probability of at least some severe outcome is high. As prevention increases, the probability of (at least some) severity decreases, so risk decreases. Only the "starting value" of the distribution changes as prevention shifts; the shape of the distribution is unaffected.

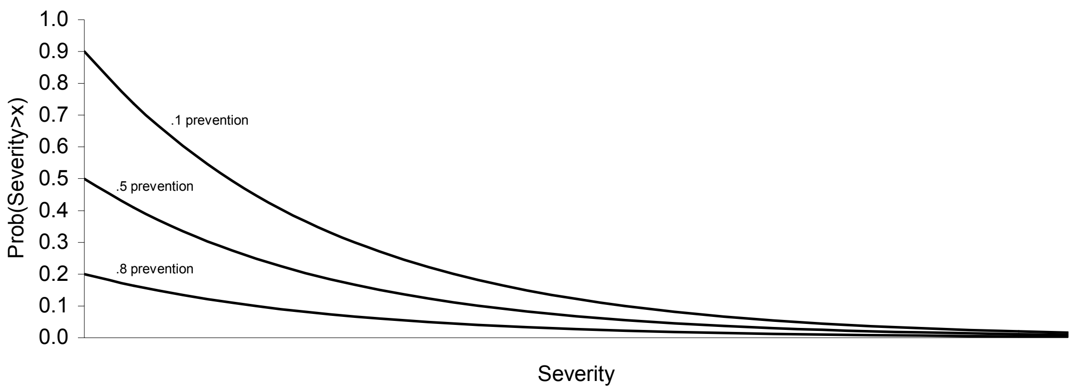

Figure 16. Effect of Prevention on Distribution of Risk

3.2.4.2. Mixed Discrete and Continuous Probability Distributions

This section describes the mathematics that allow prevention, where the risk distribution "starts" at some value less than one. In actuarial science, a claim distribution can be a mixture of a discrete and continuous random variable when there is some nonzero probability that the claim value is zero (Kaas *et al.*, 2001:22). Let X be a discrete random variable on the occurrence of an adverse event such that

$$X = \begin{cases} 0 & \text{with probability } p \\ 1 & \text{with probability } 1 - p \end{cases}$$

Let $Y|X$ be a continuous, exponentially distributed random variable of the severity of the adverse event, conditional on whether that event occurs. That is, Y assumes the value 0 if $X = 0$ and $Y \sim \text{exponential}(\lambda)$ if $X = 1$. Then Y is the unconditional random variable on severity, combining X and $Y|X$.

$$\text{P}\{Y > y\} = \text{P}\{Y > y \mid X = 0\}\text{P}\{X = 0\} + \text{P}\{Y > y \mid X = 1\}\text{P}\{X = 1\}, \tag{14}$$

$$\text{P}\{Y > y\} = 0\, p + \text{e}^{-\lambda y}\,(1 - p), \tag{15}$$

$$\text{P}\{Y > y\} = \text{e}^{-\lambda y}\,(1 - p). \tag{16}$$

In the more general case, when $Y|X$ follows some general distribution, $S(y)$, the risk distribution function of Y is the severity distribution function of $Y|X$ multiplied by one minus the prevention value.

$$\text{P}\{Y > y\} = \text{P}\{Y > y \mid X = 0\}\text{P}\{X = 0\} + \text{P}\{Y > y \mid X = 1\}\text{P}\{X = 1\}, \tag{17}$$

$$\text{P}\{Y > y\} = \text{P}\{Y > y\}(1 - p), \tag{18}$$

$$\text{P}\{Y > y\} = S(y)\,(1 - p). \tag{19}$$

3.2.4.3. Mitigation

Separate from prevention, a second possible effect of a change in capability is *mitigation*, a reduction in severity if, despite one's best efforts, an adverse event does occur. While prevention affects the risk distribution, it has no effect on the severity distribution. Mitigation changes the distribution of severity.

Like prevention, mitigation may be a monotonically increasing function of capability. That is, an increase in capability always reduces the severity of the outcome. A case could exist, however, where increasing capability increases severity.

For example, if improved mobility allows friendly forces to respond more rapidly to a crisis, they may be more vulnerable to attack.

The effect of mitigation on the severity distribution function is not as clear as the effect of prevention on the risk distribution function. An unmitigated severity distribution stochastically dominates the mitigated severity distribution, but it is not obvious how the distribution might change shape. (See Section 3.3.1 for explanation of stochastic dominance.)

If mitigation equally affects all levels of severity, it makes the most sense for mitigation to change a scale parameter of the distribution. The scale parameter defines the measurement of the range of values. Changing the scale parameter spreads or tightens the distribution, while keeping the same essential shape (Law and Kelton, 2000:198).

If mitigation primarily affects just part of the distribution, however, a change in the shape parameter will be necessary. This would occur if an increase in capability only mitigated the most catastrophic severities, for example.

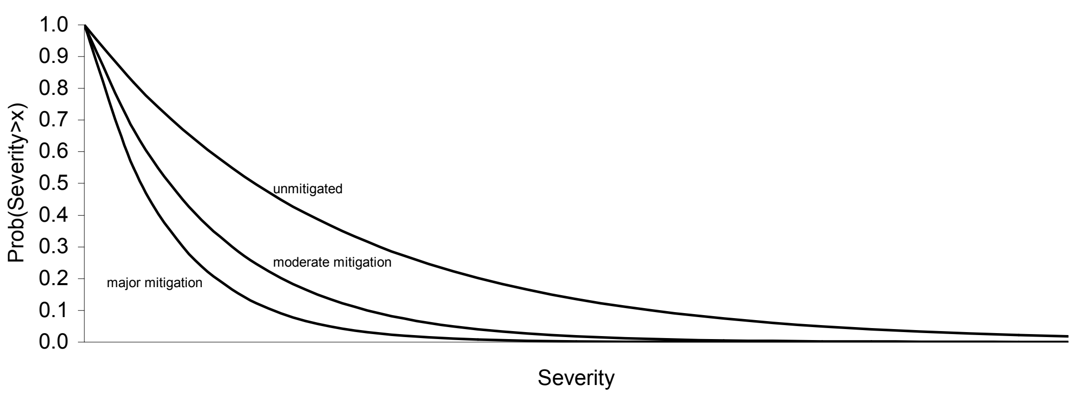

Figure 17. Effect of Mitigation on the Distribution of Severity

3.2.4.4. Overall Model

When both prevention and mitigation are included, the overall distribution of risk is

$$\mathrm{P}\{\mathrm{Severity} > x\} = (1 - p)\,\mathrm{P}\{\mathrm{Severity}_{\mathrm{mitigated}} \le y\} \tag{20}$$

Increasing prevention reduces the probability that any adverse event occurs, while mitigation changes the shape or scale of the distribution of random variable of severity. It will be necessary to model each mitigation effect on a case by case basis.

3.3. Measuring Risk

After a risk has been fully modeled, including a mapping from ordinal severity categories to a continuous quantitative index and conditioning on capability, risks can be measured. A risk measure is a number derived from a risk distribution that summarizes the distribution in a single value. The remainder of this chapter explains four basic tools for the measurement of risk: expectation, conditional expectation, risk-value measurement and distortion functions.

There are two features of risk a good risk measure will capture. First, it should include some aspect of the variation in the outcome. For two risks with the same expected value, the one with the greater range or variability is generally considered more risky. Second, a risk measure should capture something of the undesirable consequences of outcomes. For two risks with the same shape of distribution, the one with the higher expected severity is generally considered more risky (Sarin and Weber, 1993:139).

If the analytic goal is to rank risks, risk measurement is only necessary when it is unclear from the distributions which risk is less desirable than another. When the distributions can be ranked without using measurement, risks are stochastically ordered. To understand stochastic orders, it is necessary to explain the concept of stochastic dominance.

3.3.1. Stochastic Dominance

The cleanest ranking of risks occurs when one risk stochastically dominates another. When one risk is stochastically dominant, it has a greater probability of excessive severity at all points and the distribution functions never cross (Kaas *et al.*, 2001:226). Two such risks are shown in Figure 18, with the solid line risk stochastically dominantly over the dotted line risk. Mathematically, consider two risks with respective risk functions $R_1(x)$ and $R_2(x)$. The first risk is stochastically greater than the second if $R_1(x) \geq R_2(x)$ for all x (Kulkarni, 1995:586).

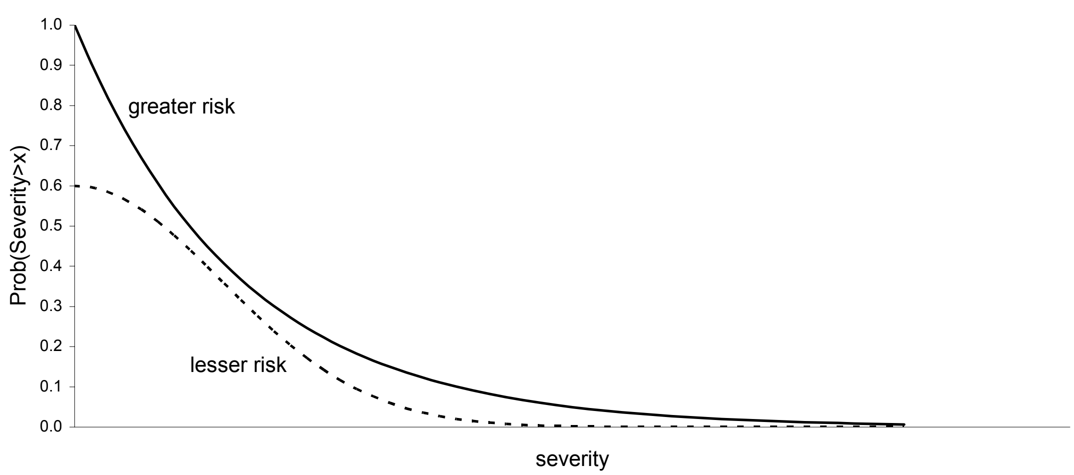

Figure 18. Stochastically Dominant Risk

When stochastic dominance does not exist, or a decision maker requires a quantification of the differences between risks, one of four risk measurement tools is appropriate. Under dominance, the greater risk will always have a higher risk measure. This fulfills the coherent risk measure axiom of monotonicity (Artzner *et al.*, 1999:210). The next four sections explain each of these measures, along with their advantages and disadvantages.

3.3.2. Mathematical Expectation Measure

The simplest measure of risk is the mean or expected severity. The mathematical expectation of a risk can be calculated using the probability density function, $f(x)$, or the risk distribution function, $R(x)$, by integrating over the entire range of severities (Kulkarni, 1995:562) as shown in Equation (21).

$$\text{Risk Measure}_{\text{expectation}} = \text{E}[X] = \int_0^\infty x f(x) dx = \int_0^\infty R(x) dx . \qquad (21)$$

The advantage of expectation as a risk measure is its common use and familiarity to decision makers. The major disadvantage is that the expectation is largely unaffected by changes in the tail of the distribution, leading decision makers to ignore the highly unlikely but catastrophic outcomes. The risk assessment and management process is generally most concerned with those catastrophic outcomes (Haimes, 1998:17). The three remaining risk measures seek to overcome this shortcoming of expectation by giving extra consideration to the extreme outcomes.

3.3.3. Risk-Value Measure

Risk-value begins to deal with the major problem of expectation by including the second moment of the distribution, the variance. In a risk-value measure the decision maker commits to some tradeoff between (expected) value and the uncertainty associated with an outcome (Sarin and Weber, 1993:136).

The simplest risk-value measure is a linear combination of expectation and variance, or expectation and standard deviation (Sarin and Weber, 1993:137).

$$\text{Risk measure}_{\text{risk value}} = a\,\mathrm{E}[X] - (1-a)\,\sigma_X \qquad (22)$$

where σ_X is the standard deviation of X and a is a value between 0 and 1.

The variance of the risk distribution can be calculated similarly to the expectation, with the probability density function or the risk distribution function. Equation (23) shows the common form of variance calculation, and Equations (24) and (25) follow as forms specific to risk distributions with distribution function $R(x)$.

$$\mathrm{Var}[X] = \int_0^\infty x^2 f(x)dx - (E[X])^2 \qquad (23)$$

$$= 0^2 p(X=0) + \int_0^\infty x^2 \frac{d}{dx}(1-R(x))dx - \left(\int_0^\infty R(x)dx\right)^2 \qquad (24)$$

$$= \int_0^\infty x^2 d(1-R(x)) - \left(\int_0^\infty R(x)dx\right)^2. \qquad (25)$$

The standard deviation follows as the square root of the variance.

For an exponentially distributed risk, with probability (1-p) of occurrence of an adverse event and parameter λ, the expectation and variance can be calculated as

$$\mathrm{E}[X] = \int_0^\infty R(x)dx \qquad (26)$$

$$= (1-p)\int_0^\infty e^{-\lambda x}dx \qquad (27)$$

$$= (1-p)/\lambda. \qquad (28)$$

$$\mathrm{Var}[X] = \int_0^\infty x^2 f(x)dx - (E[X])^2 \qquad (29)$$

$$=\int_0^\infty x^2 (1-p)\lambda e^{-\lambda x} dx - \left(\frac{1-p}{\lambda}\right)^2 \quad (30)$$

$$=\frac{1-p^2}{\lambda^2}. \quad (31)$$

One advantage of the risk-value measure is that it does not require that the entire distribution be specified. If the analyst can determine just the first two moments of the distribution, a risk-value measure can be calculated. A second advantage is that the measure can be plotted against the tradeoff parameter, a, so the decision maker can visualize the tradeoff between expectation and variance.

Consider two risks shown in Figure 19. The solid line risk, which has a greater probability of an adverse event occurring, has the larger expected severity. The dotted line risk, which has a lower probability of an adverse event occurring but a heavier distribution tail, has a larger variance.

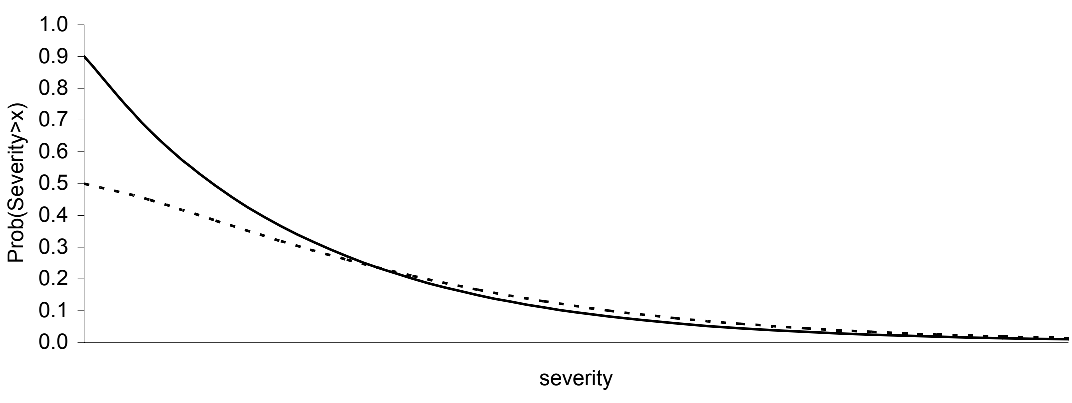

Figure 19. Undominated Risks

Figure 20 shows the tradeoff in the risk-value measure as priority is moved from expectation to standard deviation. When standard deviation is the most highly weighted component, the heavier-tailed dotted line risk has the greater risk measure. When expectation is weighted higher than 0.6, the solid line risk has the greater risk measure. When uncertainty in outcome, measured by the standard deviation, is a significant consideration, the dotted-line risk should be considered more risky. If the uncertainty of outcome is relatively unimportant, and the focus is almost exclusively on expected severity, the solid-line risk is the most significant.

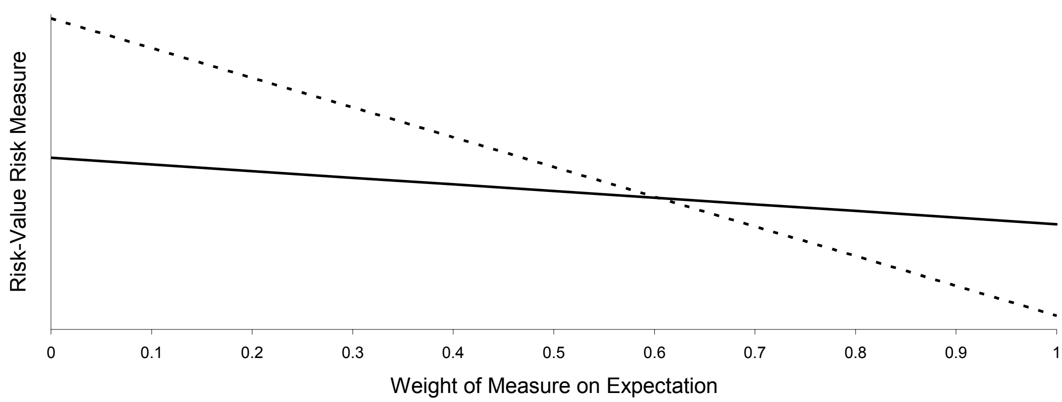

Figure 20. Risk-Value Measure of Risk

A disadvantage of the risk-value measure is the uncertainty of the appropriate tradeoff value. Increasing the weight on standard deviation does not necessarily increase the focus on catastrophic events. In addition, breakpoints where one risk measure crosses another do not have a clear interpretative value.

3.3.4. Conditional Expectation Measure

Conditional expectation is a third possible measure of risk. This measure completely ignores the low severity portions of the risk distribution, focusing on the distribution tail and the worst possible outcomes as shown in Figure 21. The risk measure is the expected severity, given that severity is greater (worse) than some target or accepted value. That value can either be a severity threshold, or a quantile of the distribution (Benati, 2003:574).

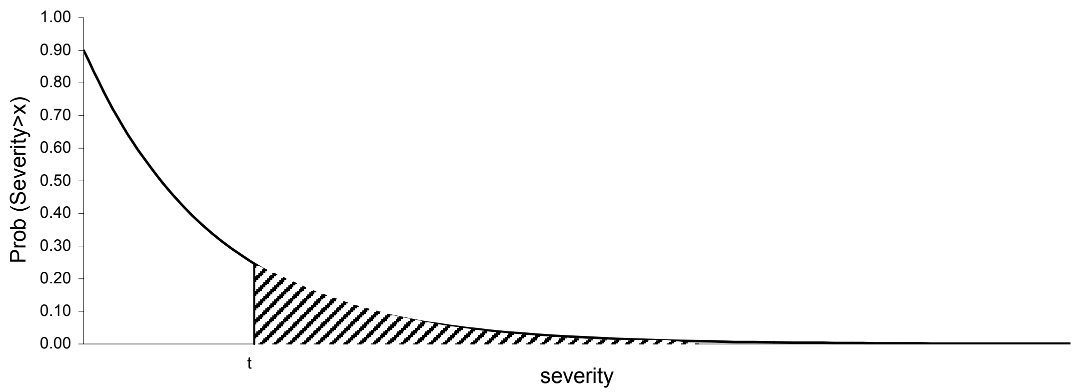

Figure 21. Conditional Severity

The calculation of the conditional expectation risk measure comes from the calculation of residual system life in reliability. Residual life is the expected remaining working time of a system, given that the system has already been in operation for a specified period of time (Ebeling, 1997:34). For this study, the conditional expectation risk measure, instead of conditioning on time, conditions on some severity threshold, *t*.

$$\text{Risk Measure}_{\text{conditional expectation}} = \text{E}[\text{Severity} \mid \text{Severity} > t] \qquad (32)$$

$$= t + \frac{1}{R(t)}\int_t^{\infty} R(x)dx \qquad (33)$$

The conditional threshold can be determined in two different ways. First, the threshold can be specified directly as a severity value. For example, the risk measure could be expected severity, given that severity is greater than *minor*. Alternatively, the threshold can be calculated from a specified quantile of the distribution. For example, the risk measure could be the expected severity, given that severity is in the 80th percentile of the distribution.

To calculate the threshold based on a specified quantile requires the inverse of the risk distribution function. For the exponential distribution

$$R(\text{threshold}) = \alpha = (1\text{-}p)\, e^{-\lambda\ \text{threshold}} \quad (34)$$

$$\text{threshold} = -\ln\,(\alpha/(1\text{-}p))\,/\,\lambda \quad (35)$$

With a specified severity threshold or distribution quantile, conditional expectation provides a single risk measure. The measure can be plotted as a function of the chosen severity threshold to show the analyst or decision maker how the risk ranking might vary. Figure 22 shows the risk measure of the risks in Figure 21 as a function of specified severity threshold. When the threshold is low, and most of the distribution is considered in the calculation, the solid line risk has the higher measure and is considered more risky. As the severity threshold increases, the dotted line risk, with a thicker distribution tail, becomes the risk with the higher measure.

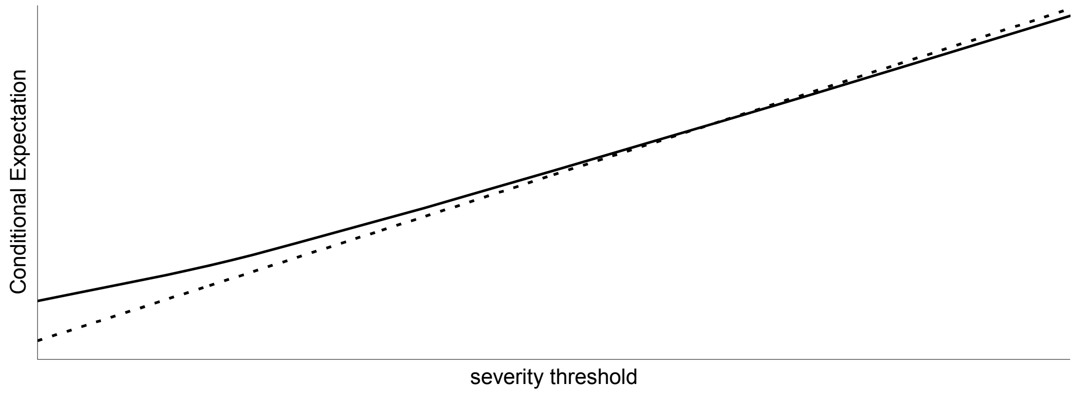

Figure 22. Conditional Expectation as a Function of Severity Threshold

Figure 23 shows the alternative formulation of the conditional expectation risk measure. The measure is plotted as a function of the proportion of the distribution, α, included in the expectation calculation. As α decreases, less of the distribution is included in the calculation, and the risk measure increases for both risks. Similar to the results from the specified threshold approach, when the risk measure is calculated in the distribution tail only, the heavier-tailed solid risk has the higher measure. As a larger fraction of the distributions are included, the dotted line risk has the higher risk measure.

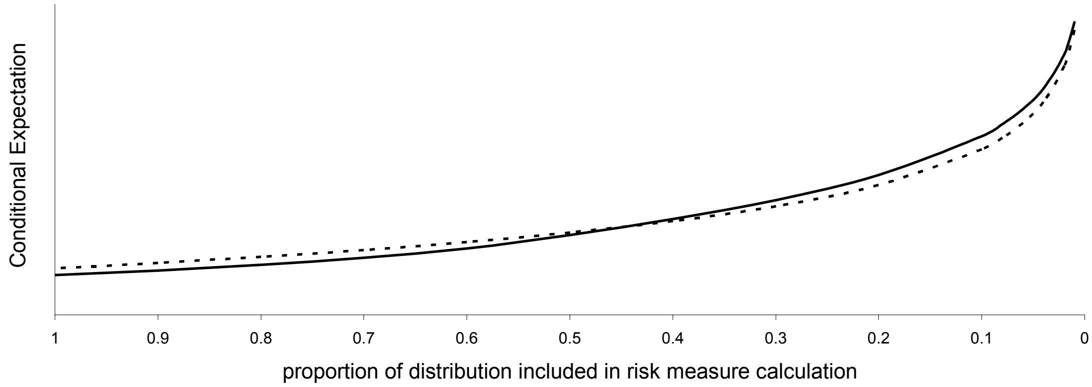

Figure 23. Conditional Expectation as a Function of Quantile

The primary advantage of the conditional expectation risk measure is its focus on the most severe possible outcomes. If the tail of the risk distribution is properly specified, the portion of the distribution in the low severity outcomes does not have to be correct. In addition, the prevention variable is not required at all, since the risk measure is calculated assuming an adverse event does occur. In effect, this risk measure does not distinguish between the risk and severity distributions, and either can be used in calculations. Finally, the measure is relatively easy to explain to a decision maker, and allows the analyst to vary the amount of the distribution used in the calculation to perform a sensitivity analysis.

The biggest disadvantage of the conditional expectation risk measure is that it ignores a large portion of the distribution. Two risks with similar tails but different prevention parameters will have similar risk measures under this approach. Decision makers concerned about the probability of occurrence will not be able to properly indicate their preferences with this risk measure. Thus the conditional expectation risk measure is only appropriate when the entire focus of the decision maker is on the worst possible outcomes.

3.3.5. Distorted Expectation Measure

The final proposed measure of risk is the distorted expectation risk measure. This measure "distorts" the risk distribution function and then calculates the expectation of the distorted function. Several distortion functions are available. All of them, however, re-weight the densities, emphasizing more catastrophic severities and deemphasizing—but still including—less catastrophic ones. This avoids the problem of conditional expectation's ignoring a portion of the distribution (Wirch and Hardy, 1999:337).

Figure 24 shows a density function and its distortion. The distortion "pushes" density into worse severities.

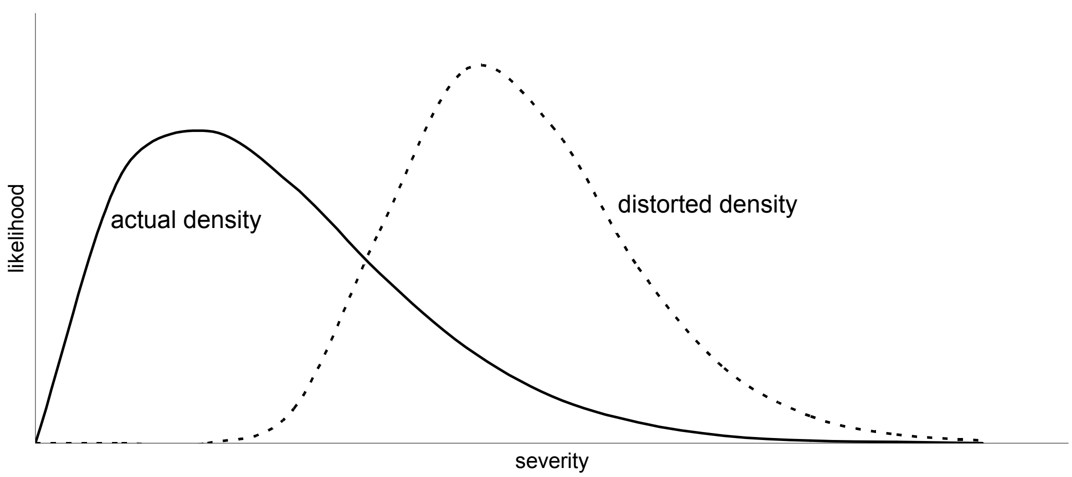

Figure 24. Actual and Distorted Severity Density Functions

Figure 25 shows a risk distribution and its distortion. The distorted distribution dominates the actual distribution at all levels of severity, reflecting the fact that the probabilities have been rescaled toward the worse severities. This places a greater emphasis on higher severities, which may be so unlikely as to have little effect on the undistorted expectation. Under the emphasis caused by distortion, these high severities are effectively given a higher priority. If the decision maker has no desire to emphasize the higher severities, undistorted expectation is an appropriate measure of risk.

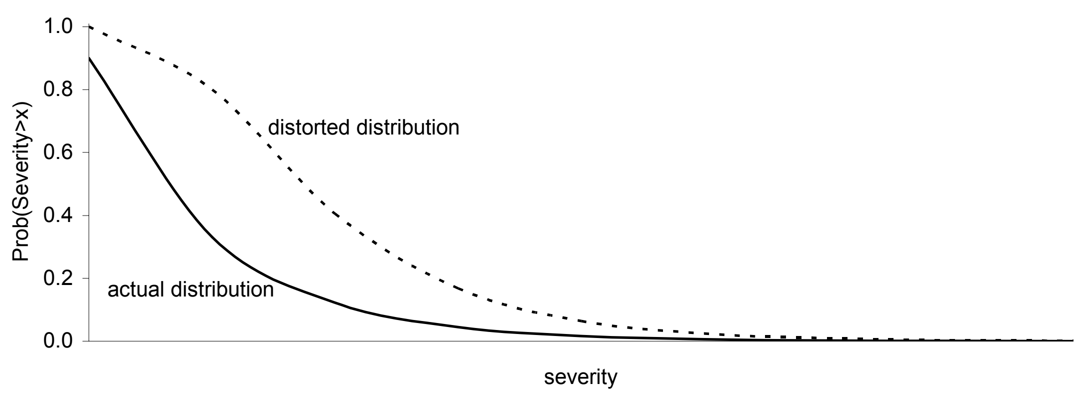

Figure 25. Actual and Distorted Risk Distribution Functions

The actuarial science literature identifies six different distortion functions, though some are special cases of others (McLeish and Reesor, 2003:141). All distortion functions operate on the risk distribution. The distribution can be distorted by applying the chosen distortion function, $g(u)$, to the risk distribution function as follows:

$$R_{\text{distorted}}(\text{severity}) = g(R(\text{severity})) \tag{36}$$

Regardless of the distortion function used, the risk measure is the expected value of the distorted distribution (McLeish and Reesor, 2003:137), calculated as the integral of the distorted distribution over the entire range of severity. Increasing the distortion, which increases the relative weight on the more extreme outcomes, will increase the value of the risk measure.

$$\text{Risk measure}_{\text{distorted expectation}} = \int_0^\infty g(R(x))dx \tag{37}$$

The *gamma-beta distortion* requires three parameters, a, b and c such that $0<a\leq 1$, $b\geq 1$, $c\geq 0$, and is the most general and flexible distortion (McLeish and Reesor, 2003:141). If the parameters a and b have values of one and the c parameter approaches infinity, there is no distortion. Decreasing a or c, or increasing b increases the amount of distortion, shifting the distribution into the tail.

$$g_{\gamma\beta}(u) = \frac{\int_0^u t^{a-1}(1-t)^{b-1} e^{-t/c} dt}{\int_0^1 t^{a-1}(1-t)^{b-1} e^{-t/c} dt} \tag{38}$$

The *beta distortion* allows the c parameter in the gamma-beta distortion to approach infinity, leaving a distortion function with two parameters (McLeish and Reesor, 2003:141). As with the gamma-beta distortion, if the a and b parameters equal 1 there is no distortion. Decreasing the former or increasing the latter increases distortion.

$$g_{\beta}(u) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^u t^{a-1}(1-t)^{b-1} dt \tag{39}$$

The *proportional hazards distortion* is the beta distortion with the *b* parameter held at one, leaving a single parameter with a simple distortion function (McLeish and Reesor, 2003:141). This parameter is a measure of risk acceptance. When *a* is one, there is no distortion, representing a risk neutral position. Decreasing the *a* parameter because of increasing risk aversion increases the distortion.

$$g_{\text{ph}}(u) = u^a \tag{40}$$

The *dual-power transform* is the beta distortion with the *a* parameter held at 1, leaving a single parameter, κ, which is equivalent to the *b* parameter (McLeish and Reesor, 2003:142). When κ=1 there is no distortion; increasing κ, corresponding with increasing risk aversion, increases distortion so that

$$g_{\text{dp}}(u) = 1 - (1 - u)^b. \tag{41}$$

Another special case of the gamma-beta distortion is the gamma distortion, which holds the *b* parameter constant at one (McLeish and Reesor, 2003:142). The *a* and *c* parameters are free to vary, and decreasing either of them increases the amount of distortion.

$$g_{\gamma\beta}(u) = \frac{\int_0^u t^{a-1} e^{-t/c}\, dt}{\int_0^1 t^{a-1} e^{-t/c}\, dt} \tag{42}$$

The final special case of the gamma-beta distortion is the *exponential distortion*, which holds a and b at one, with the c parameter free to vary (McLeish and Reesor, 2003:142). Decreasing the single parameter increases the distortion.

$$g_{\exp}(u) = (1 - e^{-u/c}) \,/\, (1 - e^{-1/c}) \tag{43}$$

In addition to the gamma-beta distortion and its special cases, the *normal distortion* can be used. This distortion, with one parameter, c, uses a standard normal and inverse standard normal distribution (McLeish and Reesor, 2003:142).

$$g(u) = \Phi[\Phi^{-1}(u) - c] \tag{44}$$

The literature does not provide specific guidance on when one distortion is preferred to another (Reesor, 2003). However, the dual-power transform has two particular advantages. First, it requires only a single parameter, while many of the other distortions require two or three. Second, that parameter (κ) has a meaningful interpretation. When κ is an integer, the resulting risk measure can be considered as the expectation of the worst result in κ sample observations of the risk (Wirch and Hardy, 1999:340). Using the single-parameter dual-power distortion, risks can be compared graphically by plotting them versus the value of the distortion parameter, κ.

Distorted expectation requires the entire risk distribution to be known, and calculates a measure using the entire distribution, an advantage over the conditional expectation risk measure. Some distorted measures are complicated, requiring multiple parameters, but the dual-power distortion requires a single parameter that serves as a measure of risk aversion. Using the dual-power distortion the analyst can rank risks at several levels of risk aversion, including the undistorted expectation, a risk-neutral measure.

The primary disadvantage of distorted expectation is its computational complexity. As distortion increases the calculation becomes more and more complex, and depending on the distribution function may not be analytically tractable. In addition, unlike a risk-value measure, distorted expectation requires the entire risk distribution to be known and specified. While the fact that the distorted expectation risk measure *uses* the entire distribution is an advantage, the requirement that the entire distribution be specified places a higher demand on subject matter experts.

3.4. Methodology Conclusions

In summary, risk can be mathematically modeled by treating it as a random variable with an associated probability distribution, including both the probability of occurrence of an adverse event and a distribution of the possible severities if an event occurs. An exponential random variable has the advantage of simplicity, and may be useful for modeling military risk when little is known about the distribution of possible outcomes. Risk distributions can be summarized using a risk measure.

The simplest risk measure, unconditional and undistorted expectation, serves as a baseline for the other three measures. The risk-value measure is equal to the expectation when all the weight is on expectation and becomes less like expectation as more weight is placed on the standard deviation. The conditional expectation measure is equal to the expectation measure when conditioning on the entire distribution, and becomes less like expectation as the severity threshold moves farther into the distribution tail. Finally, the distorted expectation measure is equal to expectation when there is no distortion, and moves away as distortion increases. Each of the alternatives to expectation, then, is a way to increase focus on the more severe outcomes that are highly unlikely, but potentially so catastrophic that they require the decision maker's primary attention.

A risk-value measure requires only the first and second moments of a risk distribution, but does not necessarily offer a way to increase focus on the worst outcomes. A conditional expectation risk measure focuses exclusively on the distribution tail, ignoring low severity regions of the distribution. When a risk distribution can be completely specified, a distorted expectation measure, using the dual-power distortion function on the risk distribution offers a flexible tool for decision makers to summarize risk in a single number.

IV. Numerical Illustration and Results

4.1. Implementation Overview

This chapter provides an illustration of the use of risk modeling and risk measurement techniques presented in Chapter III. The nine top-level capabilities from the CRRA master capabilities library are reviewed. Using four future global scenarios, each capability is given a notional, associated risk distribution. These risk distributions are measured using the dual-power distorted expectation. Finally, each risk measure is used to re-weight objective function coefficients in a linear program to suggest acquisition priorities.

4.2. Optimization of Risk and Capability Alternatives

The goal of the CRRA is to integrate assessments of current capability and risk of capability shortfalls, suggest appropriate courses of action and ultimately provide guidance to the acquisition process (Jumper, 2002). Future systems purchased by the Air Force purchases should reduce the shortfalls identified by the CRRA process. Risk measures provide a way not only to prioritize capability shortfalls but also to adjust the relative value of potential systems under consideration.

If the Air Force has a set of possible future systems, each providing some additional capability, and a budget constraint limiting the number of systems that can actually be acquired, the problem can be formulated as a mathematical program. The objective is to maximize total value, while staying within the budget. Risk measures can be used to adjust the values of each system, based on the relative importance of the capability shortfall that system starts to close.

Under this construction, risk becomes a weighting of the importance of each capability shortfall. Shortfalls identified as "more risky" will have higher measures of risk, effectively increasing the value of closing the capability gap, while shortfalls identified as "less risky" will have smaller measures of risk, and the value of any additional capability will be reduced.

For example, consider risks associated with each of the nine top-level capabilities identified by the Capabilities Review and Risk Assessment process. The full Master Capabilities Library is included in Appendix B. The nine broad capabilities are as follows:

- *Surveillance & reconnaissance* involves conducting missions to satisfy the intelligence requirements of commanders.
- *Intelligence* is developing "knowledge resulting from the collection, processing, integration, analysis, evaluation, and interpretation of available information concerning foreign countries or areas."
- *Command & control* is "the exercise of authority and direction by a properly designated commander over assigned and attached forces."
- *Communications* is the representation, transfer, interpretation and processing of data between people and machines.
- *Force application* means engaging "a variety of targets throughout the battlespace."
- *Force projection* is the means to "extend national power around the globe in a timely manner."
- *Protection* involves "offensive and defensive actions required to respond to a full spectrum of threats and protect forces."
- *Preparation & sustainment* are the "activities required to establish operating locations, generate the mission … and create forces."
- *Force creation* is the organizing, equipping, and training of combat and support personnel.

A shortfall in any of these nine capabilities creates a situation in which negative consequences could occur. The next section uses a notional framework of the future state of the world to estimate the likelihood and severity of the consequences associated with the nine identified capabilities.

4.3. Future Scenarios

Projecting the future security environment is a difficult task. Today, the United States is the preeminent military power in the world, but still faces numerous threats. While the U.S. does not appear to face a threat to its global power in the next few years, within decades the circumstances could differ significantly. Because the acquisition process for implementation of new technologies can be long, considering a different future is an important exercise.

One way to focus thinking for estimating risk in the future is scenario analysis. After falling out of favor, scenario analysis is returning as a popular form of risk assessment in the corporate world as risk analysts broaden their scope from purely financial risks to risks of terrorist attack, loss of company reputation, and supply or operations failures (Survey of Risk, 2004:14).

In April 1996 a team of Air Force officers produced a report, *Alternate Futures for 2025: Security Planning to Avoid Surprise*, suggesting several directions for global security. They developed these scenarios by creating three dimensions of global politics. The first, "American world view," is a measure of the degree to which the United States interacts with the rest of the world and ranges from "domestic" to "global". The second dimension, "ΔTeK", is a measure of the growth and proliferation of technology and ranges from "constrained" to "exponential." The third dimension, "world power grid," is a measure of the dispersion of power and ranges from "concentrated" to "dispersed" (Englebrecht, 1996:x). Combining these three dimensions the report team created four visions for the future, focusing on the extreme positions of the three dimensions.

In order to illustrate the approaches discussed in this thesis, assume subject matter experts provide their best estimates of the likelihood of future severities. The following data on future risk are purely *notional* predictions of the future for each of the four scenarios. For some capabilities, assume the experts provided an overall chance of an adverse event happening, and an average or expected severity if an adverse event occurs. For other capabilities, assume the experts provided an overall chance of occurrence of an adverse event and two exceedance probabilities: the chance that severity will be worse than the *modest* severity category and the chance that severity will be worse than the *major* severity category.

The *Gulliver's Travails* vision assumed a global American world view, concentrated technology and dispersed global power. In this future world, the United States military struggles with worldwide commitments and diverse operations (Englebrecht, 1996:xi). For the notional example, assume subject matter experts assess future risk in the *Gulliver's Travails* scenario according to the parameters specified in Table 8. In this scenario *surveillance & reconnaissance*, *intelligence*, *command & control*, *communications* and *force projection* are the capability shortfalls with relatively high risk.

Table 8. Notional Risk Data for *Gulliver's Travails*

Capability shortfall	Chance of adverse event	Avg severity if event occurs	Prob of severity > *modest*	Prob of severity > *major*
Surveillance & reconnaissance	0.96	---	0.740	0.0700
Intelligence	0.62	3.125	---	---
Command & control	0.68	9.375	---	---
Communications	0.96	---	0.560	0.0500
Force application	0.83	---	0.580	0.0060
Force projection	0.56	---	0.270	0.0080
Protection	0.08	18.75	---	---
Preparation & sustainment	0.12	---	0.090	0.0100
Force creation	0.02	9.375	---	---

In the second vision, *Zaibatsu*, a domestic American world view combines with exponential technology growth and power concentrated in a few multinational corporations to form a superficially peaceful world. The U.S. military struggles to remain relevant in this future, where the largest security threat comes from instability due to income inequity (Englebrecht, 1996:xii). For the notional example, assume subject matter experts assess future risk in the *Zaibatsu* scenario according to the parameters specified in Table 9. In this scenario *surveillance & reconnaissance*, *intelligence*, *command & control*, *communications*, *protection*, *preparation & sustainment* and *force creation* are the capability shortfalls with relatively high risk.

Table 9. Notional Risk Data for *Zaibatsu*

Capability shortfall	Chance of adverse event	Avg severity if event occurs	Prob of severity > *modest*	Prob of severity > *major*
Surveillance & reconnaissance	0.40	---	0.060	0.0080
Intelligence	0.06	3.125	---	---
Command & control	0.09	9.375	---	---
Communications	0.24	---	0.160	0.0060
Force application	0.05	---	0.005	0.0002
Force projection	0.11	---	0.060	0.0100
Protection	0.56	9.375	---	---
Preparation & sustainment	0.50	---	0.420	0.0200
Force creation	0.50	3.125	---	---

In the third vision, *Digital Cacophony*, America maintains its global interests in the face of exponential technological growth and dispersed global power. The main threat faced by the U.S. military in this scenario is advanced weapons of mass destruction and cyber attacks (Englebrecht, 1996:xii). For the notional example, assume subject matter experts assess future risk in the *Digital Cacophony* scenario according to the parameters specified in Table 10. In this scenario *force application, protection*, *preparation & sustainment* and *force creation* are the capability shortfalls with relatively high risk.

Table 10. Notional Risk Data for *Digital Cacophony*

Capability shortfall	Chance of adverse event	Avg severity if event occurs	Prob of severity > *modest*	Prob of severity > *major*
Surveillance & reconnaissance	0.59	---	0.540	0.0500
Intelligence	0.77	9.375	---	---
Command & control	0.60	3.125	---	---
Communications	0.87	---	0.410	0.0500
Force application	0.76	---	0.580	0.1100
Force projection	0.90	---	0.250	0.0070
Protection	0.54	3.125	---	---
Preparation & sustainment	0.66	---	0.570	0.0070
Force creation	0.61	9.375	---	---

In the final future vision of the world, *King Khan*, the United States role in the world shrinks and a peer competitor in Asia takes over as the primary global power. The U.S. military faces drastically reduced budgets and must prioritize which capabilities it will keep (Englebrecht, 1996:xii). For the notional example, assume subject matter experts assess future risk in the *King Khan* scenario according to the parameters specified in Table 11. In this scenario *force application* and *force projection* are the capability shortfalls with relatively high risk.

Table 11. Notional Risk Data for *King Khan*

Capability shortfall	Chance of adverse event	Avg severity if event occurs	Prob of severity > *modest*	Prob of severity > *major*
Surveillance & reconnaissance	0.20	---	0.190	0.0400
Intelligence	0.47	3.125	---	---
Command & control	0.35	18.75	---	---
Communications	0.04	---	0.010	0.0020
Force application	0.45	---	0.400	0.0100
Force projection	0.16	---	0.004	0.0003
Protection	0.21	3.125	---	---
Preparation & sustainment	0.26	---	0.130	0.0200
Force creation	0.23	18.75	---	---

4.4. Overall Probability Estimation

Overall risk distributions can be built based on the likelihood that any of these scenarios is likely to occur. Using conditional expectation, the likelihood of an adverse event can be calculated as the sum of the severity given a scenario, times the likelihood of the scenario.

$$\begin{aligned} P\{\text{Severity}>0\} = & P\{\text{Severity}>0 \mid \text{Gulliver's Travails}\}\, P\{\text{Gulliver's Travails}\} \\ & + P\{\text{Severity}>0 \mid \text{Zaibatsu}\}\, P\{\text{Zaibatsu}\} \\ & + P\{\text{Severity} \mid \text{Digital Cacophony}\}\, P\{\text{Digital Cacophony}\} \\ & + P\{\text{Severity}>0 \mid \text{King Khan}\}\, P\{\text{King Khan}\}. \end{aligned} \tag{45}$$

Table 12 shows the combined probabilities and severities assuming each scenario is equally likely. The scenarios do not have to be equally weighted however. For illustrative purposes, assume severity doubles for each categorical step increase, and assume catastrophic severity is defined as any severity greater than 100. For the categorical specifications of expected severity the midpoint of the category interval has been used to calculate an average.

Table 12. Combined Quantitative Risk Descriptions

Capability shortfall	Event likelihood	Avg if event occurs	Prob of severity > *modest*	Prob of severity > *major*
Surveillance & reconnaissance	0.54	---	0.38	0.04
Intelligence	0.48	4.7	---	---
Command & control	0.43	10.2	---	---
Communications	0.53	---	0.29	0.03
Force application	0.52	---	0.39	0.03
Force projection	0.43	---	0.15	0.01
Protection	0.35	8.6	---	---
Preparation & sustainment	0.39	---	0.30	0.03
Force creation	0.34	10.2	---	---

For risks with likelihood of occurrence and average severity, an exponential distribution can be fit. The risk distribution can be defined as

$$\mathrm{P}\{\text{Severity} > x\} \equiv R(x) = (1\text{-}p)\, e^{-\lambda x} \qquad (46)$$

The p parameter is the complement of the probability of occurrence. The λ parameter can be calculated using the specified average severity. Because this was specified as the average *if* an event occurs, the calculation is made without regard for p.

$$\text{Average severity} = \int_0^\infty R(x)dx = \int_0^\infty e^{-\lambda x} dx \qquad (47)$$

$$\lambda = 1 \,/\, (\text{Average severity}) \qquad (48)$$

For risks with estimated likelihood of occurrence and two additional probability estimates, a Weibull distribution can be fit. The risk distribution can be defined as

$$\mathrm{P}\{\text{Severity} > x\} \equiv R(x) = (1\text{-}p)\, \exp(-\lambda^{\beta} x^{\beta}) \qquad (49)$$

The p parameter is the complement to the probability of occurrence. The λ and β parameters can be calculated by simultaneously solving the other two specified probabilities.

$$R(\text{modest}) = R(12.5) = (1\text{-}p)\exp(-\lambda^{\beta}12.5^{\beta}) \quad (50)$$

$$R(\text{major}) = R(50) = (1\text{-}p)\exp(-\lambda^{\beta}50^{\beta}) \quad (51)$$

$$\beta = \frac{\ln\left[\ln\left(S(\text{major})\big/(1-p)\right) - \ln\left(S(\text{modest})\big/(1-p)\right)\right]}{\ln\left(\text{major}\big/\text{modest}\right)} \quad (52)$$

$$\lambda = \exp\left\{\frac{1}{\beta}\ln\left[\frac{-\ln\left(S(\text{modest})\big/(1-p)\right)}{\text{modest}^{\beta}}\right]\right\} = \exp\left\{\frac{1}{\beta}\ln\left[\frac{-\ln\left(S(\text{major})\big/(1-p)\right)}{\text{major}^{\beta}}\right]\right\} \quad (53)$$

Table 13 shows the parameters of the fitted distributions and Figure 26 shows these risks graphically.

Table 13. Risk Distribution Parameters

	p	λ	β
Surveillance & reconnaissance	0.4625	0.0345	1.4528
Intelligence	0.5200	0.1024	---
Command & control	0.5700	0.0423	---
Communications	0.4725	0.0425	1.1357
Force application	0.4775	0.0351	1.6393
Force projection	0.5675	0.0643	0.9800
Protection	0.6525	0.0404	---
Preparation & sustainment	0.6150	0.0327	1.7018
Force creation	0.6600	0.0335	---

Note that an exponentially distributed risk is equivalent to a Weibull distributed risk with a β parameter equal to one.

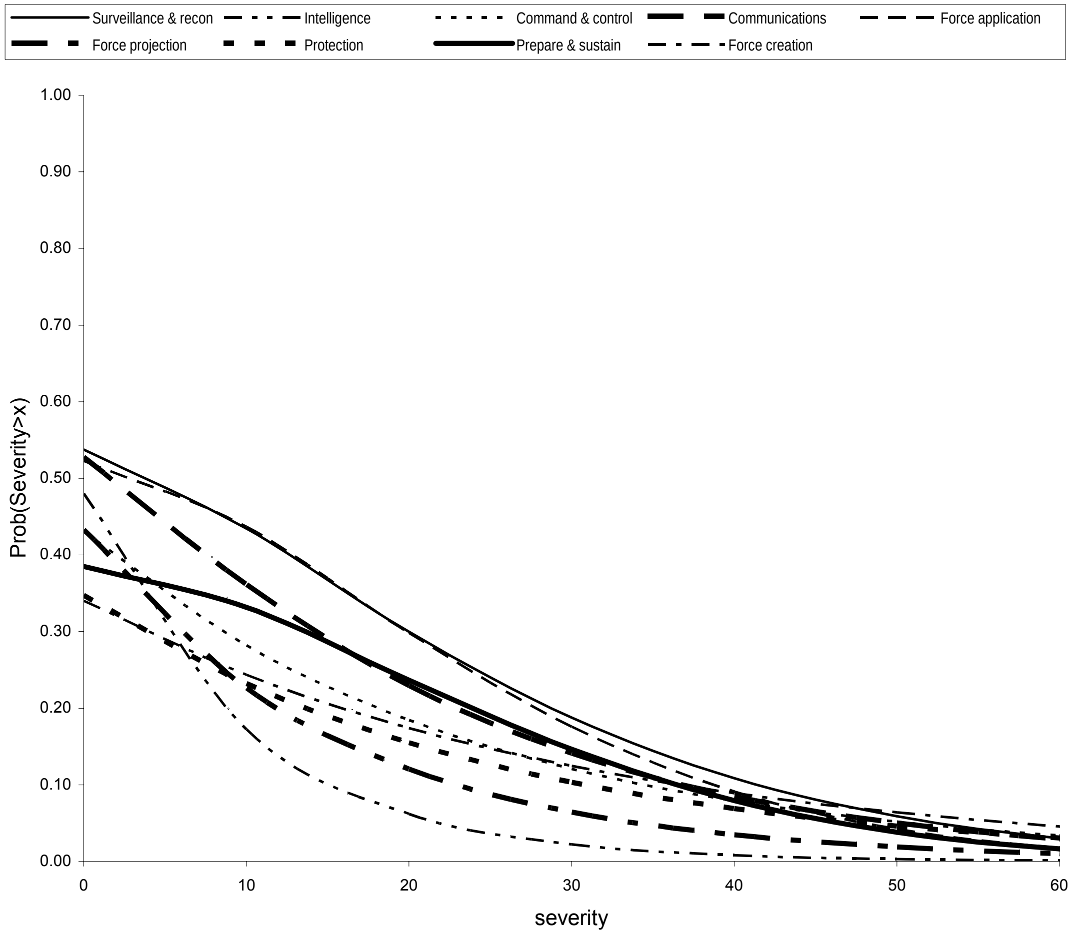

Figure 26. Graphical Depiction of Risks Associated with Nine Capabilities

4.5. Risk Prioritization and Measurement

With risk distributions fully specified, risks can be analyzed for prioritization and measurement. The first step is to identify any sets of risks with stochastic dominance. This provides ordinal ranking, but has no associated value. Risk measurement then follows for all nine risks using the distorted expectation risk measure.

4.5.1. Stochastic Dominance

These risks can be analyzed for stochastic dominance by setting the distribution functions equal to each other. If a solution exists to this equation, and the solution is not a point of tangency, the distribution functions cross and neither risk dominates the other.

$$R_1(x) = R_2(x) \tag{54}$$

$$(1-p_1)e^{-\lambda_1^{\beta_1}x^{\beta_1}} = (1-p_2)e^{-\lambda_2^{\beta_2}x^{\beta_2}} \tag{55}$$

$$\lambda_1^{\beta_1}x^{\beta_1} - \lambda_2^{\beta_2}x^{\beta_2} - \ln\left(\frac{1-p_1}{1-p_2}\right) = 0 \tag{56}$$

In the example, every pair of risks has a solution to this equation with the exception of the *surveillance & reconnaissance* and *preparation & sustainment* risks, shown in Figure 27. Over the entire severity range, *surveillance & reconnaissance* has a greater probability, so it stochastically dominates *preparation & sustainment*, and will have a larger risk measure regardless of what risk measure is used.

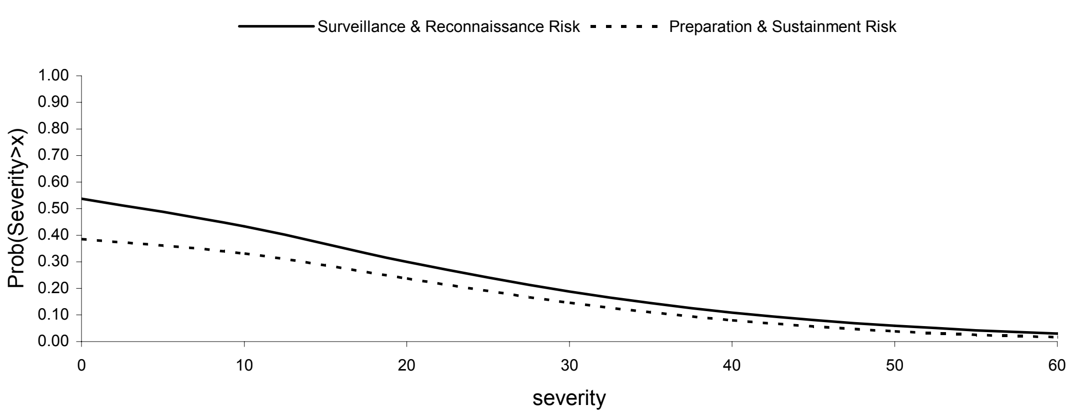

Figure 27. Stochastically Ordered Risks

4.5.2. Risk Measurement

The expected value of the exponential distributed risks is $\frac{1-p}{\lambda}$ and the expected value of the Weibull distributed risks is $(1-p)\left(\frac{1}{\lambda\beta}\right)\Gamma\left(\frac{1}{\beta}\right)$. Table 14 shows the expected severities of each of the nine risks.

Table 14. Expected Severity Risk Measure

Risk	Expected Severity	Priority Order
Surveillance & reconnaissance	14.1	(1)
Intelligence	4.7	(9)
Command & control	10.2	(5)
Communications	11.8	(3)
Force application	13.3	(2)
Force projection	6.8	(8)
Protection	8.6	(7)
Preparation & sustainment	10.5	(4)
Force creation	10.2	(6)

Under the expected value measure of risk, the *surveillance & reconnaissance* capability has the greatest risk, followed by the *force application* capability. The *intelligence* capability has the smallest associated risk. Again, it is important to note that these are all purely notional values.

The distorted expectation risk measure can be calculated by

$$\text{Risk Measure}_{\text{distorted}} = \int_0^\infty \left(1 - [1 - R(x)]^\kappa\right) dx \tag{57}$$

where larger κ creates larger distortion and corresponds to greater risk aversion, more emphasis on the distribution tail. When κ is one there is no distortion and the resulting measure is the undistorted expectation. Under increasing values of κ, the resulting expectation is the κth order statistic, that is, the expected worst outcome if κ samples are taken of the random variable.

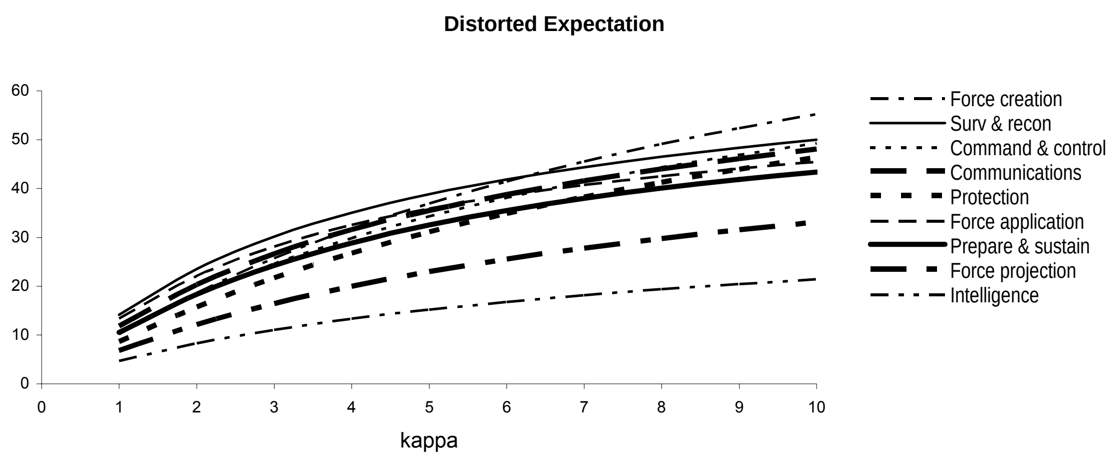

Figure 28. Distorted Expectation Risk Measure as a Function of Risk Aversion

Table 15 shows the values of the distorted expectation risk measure for the nine risks at three levels of distortion. With no distortion *surveillance & reconnaissance* is the greatest risk, followed by *force application*. *Intelligence* is the capability with the smallest amount of risk. With some distortion, *surveillance & reconnaissance* remains the greatest risk, but *force creation* is the second highest. As distortion increases even more, *force creation* becomes the greatest risk, followed by *surveillance & reconnaissance*.

Table 15. Distorted Expectation Risk Measure

Risk	No distortion (κ=1)	Low distortion (κ=5)	High distortion (κ=10)
Surveillance & Reconnaissance	14.1 (1)	38.8 (1)	50.0 (2)
Intelligence	4.7 (9)	15.2 (9)	21.4 (9)
Command & Control	10.2 (5)	34.3 (5)	49.3 (3)
Communications	11.8 (3)	35.5 (4)	48.1 (4)
Force Application	13.3 (2)	35.9 (3)	45.5 (6)
Force Projection	6.8 (8)	23.0 (8)	33.1 (8)
Protection	8.6 (7)	31.1 (7)	46.4 (5)
Preparation & Sustainment	10.5 (4)	32.5 (6)	43.4 (7)
Force Creation	10.2 (6)	36.9 (2)	55.3 (1)

Force creation has the largest ascent in the rankings from zero to high distortion, moving from the sixth highest risk to the greatest. *Command & control*, moving from fifth to third, and *protection*, moving from seventh to fifth, also display changes in the rankings. Each of these risks has a relatively thick tail in its distribution, so increasing the amount of distortion increases these risk measures most significantly.

Risks that fall in the rankings as distortion increases include *force application*, from second to sixth, and *preparation & sustainment*, from fourth to seventh. These risks have relatively thin tails in their distributions, so increasing the amount of distortion has a small effect on the risk measure.

4.6. Potential Systems

Ultimately the goal of the CRRA process is to guide future system acquisition. Risk measurement can play a role by quantifying the danger of a capability shortfall and weighting the importance of closing that gap.

This section suggests a notional combination of future systems the Air Force might consider. A limited budget means that every system on the list cannot be acquired, so the goal is to determine the optimal mix of systems.

In a December 2003 press release, the Air Force announced six capability shortfalls identified by the first iteration of the CRRA process (U.S. Air Force, 2003).

- Global information grid: The Air Force must create a massive system to collect, process, and disseminate information for policy-makers and service personnel.
- Battle space management: The service must create a useful operational picture and implement war planning based on combat effects.
- Fleeting and mobile targets: The service must reduce the time to find, track and destroy enemy forces.
- Battle damage assessment: The Air Force should build a toolkit and definitions for commanders to analyze combat effects.
- Base defense: Roles and responsibilities between the Air Force and the other services must be clarified.
- Cargo airlift: The Air Force should begin a formal review of requirements and prepare for possible force structure changes.

Suppose that, based on these shortfalls, the Air Force identifies six systems for possible future acquisition. Note that these are purely notional potential acquisition projects.

- Enhanced globally-accessible intelligence database
- New heads-up display for fighter aircraft to increase pilot battlespace awareness
- Standoff missile designed for use against mobile targets
- Unmanned aerial vehicle with sensors specific for battle damage assessment
- Detection equipment for chemical, biological and explosive devices at base entry points
- Additional strategic aerial refueling aircraft

It is assumed that the intelligence database provides value in closing the shortfall associated with the intelligence, command & control, communications, force application, protection, preparation & sustainment and force creation capabilities. The fighter aircraft HUD provides value in closing the shortfall associated with the surveillance & reconnaissance, intelligence, command & control, force application, force projection and force creation capabilities. The standoff missile provides value in closing the shortfall associated with the command & control, force application and force projection capabilities. The UAV provides value in closing the shortfall associated with the surveillance & reconnaissance, intelligence, communications and force application capabilities. The detection equipment provides value in closing the shortfall associated with the surveillance & reconnaissance, intelligence, protection, preparation & sustainment, and force creation capabilities. The refueling aircraft provides value in closing the shortfall associated with the force application, force projection, preparation & sustainment, and force creation capabilities.

Table 16 shows the *notional* values of each of these potential systems for closing the nine capability shortfalls, as well as the notional cost of each system. Where no number is specified, the system is assumed to have negligible impact in closing the capability shortfall.

Table 16. Notional Reduction in Capability Shortfall by Potential Systems

Capability Shortfall	Intel database	Fighter HUD	Standoff missile	BDA UAV	Base detection	Tankers
Surveillance & reconnaissance	---	0.19	---	0.26	0.26	---
Intelligence	0.46	0.21	---	0.12	0.68	---
Command & control	0.34	0.19	---	---	---	0.23
Communications	0.14	0.42	---	0.36	---	0.05
Force application	0.10	0.21	0.92	0.30	---	0.10
Force projection	---	0.16	0.54	---	---	0.11
Protection	0.16	---	---	---	0.25	---
Preparation & sustainment	0.19	---	---	---	0.31	0.48
Force creation	0.10	0.05	---	---	0.33	0.36
Cost	7	7	10	9	6	9

4.7. Optimization Solution

The optimal mix of systems to acquire can be solved using a "knapsack problem" in which each potential system is assigned a binary decision variable that takes on a value of one if the system is included in the set of systems to be acquired and zero if it is excluded. An objective function to maximize is the sum of the binary decision variables, with each variable multiplied by the relative value its system provides.

Consider the set of nine desired capabilities and the set of six possible systems for acquisition. An objective function of the acquisition problem to be optimized is

$$\text{Maximize} \sum_{i=1}^{9} \left(RM_i \sum_{j=1}^{6} v_{ij} x_j \right) \tag{58}$$

$$\text{subject to} \sum_{j=1}^{6} c_j x_j \le b \tag{59}$$

where

RM_i is the measure of risk associated with capability i, taking into account both the probability of occurrence of an adverse event and the distribution of severity if an adverse event does occur, $i = 1, 2, \ldots 9$

v_{ij} is the value that system j brings to capability i, $i = 1, 2, \ldots 9$, $j = 1, 2, \ldots 6$

x_j is a binary decision variable equal to one if system j is to be acquired and zero if system j is to be rejected, $j = 1, 2, \ldots 6$

c_j is the cost of system j

b is the maximum available budget

Because the particular values of the risk measures do not have any meaning beyond their relative relationship to each other, the *value* of the objective function at optimality does not have a meaningful interpretation. Therefore, the usefulness of the mathematical program is to identify the optimal mix of systems.

Objective function coefficients are the amount of the capability shortfall closed by the system in question, adjusted for risk. The values in Table 16 are multiplied by the associated risk measures. This results in inflated values for capability shortfalls with higher risks and deflated values for capability shortfalls with lower risks.

Ignoring risk entirely and using a 23-unit budget, the optimal acquisition plan is the intelligence database, standoff missile and base detection system. This leaves no slack in the budget. If the budget is reduced by one to 22, those three systems have become too expensive and the missile is exchanged for the HUD in the optimal mix of systems. Under a budget increase there is no change in the optimal system mix until the budget reaches 29. At that level it is possible to acquire four of the six systems and the optimal mix includes the intelligence database, HUD, base detection equipment and the new tankers.

Including risk in the objective function—by using the undistorted expectation risk measure—results in an optimal system mix of HUD, standoff missile and base detection equipment. The HUD is selected instead of the database because it provides value to *surveillance & reconnaissance*, the highest ranked risk in the notional example. The database does not provide any *surveillance & reconnaissance* value and the largest portion of its value is in *intelligence*, which is the lowest ranked risk.

As with the risk-ignored solution, the optimal system mix is more sensitive to budget reduction than budget increase. Reducing the budget by one to 22 changes the optimal mix to the HUD, base detection equipment and new tankers. The budget must increase to 29 before the optimal mix changes. Under that budget the optimal mix includes the database, HUD, base detection equipment and new tankers.

Increasing risk aversion continues to change the optimal mix of systems to acquire. Under low distortion (κ=5) the result is the same as for the undistorted expectation risk measure. Raising the distortion to κ=10 for a distorted expectation risk measure results in an optimal mix of HUD, base detection equipment and new tankers. The tankers provide the most value of any system to *force creation*, which is the highest ranked risk under high distortion.

Tightening the budget by two to 21 changes the optimal mix from the tankers to the less expensive intelligence database. Increasing the budget by five to 29 allows for an additional system to be acquired, and the optimal mix includes the database, HUD, detection equipment and tankers.

In this notional analysis, the UAV is not included in the optimal acquisition mix in any of these scenarios because of its relatively high cost and low value to the nine capabilities. Under any of the risk measures, including the risk-excluded alternative, the UAV cost must fall from nine to three before it will become part of the optimal mix. Alternatively, the UAV will become part of the optimal mix (using the κ=5 distorted expectation risk measure) if its value increases from 0.26 to 0.52 for *surveillance & reconnaissance*, 0.12 to 0.78 for *intelligence*, 0 to 0.30 for *command & control*, 0.36 to 0.65 for *communications*, 0.30 to 0.58 for *force application*, 0 to 0.44 for *force projection*, 0 to 0.33 for *protection*, 0 to 0.31 for *preparation & sustainment*, or 0 to 0.28 for *force creation*. If none of these changes to the UAV program parameters are realistic, the UAV may be eliminated from discussion to simplify the problem.

Table 17 summarizes the results for a 23-unit budget. The problem formulations are included in Appendix C. The base detection equipment is included in the optimal system mix regardless of risk measure and the unmanned aerial vehicle is never included in the optimal mix.

Table 17. Optimization Summary Results

	Risk Measure			
	None (risk measure=1)	Undistorted Expectation (κ=1)	Low Distorted Expectation (κ=5)	High Distorted Expectation (κ=10)
Total cost	23	23	23	22
Intel database	buy	---	---	---
Fighter HUD	---	buy	buy	buy
Standoff missile	buy	buy	buy	---
BDA UAV	---	---	---	---
Base detection	buy	buy	buy	buy
Tankers	---	---	---	buy

The mathematical optimization program can be made more robust and flexible than the example given. The decision variables could be relaxed, for example, from binary variables to any real value between zero and one. This would allow systems to be acquired at less than full capability (Bretschneider, 1993:130). Alternatively, the decision variables could be allowed to take on integer values greater than one to model acquiring multiple copies of a single system.

The constraint set can also be developed further, including more than just a single budget constraint. In addition to budget or resource availability constraints, integer and linear programming problems often include program balance constraints, where the acquisition of one system either requires or does not allow the acquisition of another (Bretschneider, 1993:130). For example, if the HUD and UAV should not both be acquired, a constraint

$$HUD + UAV \leq 1 \quad (60)$$

prevents both systems from being included in the optimal mix.

4.8. Illustration Summary

This chapter illustrates one possible application of the methodology explained in Chapter III. The illustration considers nine high-level capabilities from the CRRA Master Capabilities Library, and four possible scenarios for future security environments the U.S. Air Force may find itself facing. This framework allows subject matter experts to make considered judgments about the probability and severity of future events. These judgments can be combined into risk distributions.

The most difficult part of risk prioritization and measurement is correctly determining the parameters of the risk distribution. The exponential and Weibull distributions offer the advantage of requiring relatively few inputs from subject matter experts. When distributions have been specified, a distorted expectation risk measure can be used to summarize each distribution in a single number for ranking purposes. The amount of distortion must be determined by the analyst and decision maker according to the appropriate amount of risk aversion. The distortion can be varied for sensitivity analysis, but should not be adjusted in order to give the "right" answer.

The most natural use of a risk measure is to rank or prioritize risks. This chapter suggests a further use, as an adjustment to the objective function coefficients of a mathematical program to determine the optimal mix of new systems under a limited budget. The greatest risks, with high risk measures, inflate the relative value of systems that close capability shortfalls, while the smallest risks, with low risk measures shrink the relative value of systems that close those shortfalls.

In addition to traditional sensitivity analysis performed on the budget or the objective function coefficients, in this methodology the amount of distortion in the risk measure can also be adjusted to test for the sensitivity of the optimal acquisition mix. In the notional example, excluding risk from the acquisition decision resulted in a different optimal solution from a solution using an undistorted expectation risk measure. High distortion produced a third different optimal solution. These changes in the solution represent optimal decision making at different levels of risk aversion. Increasing the level of distortion corresponds to increased risk-aversion, and a greater focus on the worst possible outcomes.

V. Conclusions and Recommendations

5.1. Background and Literature Conclusions

The study of risk works its way into many disciplines, and the military may be able to enhance its own understanding of risk by borrowing from these approaches. In the private sector, significant progress has been made toward the analytic goal of understanding and quantifying risk. The growth in information technology, and the amount of data collected on, for example, life spans, earthquakes and stock market volatility, have allowed for increasingly complex mathematical models for understanding and describing risk (Survey of Risk, 2004:4). Military risk lacks the voluminous quantitative data of financial markets and mechanical components. Building risk distributions, then, relies heavily on subjective expert forecasts.

The general field of decision analysis offers a number of tools relevant to the risk ranking problem. Particularly when very little is known about the underlying distributions of risks, and subjective expertise plays a large role, techniques like non-parametric decision rules and ranking algorithms like the lexicographic or ELECTRE methodologies may be valuable. When possible, however, estimating more complete distributions of risk will provide more insight.

Engineering approaches to risk focus on breaking complex systems into component parts for simpler analysis. The CRRA master capabilities library will, ideally, be an exhaustive list of mutually exclusive Air Force capabilities, broken down into simple, measurable sub-capabilities. This is effectively the componentization of the complex system of the United States Air Force. As the library grows more detailed, estimating risk parameters should become easier, but more estimation will be required because of the number of lowest-level capabilities.

Financial risk approaches have emphasized the importance of using good measures of risk to summarize and rank risk distributions. The variety of risk measures proposed in the literature suggests that there is no single best equation and decision makers and analysts must choose a measure appropriate for their particular situation and goals (Reesor, 2003).

The methodology in this thesis borrows most extensively from actuarial science, which may be the field most useful in the study of risk in a military context. Insurance firms, with a portfolio of policies, manage multiple risks simultaneously. Some of these risks will result in net losses to the firm, while many will never involve a claim. Similarly, the military must prepare to use many different capabilities even though a significant number—perhaps a majority—will never be employed in combat. The actuarial collective risk model, which considers claim disbursements as a temporal stochastic process, may have application in the assessment of risk in military logistics.

5.2. Methodology Conclusions

The first step in modeling risk is the determination and specification of severity. The CRRA has taken significant steps in this process by identifying eight factors to consider when estimating severity and defining six qualitative severity categories (Appendix A). This thesis proposes a way to translate these categorical severities into an index that can be mathematically manipulated. Categories express severities in rank order; an index shows the relative differences and the ratio of minor severities that are equivalent to catastrophic ones.

Displaying risk visually for decision makers can be accomplished with probability density functions, showing the likelihood over the range of possible outcomes, or complementary distribution functions, indicating the probability of severity exceeding some value. A severity distribution function, $S(x)$, assumes an adverse event will occur and is the probability that severity will be greater than x. A risk distribution function, $R(x)$, modifies this severity distribution function to allow for the possibility that no adverse event occurs. In general this risk distribution function is the preferred mathematical description of risk and can be shown graphically to compare risks as shown in, for example, Figure 26.

This thesis proposes using an exponential distribution or Weibull distribution to model risk. These distributions are simple to calculate, requiring relatively little subjective input, and in the case of the exponential distribution have the intuitively-pleasing property that lower severities are more likely than higher severities. In addition, the exponential and Weibull distributions are used in practice in actuarial science (Vázquez-Abad and LeQuoc, 2001:78).

Many authors have proposed measures of risk to summarize distributions into a single number. This thesis examines four that may be appropriate for measuring military risk. Expectation, the most common risk measure, serves as a baseline for the other three measures. The risk-value measure, combining expectation with standard deviation, only requires distribution moments to be known and not the full distribution. Conditional expectation calculates a risk measure exclusively using the worst possible outcomes, focusing the decision maker on the most extreme severities. All other things equal, however, the preferred risk measure is distorted expectation, using a dual-power distortion measure. This measure calculates using the entire distribution, re-weighting probabilities based on the relative importance of the highest possible severities. The methodology results in a ranking of risks, a measure of risks which provides at least some indication of the relative difference between the riskiness of the various shortfalls, the ability to vary the measure of risk based on decision maker risk aversion.

5.3. Future Research Opportunities

Risk in this thesis only covers negative consequences or severities. This is in accordance with actuarial science and mechanical and environmental engineering risk assessment. The financial literature however, considers risk in both the positive and negative directions. Investment returns can be higher or lower than expected, a factor that must be considered in building a portfolio.

In the assessment of military capabilities, positive risk may be equivalent to redundancy. The CRRA may benefit from considering not only the capabilities where the Air Force suffers from a shortfall, but also those capabilities where a surplus exists. Future study may identify and measure the inefficiencies of this "upside" risk.

This research proposes the exponential distribution, and the related Weibull distribution, as possible models for military risk, primarily because of their simplicity. Future study could confirm the usefulness of these distributions or suggest others as more appropriate. Insurance firms often model claims with a Pareto distribution, for example, in addition to their use of the exponential and Weibull distributions (Vázquez-Abad and LeQuoc, 2001:78). The dual-power distortion is recommended because of its simplicity and interpretability. Future research to confirm the appropriateness of this distortion, or another more appropriate distortion, would be valuable to the military community and the larger risk analysis field (Reesor, 2003).

The CRRA has defined severity according to eight factors. All of the risk distributions and measures considered in this thesis require these factors be simplified into a single dimension of severity. A multidimensional risk distribution might more precisely describe military risk. Other risk measurement techniques would be required, however. Alternatively, a study of how to objectively combine severity values from all eight factors into a single index would enhance the proposed methodology.

Risk management is by nature defensive (Survey of Risk, 2004:13). Enormous sums can be spent on an issue that appears potentially harmful. If nothing negative happens, however, it is not necessarily clear if harm was prevented by the expense or whether the risk management actions were wasteful. A study of the past efforts by the military to prevent or mitigate perceived risk may provide insight into both the assessment of risk and the actions that can be taken to reduce it.

5.4. Final Recommendations

To properly assess risk and include risk as one factor in future Air Force system acquisition decisions, probability must be considered part of the analysis. For some capability shortfalls there may be some probability that no severity occurs. Even if there is certainty that some adverse event will occur if a shortfall exists, the precise severity of the outcome probably cannot be determined. Probability distributions allow this range of possibilities to be described mathematically.

Risks should be analyzed for stochastic ordering. When two risk distribution functions never cross, they are stochastically ordered, and one risk stochastically dominates the other. This does not necessarily mean that the outcome of the lesser risk will always be less than the greater, but the stochastically dominant risk can always be considered more risky.

When stochastic dominance does not occur, or a quantification of risk is required, the distorted expectation risk measure offers a flexible, mathematically rigorous way to summarize the risk distribution in a single number. Unlike some other risk measures, it includes the entire distribution in the calculation, and can be adjusted to reflect decision maker risk aversion.

One of the primary responsibilities of senior leadership—in the corporate world or the military—is the management of risk. Among all the aspects of the future for which leadership must prepare, risk is a particular challenge because it involves a range of possible outcomes and not an exact target (Survey of Risk, 2004:12). This research offers a rigorous approach for the modeling and measuring of risk associated with shortfalls in Air Force capabilities.

Appendix A: Severity Categories and Descriptions

(AFSAA briefing, 2003)

	Minor	Modest	Substantial	Major	Extensive	Catastrophic
Achievement of Objectives	All major objectives achieved. Strong initial strategy requires few/no adjustments. Objectives achieved on time.	All major objectives achieved. Strong initial strategy requires modest adjustments. Few operational delays. Few delays in achieving campaign objectives.	All major objectives achieved, but strategy adjustments required along the way. Some operations slowed. Achievement of a major objective delayed.	One or more major objectives in jeopardy of not being achieved. Several major strategy adjustments required. Advances toward objectives slowed/stalled. Delayed achievement of campaign's major objectives.	One or more major objectives not achieved. Inadequate strategy requires many major adjustments. Advances toward objectives stalled. Major time pressures to achieve objectives and end campaign.	Major objectives not achieved. No strategy adjustments will allow objectives to be achieved. Time pressures force a decision to end the campaign without achieving objectives.
Friendly Casualties	Few citizens/troops killed/ injured. Citizens overseas threatened.	Tens of citizens/troops killed/ injured. Citizens overseas attacked/ injured.	Hundreds of citizens/troops killed/ injured. Citizens overseas attacked/ taken hostage.	Hundreds to thousands of citizens/troops killed/ injured. Citizens overseas killed/ taken hostage.	Thousands to tens of thousands of citizens/troops killed/ injured. Citizens overseas killed/ taken hostage.	Hundreds of thousands of citizens/troops killed/ injured. Many citizens overseas killed/ taken hostage.
Friendly Capability	Air/ land/ sea/ space control unchallenged. No combat losses. All mutual support requests fulfilled.	Superiority achieved in/ over all areas on time; no holdout areas. Enemy capabilities do not disrupt any missions. Almost all requests for mutual support fulfilled.	Superiority in/ over enemy territory delayed; a few holdout areas avoided. Enemy capabilities disrupt some missions. Most requests for mutual support fulfilled.	Superiority in/ over enemy territory not completely achieved; a few areas continuously avoided. A few unanswered challenges from enemy capabilities. Mutual support only for high priority needs.	Superiority in/ over enemy territory limited in area or duration; some areas avoided. Some unanswered challenges from enemy capabilities. Mutual support very limited.	Superiority limited to friendly territory or achieved only for specific missions; significant areas avoided. Major unanswered challenges from enemy capabilities. Unable to provide mutual support.
Friendly Infrastructure	No loss of critical infrastructure.	Local/ limited damage to critical infrastructure. No regional damage or loss.	Local damage to critical infrastructure. No regional damage or loss.	Local damage/ loss of critical infrastructure. Regional infrastructure affected.	Some damage to friendly centers of gravity. Regional damage/ loss of critical infrastructure.	Friendly centers of gravity damaged or destroyed. Widespread damage/loss of critical infrastructure.
Collateral Damage	Few to dozens killed or injured in collateral damage. Local damage/ destruction to buildings/ infrastructure.	Dozens to hundreds killed or injured in collateral damage. City-wide damage/ destruction to buildings/ infrastructure.	Hundreds to thousands killed or injured in collateral damage. Regional damage/ destruction to buildings/ infrastructure.	Thousands to tens of thousands killed or injured in collateral damage. Regional damage/ destruction to buildings/ infrastructure.	Tens of thousands killed or injured in collateral damage. Multi-region damage/ destruction to buildings/ infrastructure.	Hundreds of thousands killed or injured in collateral damage. Widespread damage/ destruction to buildings/ infrastructure.
Enemy Escalation / WMD	Enemy offensives stopped as they are started. No threats to friendly bases. Continuous monitoring of known CBRNE sources.	Enemy offensives stopped in their early stages. Direct, credible threats to friendly bases. Threat of CBRNE use/attack possible.	Enemy offensives make some gains before being driven back. A friendly base attacked and damaged. Credible threat of CBRNE use/attack.	Enemy offensives make significant gains before being driven back. More than one friendly base attacked. Some CBRNE attacks, but we have adequate detection and warning.	Enemy offensives make significant gains. Widespread attacks on friendly bases. Some enemy use of CBRNE weapons. No warning for half the attacks; adequate warning for half the attacks.	Enemy offensives make gains we cannot counter. Widespread attacks on friendly bases. Widespread use of CBRNE weapons with no warning for most attacks. Detection occurs after attack.

	Minor	Modest	Substantial	Major	Extensive	Catastrophic
U.S. National Integrity	No enemy advances toward US territory/ airspace. No terror attacks/incidents on US territory.	No enemy advances toward US territory/ airspace. No terror incidents on US territory.	Embassies fired on. Conventional enemy forces observe US territory/ airspace; are prevented from encroaching. Terror attack with conventional arms/explosives on US territory.	Conflict is non-nuclear but involves terrorism, chemical, bio, or radiological strikes on US territory. Embassies occupied. Conventional enemy forces encroach upon US territory/ airspace, but do not fire on it.	US survives as a nation. Active conflict where enemies fire on US territory/ penetrate US airspace. Single nuclear strike on US territory. CBRNE incidents.	National survival threatened, loss of territorial integrity. Long term exhausting war. Entire nation focused on resolving conflict. Some nuclear strikes on US territory.
U.S. Government Function	State or federal first responders may go on heightened alert. No recovery action(s) required.	State government(s) executes well prepared recovery actions. Federal government assistance not needed.	State government falters occasionally in executing recovery plans. Federal government assistance necessary.	Attack recovery is difficult. Federal government focuses on it above all else. State government focuses on it above all else. Federal government assistance required for response.	Federal and state governments struggle to cope with attack(s). Losing war would mean ideological and cultural realignment.	Survival of a functioning government is threatened. Losing war would mean ceding sovereignty or occupation.

Appendix B: Master Capabilities Library

(AFSAA, 2004)

1. **Surveillance & Reconnaissance.** The capability to successfully conduct surveillance and reconnaissance missions to satisfy Commanders' Priority Intelligence Requirements (PIRs).
   - 1.1. Surveillance. The capability to systematically and continuously observe aerospace, surface or subsurface areas, places, persons, or things, by visual, aural, electronic, photographic or other means. (Joint Pub 1-02)
     - 1.1.1. Conduct maritime surface/terrestrial surveillance
     - 1.1.2. Conduct maritime subsurface/subterranean surveillance
     - 1.1.3. Conduct air surveillance
     - 1.1.4. Conduct space surveillance
     - 1.1.5. Conduct environmental surveillance
     - 1.1.6. Conduct information surveillance
   - 1.2. Reconnaissance. The capability to conduct transitory missions to obtain by visual observation or other detection methods, specific information about the activities and resources of an adversary or potential adversary, or to secure data concerning the meteorological, hydrographic, or geographic characteristics of a particular area. (AFDD 2-5.2)
     - 1.2.1. Conduct maritime surface/terrestrial reconnaissance
     - 1.2.2. Conduct maritime subsurface/subterranean reconnaissance
     - 1.2.3. Conduct air reconnaissance
     - 1.2.4. Conduct space reconnaissance
     - 1.2.5. Conduct environmental reconnaissance
     - 1.2.6. Conduct information reconnaissance
2. **Intelligence.** An integrated capability to provide accurate, timely information and thereby achieve the Predictive Battlespace Awareness (PBA) required to plan and conduct operations. (AFDD 2-5.2) It is the capability to develop information and knowledge as the result of collection, processing and exploitation, analysis, evaluation, and interpretation of available information concerning foreign countries or areas (e.g. geographic, technological, etc.). (Joint Pub 2-01.) General Categories of intelligence include Imagery Intelligence (IMINT), Signals Intelligence (SIGINT), Human Intelligence (HUMINT), Measurement and Signature Intelligence (MASINT) and Open Source Intelligence (OSINT).
   - 2.1. Processing and Exploitation. The capability to exploit and convert raw data info forms of information that can be readily used by intelligence and environmental analysts/experts. Processing and exploitation tasks include initial interpretation, data conversion and correlation, document translation, and decryption, as well as providing the processed information to follow-on phases of analysis.
     - 2.1.1. Interpret and convert IMINT data
     - 2.1.2. Decrypt and correlate SIGINT data
     - 2.1.3. Translate and correlate HUMINT data
     - 2.1.4. Conduct data conversion and correlate MASINT data
     - 2.1.5. Interpret and correlate OSINT data
     - 2.1.6. Process Mapping and Geodesy data
     - 2.1.7. Process and build a coherent picture of the natural environment
   - 2.2. Analysis and Production. The capability to integrate, analyze, evaluate, interpret and fuse processed information to create intelligence and environmental products in the appropriate media that will satisfy the PIRs, other user requirements, or Battlespace Awareness. Information becomes intelligence and environmental impacts knowledge at the conclusion of this phase.
     - 2.2.1. Produce Indications and Warning (I&W)
     - 2.2.2. Produce Current Intelligence

2.2.3. Produce Targeting Intelligence
2.2.4. Produce General Military Intelligence
2.2.5. Produce Scientific/Technical Intelligence
2.2.6. Produce current and predicted environmental impacts knowledge

2.3. Dissemination and Integration. The capability to format and disseminate intelligence and environmental products to the requestor/consumer. The intelligence cycle is complete when the requestor/consumer integrates the intelligence into decision making and planning processes.

2.3.1. Provide Indications and Warning (I&W)
2.3.2. Provide Current Intelligence
2.3.3. Provide Targeting Intelligence, to include Battle Effects Assessments
2.3.4. Provide General Military Intelligence
2.3.5. Provide Scientific/Technical Intelligence
2.3.6. Provide Precise Mapping and Geodesy Information
2.3.7. Disseminate and integrate environmental impacts knowledge

2.4. Predictive Battlespace Awareness. The capability to correlate and fuse patterns of enemy activity and subsequent events to predict adversary intent and/or potential future enemy courses of action. PBA is used to enable effects based planning, execution and assessment of an operation or operations in a theater. Fusing all sources of data/intelligence to produce intelligence assessments inside the enemy's decision loop. Providing this information to commanders in time to protect friendly forces from enemy attack or to maximize the element of surprise.

3. **Command & Control.** The exercise of authority and direction by a properly designated commander over assigned and attached forces in the accomplishment of the mission. Command and control functions are performed through an arrangement of personnel, equipment, communications, facilities, and procedures employed by a commander in planning, directing, coordinating, and controlling forces and operations in the accomplishment of the mission. (JP 1-02 April 2001) Operations requiring C2 include Surveillance & Reconnaissance, Intelligence, Communication, Force Application, Force Projection, Protection, and Preparation & Sustainment.

3.1. Monitor. The reception, monitoring, maintenance, integration, and display of information on global actions, critical events, and crisis areas, to include the status of friendly and non-friendly forces, rules of engagement (ROE), treaties, agreements, and physical environmental conditions.

3.1.1. Receive information from all sources
3.1.2. Monitor information from all sources
3.1.3. Maintain information from all sources
3.1.4. Integrate information from all sources
3.1.5. Display information from all sources

3.2. Assess. Determine the nature and impact of conditions and events to include the military implications of intelligence indicators, environmental effects, and orders of battle. Implies ability to develop total situational awareness, evaluate threats and opportunities and to provide early warning and attack assessment to:

3.2.1. Determine and assess the nature and impact of critical events in the battlespace
3.2.2. Assess status of resources
3.2.3. Assess implications of fused, all source intelligence assessment combined with current and predicted environmental impacts knowledge
3.2.4. Assess events relative to rules of engagement (ROE), treaties and agreements
3.2.5. Assess termination options, conditions, proposals

3.3. Plan. Formulate the operational objectives, generate force lists, and force movement requirements and develop, evaluate, and select courses of action and plans for friendly forces.

3.3.1. Formulate Military Objectives

        3.3.2. Develop potential COAs/Plans
        3.3.3. Evaluate COAs/Plans
        3.3.4. Select COA/Plan
        3.3.5. Merge Generate and tailor force list and force Movement requirements
        3.3.6. Develop Joint Air Operation Plan (JAOP), ISR Collection plans, Air Control Order, Area Air Defense Plan, Air & Space Tasking Order, and other directives and orders as required.
        3.3.7. Coordinate planning with multi-agency partners, including military, national, civil, and commercial organizations
        3.3.8. Plan Military Assistance to Civil Authorities (MACA) Operations
    3.4. Execution Authority. Conduct dynamic battle management and control, and adjust operations as circumstances change.
        3.4.1. Disseminate information
        3.4.2. Convey execution authority for COA/plan
        3.4.3. Retask based on effects based operation assessment
        3.4.4. Interoperate with multi-agency partners, including military, national, civil, and commercial organizations
        3.4.5. Execute Military Assistance to Civil Authorities (MACA) Operations
    3.5. Position, Navigation, Timing
        3.5.1. Provide position
        3.5.2. Provide navigation
        3.5.3. Provide timing

4. **Communications**. The ability to represent transfer, compute, and assure data among persons and machines.
    4.1. Transport Information. Send voice, data, imagery, or video from one location and receive it at another location(s).
        4.1.1. Provide information transport to and from any location on the globe via space, air, terrestrial or subsurface means.
        4.1.2. Prioritized Information Delivery. Based upon commander's quality of service requirements and users' needs.
    4.2. Computing and Enterprise Services. Input, store, retrieve, process, display, access, discover, and output information.
        4.2.1. Store. Retain data in any form, usually for the purpose of orderly retrieval and documentation.
        4.2.2. Retrieve. Find and bring back requested data
        4.2.3. Process. Operate on data with software applications for a specified purpose.
        4.2.4. Display. Present information for use by a person.
        4.2.5. Discover.
    4.3. Assure. Protect and defend information and information systems by ensuring their availability, integrity, authentication, confidentiality, and non-repudiation.
        4.3.1. Information.
        4.3.2. Information Systems.
    4.4. Manage and Control Network Resources and Network Systems.
        4.4.1. Network Management. Provision network resources to meet capacity requirements of the network's users and connected devices.
        4.4.2. Network Damage Assessment/Reconstitution. Automatic or manual methods to detect/assess damage or degradation and return a network to service.

5. **Force Application**. Capability to survive and engage a variety of targets throughout the battlespace by kinetic (nuclear and non-nuclear) and non-kinetic means.
    5.1. Survive and operate against air, space, surface, subsurface, maritime, information, and asymmetric/unconventional threats
        5.1.1. Gain awareness of threat prior to entering enemy detection envelope
        5.1.2. Deceive, disrupt, deny, degrade, or destroy (D5) the enemy F2T2E kill chain

5.1.3. If unsuccessful, mitigate/negate effects of engagement by threat

5.2. Neutralize Threats/Targets – The actions necessary to engage a threat or target assuming F2T2 are complete.

5.2.1. Transit to a weapons employment zone (WEZ) – Move to a position from which a weapon may be delivered against the threat or target.

5.2.2. Deliver weapons – The primary capabilities-based threat/target classes are fixed and moving/movable. Further subclasses include the full spectrum of target types located in all environments subsurface, surface, air, suborbit, space, and the infosphere. In these environments, weapons must achieve desired effects including conventional, nuclear, non-kinetic, counter- CBRNE/Low-observable/HDBT, informational, psychological, permanent, or temporary.

5.2.3. Support weapons as required from target designation through release until fuzing or effective – update threat/target track as required during flyout of weapons to ensure precise effects

5.3. Recover Personnel and Materiel - the capability to locate, authenticate, and recover downed combatants and materiel in enemy (Combat Search and Rescue), neutral, and friendly environments

5.3.1. Report

5.3.2. Locate

5.3.3. Support

5.3.4. Recover

6. **Force Projection**. The ability to project and extend national power (military and non-military) around the globe in a timely manner.

6.1. Rapid Global Delivery. The timely movement, positioning, and sustainment of military forces and capabilities through air and space, across the range of military operations.

6.1.1. Airlift. Worldwide transportation of personnel and materiel through the air which can be applied across the entire range of military operations

6.1.2. Spacelift. The delivery of satellites, payloads and materiel to or through space. Includes the capabilities of routine or on-demand launch and on-orbit repositioning of space-based assets.

6.1.3. Sealift. Worldwide transportation of personnel and materiel via sea mode of transportation which can be applied across the entire range of military operations

6.1.4. Surface Lift. Worldwide transportation of personnel and materiel via ground mode of transportation which can be applied across the entire range of military operations

6.2. Extend Air and Space Operations. The ability to increase range, loiter time, cargo load, payload and orbit life of air and space assets

6.2.1. Air Refueling: Provide the in-flight transfer of fuel between tanker and receiver aircraft for the deployment, employment, sustainment, and redeployment for all refuelable U.S. and coalition aircraft (includes fixed and rotary wing aircraft)

6.2.2. Provide on-orbit servicing: Support the inspection, repair, replacement, and/or upgrade of spacecraft subsystem components and replenish spacecraft consumables (fuels, fluids, cryogens, etc.) by another vehicle.

7. **Protect**. The integrated application of offensive and defensive actions that detect, assess, predict, warn, deny, respond, and recover, preempt, mitigate, or negate from threats against or hazards to air and space operations, critical infrastructure, and assets, and personnel based on an acceptable level of risk. Includes Full Spectrum Threat Response to all threats, including humanitarian and civilian, major accidents, natural disasters, and use of unconventional (including WMD) or conventional weapons.

7.1. Detect. The ability to detect threats to friendly resources (personnel, physical assets, or information).

7.1.1. Sense CBRNE Threats at Point and Stand-off Distances

7.1.2. Detect Health Threats: Ability to detect the effects of select nonweaponized (naturally occurring) physical, biological and chemical threats on personnel and in the environment. Ability to establish baseline levels of naturally occurring agents and health assessments of personnel and to identify increases from the baseline.

7.1.3. Detect Conventional and Unconventional Threats. Detect the full range of threats to Air Force operations, assets, and personnel including surveillance, conventional capabilities, and asymmetric capabilities.

7.1.4. Detect Information Operations Threats

7.2. Assess and Predict. Accurately assess adversary capabilities to be used against friendly personnel, physical assets, or information and precisely derive adversary courses of action planned or employed with the intent to destroy or disrupt operational readiness. Track threat and friendly location in order to predict future actions.

7.2.1. Assess and predict friendly vulnerabilities. Conduct assessments and predictive analysis to identify and predict vulnerabilities.

7.2.2. Predict threat COAs against friendly resources (personnel, physical assets, or information). Conduct predictive analysis of possible enemy COAs for the purpose of effective planning and mitigation.

7.2.3. Assess identified threats. Provide positive identification of threat and assessment of overall capability of the threat.

7.2.4. Track identified threats. Provide decision makers and responders with track/path of threat.

7.2.5. Assess friendly COAs. Conduct assessment of friendly capabilities in order to effectively plan and mitigate potential enemy COAs.

7.2.6. Track friendly forces. Provide decision makers and responders with track of friendly forces.

7.3. Warn. Disseminate threat information in a timely, accurate, and unambiguous manner.

7.3.1. Provide military decision-makers with recommended courses of action. Provide threat working group recommendation to decision makers, from base commander to higher headquarters in a timely, accurate, and unambiguous manner.

7.3.2. Provide civil authorities warning of threat and recommended courses of action. Provide an effective, timely means to communicate with civil authorities. May require foreign disclosure authority.

7.3.3. Provide military/installation populace advanced warning of threat. Provide commander's channel, public affairs, giant voice, email and other means to warn of threat.

7.3.4. Provide civil populace advanced warning of threat. Provide public affairs, email and other means to warn of threat. Off base may require foreign disclosure authority.

7.4. Deny and Respond. Includes Full Spectrum Threat Response. Support and offensively and defensively resist threats directed against friendly personnel, physical assets, or information in order to preserve operational readiness by both active and passive means. Includes Full Spectrum Threat Response to all threats, including humanitarian and civilian, major accidents, natural disasters, hazardous materiel incidents, and use of unconventional (including WMD) or conventional weapons. Respond through pre-emptive, immediate, and sustained actions.

7.4.1. Deny Conventional or Unconventional Threats

7.4.2. Respond. Provide law enforcement and security, fire protection, EOD/WMD, medical response, by lethal and/or non-lethal means, to the full spectrum of emergencies, threats, hostile acts/events.

7.4.3. Provide Assistance to Civil Authorities: Includes Military Assistance to Civil Authorities (MACA) in the US and overseas.

7.4.4. Provide Defensive Information Operations. The protection of critical information systems and infrastructure. Capabiltities that prevent paralysis of critical infrastructure and prevent unauthorized or harmful activities on AF information systems.

7.4.5. Provide Defensive Counterspace. Protect and prevent against Space Threats/Targets and environment – the capability to perform defensive counterspace operations in order to distinguish between attacks and anomalies, withstand and defend systems from attack, and reconstitute and repair space capabilities. (Note: The counter space functions of space surveillance/space situation awareness are under Surveillance and Reconnaissance, Command and Control, Communications master capabilities.)

7.5. Recover. The threat is defeated and recovery actions begin. (Residual threats may still be present). These capabilities include medical treatment and support, damage repair, cleanup actions, and actions to transition back to normal peacetime operations.

7.5.1. Recovery Operations

7.5.2. Medical treatment—restore health

7.5.3. Mortuary Operations

8. **Prepare and Sustain.** Activities required to establish operating locations, generate the mission, support and sustain the mission, and posture responsive forces.

8.1. Open & Establish Operating Locations. Assess, plan, reconfigure, modify, build, and use a supportable infrastructure (industrial, administrative, medical, living) to support the mission, personnel and equipment at specific locations from which operations are projected or supported. This includes expeditionary as well as in-garrison operating locations.

8.1.1. Provide operating location assessments. Collect and assess operational and support infrastructure and security data, and plan for the support of operations from the selected location. Includes: Collect - Collect pertinent pre-deployment data on-location and/or remotely; Survey – Confirm the validity and accuracy of collected data; Assess - Analyze location capability and operational support requirements; and Plan – Plan base lay out and security requirements.

8.1.2. Establish runways, taxiways, ramps, roads, security perimeters, and building sites. Utilize, initiate, build, and modify surface and vertical structures required to bring a base's airfield operating and support infrastructure to a functional condition or preparatory state to accomplish the assigned mission.

8.1.3. Establish utility grid. Utilize, initiate, install, and modify power (electrical), water, and wastewater infrastructure to a predetermined operational state.

8.1.4. Establish communications grid. Utilize, initiate, install, and modify a telecommunications network to a predetermined operational state.

8.1.5. Establish fuel grid. Utilize, initiate, install, and modify a fuel storage and distribution network of tanks, pipelines, and access points for aviation petroleum, oils, lubricant, and propellant requirements to a predetermined operational state.

8.1.6. Establish facilities. Utilize, initiate, construct, modify, and assemble temporary or permanent structure and infrastructure to a predetermined operational state.

8.2. Generate the Mission. Prepare and generate mission elements and payloads; initiate or launch air, space, SOF, information, and HUMRO missions; recover mission elements; and regenerate mission capability repetitively for the full range of mission operations.

8.2.1. Prepare and generate the mission element. repair, configure and inspect, and provide to operations to accomplish the assigned mission.

8.2.2. Configure mission element. Set up for specific mission (configure and load payload)

8.2.3. Support initiation/launch of mission element. Handoff to operator

8.2.4. Recover mission element. Receive and assess status of mission element

8.2.5. Prepare payload. configure for specific mission need (assemble payload, deliver for loading)

8.2.6. Prepare and configure launch and recovery apparatus. Repair, restore, and configure apparatus used for support of mission element initiation, launch, or recovery

8.2.7. Fuel mission element. Direct contact with mission element to provide POL and other propellants required for mission element initiation/launch.

8.3. Support and Sustain the Mission and Forces. "Support" directly assists, maintains, supplies, and distributes forces at the operating location to achieve the mission and maintain the operation of its infrastructure. "Sustain" maintains effective capacities of mission support for the duration of operations worldwide and distributes materiel when the executive agent role falls to Air Force.

8.3.1. Assist mission, forces, and infrastructure. Assure operation of the operating location as a platform for mission elements. (control flightline and airspace traffic, billet forces, medically treat forces, enhance human performance, pay forces, feed forces, minister to forces, administer UCMJ, PERSCO, contract management, agreements, etc.)

8.3.2. Maintain support of mission, forces, and infrastructure. Assure operating capability through repair and preservation of equipment, vehicles, runways, taxiways, ramps, roads and building sites, utility, communications, and fuels grids, facilities, and other infrastructure used in support of mission.

8.3.3. Supply support for mission, forces, and infrastructure. Receive, store, and issue all commodities needed to service and maintain the mission equipment, munitions, support equipment, vehicles, facilities and infrastructure, personnel, medical, service and administrative functions, and communications.

8.3.4. Distribution support for mission, forces, and infrastructure. Transport and deliver personnel, equipment, and commodities to user in processes of mission and support operations. Maintain effective capacities of mission support for the duration of operations worldwide. Reachback repair and resupply Major End Items and components. Provide purchasing and Supply Chain Management, Air Force Specialty (AFS) Functional Management, strategic and operational levels of distribution (in those instances where executive agent role falls to Air Force) and create and maintain Total Asset Visibility.

8.4. Posture Responsive Forces. Define, present, apportion, and process force capabilities, including execution of agreements and prepositioning strategy, to maximize responsiveness and speed of employment.

8.4.1. Define force capabilities. Define common operating and support pictures for global, theater, and operating location current and future operating environments.

8.4.2. Structure force capabilities. Organize and right-size forces to create specified effects as required by the combatant commander. (e.g. UTCs and force modules such as Open the Base, Establish the Base, etc.)

8.4.3. Apportion force capabilities. Assess and allocate force capabilities needed to meet the National Security Strategy objectives of the regional combatant commanders.

8.4.4. Process force capabilities. Form, load, move, receive, and account for the personnel, materiel, and equipment that constitute a capability.

8.4.5. Execute Support Arrangements. Negotiate and put in place interservice, coalition, and/or contract arrangements to assure responsive support.

8.4.6. Execute Prepositioning Strategy. Assess, plan, and place prescribed levels of resources and capabilities at strategic locations to meet required National Security Strategy objectives.

9. **Create the Force**. Organize, train, and equip the combat and support capabilities of the Total Force to meet global combatant commander requirements. Maintain sufficient capacities of created forces.

9.1. Organize Forces

9.1.1. Model, simulate, test, evaluate, and assess responsive forces.

- 9.1.2. Simulate force capabilities to ensure they are fully integrated into training, wargames, experiments, exercises, and operations
- 9.1.3. Define requirements for and establish responsive forces
- 9.1.4. Define requirements for and establish responsive organizations capable of integration with operations, joint, coalition and inter-agency organizations

9.2. Train. Prepare mission-ready graduates capable of providing the best available specialized expertise to the combatant commander
- 9.2.1. Conduct Flying Training
- 9.2.2. Conduct Technical Training

9.3. Educate. Develop airmen, over the span of their career, by integrating enduring leadership competencies and analytical skills
- 9.3.1. Provide Accessions Education
- 9.3.2. Provide Professional Military Education
- 9.3.3. Provide Specialized/Professional Continuing Education
- 9.3.4. Provide Degree Granting Educational Programs
- 9.3.5. Provide Citizenship Education
- 9.3.6. Provide Research and Consultation Programs

9.4. Equip
- 9.4.1. Design, develop, acquire, and modernize force elements; includes equipment, systems and personnel
- 9.4.2. Ensure and maintain, through a combination USAF/DOD agencies, industry and academia, a viable industrial base capable of research, testing, manufacturing, dismantlement, and remanufacturing to produce, sustain and modernize forces
- 9.4.3. Assure the reliability and technological superiority of materiel, equipment, and information
- 9.4.4. Assure and validate weapon system, equipment, item, materiel and IT capability across system life cycles through operational test and evaluation of operational availability and performance requirements

9.5. Recruit and Access. Seek, select, and enter quality people into active duty according to Air Force mission requirements
- 9.5.1. Access Enlisted Personnel
- 9.5.2. Access Officers
- 9.5.3. Access Health Professions and Chaplains

Appendix C: Linear Program Formulations

Potential systems for acquisition.

- Intelligence database (x_1)
- Fighter heads-up display (x_2)
- Standoff missile (x_3)
- Stealthy battle damage assessment unmanned aerial vehicle (x_4)
- Chemical weapon detection equipment (x_5)
- Tankers (x_6)

Linear program formulation with no risk measure included.

Maximize

(0 +0.46+0.34+0.14+0.10+0 +0.16+0.19+0.10) x_1
+ (0.19+0.21+0.19+0.42+0.21+0.16+0 +0 +0.05) x_2
+ (0 +0 +0 +0 +0.92+0.54+0 +0 +0) x_3
+ (0.26+0.12+0 +0.36+0.30+0 +0 +0) x_4
+ (0.26+0.68+0 +0 +0 +0 +0.25+0.31+0.33) x_5
+ (0 +0 +0.23+0.05+0.10+0.11+0 +0.48+0.36) x_6

Subject to:

$7x_1+7x_2+10x_3+9x_4+6x_5+9x_6 \leq 23$

All decision variables binary

Linear program formulation with undistorted expectation risk measure.

Maximize

(0 +4.7*0.46+10.2*0.34+11.8*0.14+13.3*0.10+0 +8.6*0.16+10.5*0.19+10.2*0.10) x_1
+ (14.1*0.19+4.7*0.21+10.2*0.19+11.8*0.42+13.3*0.21+6.8*0.16+0 +0 +10.2*0.05) x_2
+ (0 +0 +0 +0 +13.3*0.92+6.8*0.54+0 +0 +0) x_3
+ (14.1*0.26+4.7*0.12+0 +11.8*0.36+13.3*0.30+0 +0 +0) x_4
+ (14.1*0.26+4.7*0.68+0 +0 +0 +0 +8.6*0.25+10.5*0.31+10.2*0.33) x_5
+ (0 +0 +10.2*0.23+11.8*0.05+13.3*0.10+6.8*0.11+0 +10.5*0.48+10.2*0.36) x_6

Subject to:

$7x_1+7x_2+10x_3+9x_4+6x_5+9x_6 \leq 23$

All decision variables binary

Linear program formulation with low distortion (κ=5) expectation risk measure.

Maximize

(0 +15.2*0.46+34.3*0.34+35.5*0.14+35.9*0.10+0 +31.1*0.16+32.5*0.19+36.9*0.10) x_1

+ (38.8*0.19+15.2*0.21+34.3*0.19+35.5*0.42+35.9*0.21+23.0*0.16+0 +0 +36.9*0.05) x_2

+ (0 +0 +0 +0 +35.9*0.92+23.0*0.54+0 +0 +0) x_3

+ (38.8*0.26+15.2*0.12+0 +35.5*0.36+35.9*0.30+0 +0 +0 +0) x_4

+ (38.8*0.26+15.2*0.68+0 +0 +0 +0 +31.1*0.25+32.5*0.31+36.9*0.33) x_5

+ (0 +0 +35.5*0.23+35.5*0.05+35.9*0.10+23.0*0.11+0 +32.5*0.48+36.9*0.36) x_6

Subject to:

$7x_1+7x_2+10x_3+9x_4+6x_5+9x_6 \leq 23$

All decision variables binary

Linear program formulation with high distortion (κ=10) expectation risk measure

Maximize

(0 +21.4*0.46+49.3*0.34+48.1*0.14+45.5*0.10+0 +46.4*0.16+43.4*0.19+55.3*0.10) x_1

+ (50.0*0.19+21.4*0.21+49.3*0.19+48.1*0.42+45.5*0.21+33.1*0.16+0 +0 +55.3*0.05) x_2

+ (0 +0 +0 +0 +45.5*0.92+33.1*0.54+0 +0 +0) x_3

+ (50.0*0.26+21.4*0.12+0 +48.1*0.36+45.5*0.30+0 +0 +0 +0) x_4

+ (50.0*0.26+21.4*0.68+0 +0 +0 +0 +46.4*0.25+43.4*0.31+55.3*0.33) x_5

+ (0 +0 +48.1*0.23+48.1*0.05+45.5*0.10+33.1*0.11+0 +43.4*0.48+55.3*0.36) x_6

Subject to:

$7x_1+7x_2+10x_3+9x_4+6x_5+9x_6 \leq 23$

All decision variables binary

<u>Bibliography</u>

Air Force Studies and Analyses Agency. Master Capabilities Library. January 2004. www.afsaa.hq.af.mil/php/content/MCL_v54_Final.pdf

-----. "FY04 Capability Review and Risk Assessment (CRRA) Analytic Methodology." Briefing. 16-18 December 2003.

Artzner, Philippe, Freddy Delbaen, Jean-Marc Eber and David Heath. "Coherent Measures of Risk," *Mathematical Finance*, 9: 203-228 (1999).

Ayyub, Bilal M. *Risk Analysis in Engineering and Economics*. Boca Raton: Chapman, 2003.

Benati, Stefano. "The Optimal Portfolio Problem with Coherent Risk Measure Constraints," *European Journal of Operational Research*, 150: 572-584 (2003).

Bennett, Joanna C., George A. Bohoris, Elaine M. Aspinwall and Richard C. Hall. "Risk Analysis Techniques and Their Application to Software Development." *European Journal of Operational Research*. 95: 467-475 (1996).

Bratley, Paul, Bennett L. Fox and Linus E. Schrage. *A Guide to Simulation*. New York: Springer-Verlag, 1987.

Bretschneider, Stuart. "Operations Research Contributions to Evaluation of R&D Projects," in *Evaluating R&D Impacts: Methods and Practice*. Ed. Barry Bozeman and Julia Melkers. Boston: Kluwer, 1993.

Cabelo Semper, J. David and Ismael Moya Clemente. "Value at Risk Calculation through ARCH Factor Methodology: Proposal and Comparative Analysis," *European Journal of Operational Research*. 150: 516-528 (2003).

Castellacci, Giuseppe and Michael J. Siclari. "The Practice of Delta-Gamma VaR: Implementing the Quadratic Portfolio Model," *European Journal of Operational Research*. 150: 529-545 (2003).

Chankong, Vira and Yacov Y. Haimes. *Multiobjective Decision Making: Theory and Methodology*. New York: North-Holland, 1983.

Cho, C.C., P.R. Garvey and R.J. Giallombardo. "RiskNav™: A Decision Aid for Prioritizing, Displaying, and Tracking Program Risk," *Military Operations Research*, 3:25-36 (1997).

Choobineh, F. and A. Behrens. "Use of Intervals and Possibility Distributions in Economic Analysis," *Journal of Operations Research Society*, 43: 907-918 (1992).

Department of Defense. *Dictionary of Military and Associated Terms*. Joint Publication 1-02. 12 April 2001 (as amended through 17 December 2003).

-----. *Procedures for Performing a Failure Mode, Effects and Criticality Analysis*. Military Standard 1629A. 24 November 1980.

Ebeling, Charles E. *An Introduction to Reliability and Maintainability Engineering*. New York: McGraw-Hill, 1997.

Engelbrecht, Joseph A. Jr., Robert L. Bivins, Patrick M. Condray, Merrily D. Fecteau, John P. Geis II and Kevin C. Smith. "Alternate Futures for 2025: Security Planning to Avoid Surprise," April 1996.

Estrada, Javier. "Mean-Semivariance Behavior: A Note," *Finance Letters*, 1: 9-14 (2003).

Eum, Yun Seong, Kyung Sam Park and Soung Hie Kim. "Establishing Dominance and Potential Optimality in Multicriteria Analysis with Imprecise Weight and Value." *Computers & Operations Research*. 28:397-409 (2001).

Fishburn, Peter C. "Foundations of Risk Measurement: Risk as Probable Loss," *Management Science*, 30: 396-406 (April 1984).

Fleischer, G.A. *Engineering Economy: Capital Allocation Theory*. Boston: PWS-KENT, 1984.

Haimes, Yacov Y. *Risk Modeling, Assessment, and Management*, New York: Wiley, 1998.

Hamill, J. Todd, Richard F. Deckro, Jack M. Kloeber and T.S. Kelso. "Risk Management and the Value of Information in a Defense Computer System," *Military Operations Research*, 7: 61-81 (2002).

Henley, Ernest J. and Hiromitsu Kumamoto. *Reliability Engineering and Risk Assessment*. Englewood Cliffs NJ: Prentice-Hall, 1981.

Jablonowski, Mark. "Words or numbers?" *Risk Management*, 41: 47-50 (1994).

Jia, Jianmin and James S. Dyer. "A Standard Measure of Risk and Risk-Value Models," *Management Science*, 42: 1691-1705 (1996).

Jumper, John. "Chief's Sight Picture: Capabilities Review and Risk Assessment," www.xo.hq.af.mil/xor/taskforce/taskforce_docs/csafsightpicture.doc, February 2002.

Kaas, Rob, Marc Goovaerts, Jan Dhaene, and Michel Denuit. *Modern Actuarial Risk Theory*, Boston: Kluwer, 2001.

Kelly, Louise A., Mike A. Taylor and Marion J.A. Wooldridge. "Estimating the Predicted Environmental Concentration of the Residues of Veterinary Medicines: Should Uncertainty and Variability Be Ignored?" *Risk Analysis*. 23: 489-496 (2003).

Klinke, Andreas and Ortwin Renn. "A New Approach to Risk Evaluation and Management: Risk-Based, Precaution-Based, and Discourse-Based Strategies," *Risk Analysis* 22: 1071-1094 (2002).

Kulkarni, Vidyadhar G. *Modeling and Analysis of Stochastic Systems*, London: Chapman & Hall, 1995.

Lambert, James H., Nicholas C. Matalas, Con Way Ling, Yacov Y. Haimes and Duan Li. "Selection of Probability Distributions in Characterizing Risk of Extreme Events." *Risk Analysis* 14: 731-742 (1994).

Landsman, Zinoviy and Michael Sherris. "Risk measures and insurance premium principles," *Insurance: Mathematics and Economics* 29: 103-115 (2001).

Langewish, Andrew and Fred Choobineh. "Stochastic dominance tests for ranking alternatives under ambiguity." *European Journal of Operational Research*. 95: 139-154 (1996).

Law, Averill M. and W. David Kelton. *Simulation Modeling and Analysis*, Boston: McGraw-Hill, 2000.

Markowitz, H.M. *Portfolio Selection*, New York: Wiley, 1959.

McLeish, Donald L. and R. Mark Reesor. "Risk, Entropy, and the Transformation of Distributions," *North American Actuarial Journal* 7: 128-144 (2003).

Mitchell, Douglas W. and Gregory M. Gelles. "Risk-Value Models: Restrictions and Applications," *European Journal of Operational Research*, 145: 109-120 (2003).

National Research Council Commission on Geosciences, Environment and Resources. "Risk Analysis and Uncertainty in Flood Damage Reduction Studies," Washington: National Academy, 2000.

Paté-Cornell, Elisabeth and Seth Guikema. "Probabilistic Modeling of Terrorist Threats: A Systems Analysis to Setting Priorities among Countermeasures," *Military Operations Research*, 7: 5-23 (2002).

President's Commission on Critical Infrastructure Protection. *Critical Foundations: Protecting America's Infrastructures*. October 1997.

Reesor, R. Mark. Department of Applied Mathematics, University of Western Ontario. Personal Correspondence. 18 December 2003.

Rios Insua, David and Simon French. "A Framework for Sensitivity Analysis in Discrete Multi-Objective Decision-Making." *European Journal of Operations Research*. 54: 176-190 (1991).

Sarin, Rakesh K. and Martin Weber. "Risk-Value Models," *European Journal of Operational Research*, 70: 135-149 (1993).

Sarma, Mandira, Susan Thomas and Ajay Shah. "Selection of Risk-Value Models," *Journal of Forecasting*. 22: 337-358 (2003).

Smith, Emily M. "Designing for Sabotage," *Mechanical Engineering*, 124: 40-44 (2002).

"A Survey of Risk," *The Economist*, 370: special section 1-16 (22 January 2004).

U.S. Air Force. "AF Identifies Operational Shortfalls." Press Release. 17 December 2003.

Vázquez-Abad, F.J. and P. LeQuoc. "Sensitivity Analysis for Ruin Probabilities: Canonical Risk Model," *Journal of the Operational Research Society*, 52: 71-81 (2001).

Wackerly, Dennis D., William Mendenhall III and Richard L. Scheaffer. *Mathematical Statistics with Applications*, Pacific Grove CA: Duxbury, 2002.

Wirch, Julia Lynn and Mary R. Hardy. "A Synthesis of Risk Measures for Capital Adequacy," *Insurance: Mathematics and Economics* 5: 337-347 (1999).

Vita

Captain William E. Woodward was born and raised in Seattle, Washington where he graduated from King's High School in 1994. He was a distinguished graduate of the U.S. Air Force Academy in 1998, where he earned a Bachelor of Science degree in operations research and economics, and was awarded eight varsity letters in cross country and track & field. Additionally, he was named the Air Force Academy's 1998 Outstanding Scholar-Athlete. He was selected for graduate study at the University of Maryland, where he interned at the U.S. General Accounting Office and earned a Master of Public Policy degree, specializing in international security and economic policy. In 2000 he moved to Oklahoma City, Oklahoma, where he spent two years at Tinker Air Force Base as an operations analyst for the AWACS surveillance aircraft of the 552d Air Control Wing. Following graduation from the Air Force Institute of Technology, he will move to the Air Force Studies and Analyses Agency in Washington, D.C.

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 074-0188*

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From – To)*
04-03-2004	**Master's Thesis**	May 2003 – Mar 2004

4. TITLE AND SUBTITLE

MEASURING THE RISK OF SHORTFALLS IN AIR FORCE CAPABILITIES

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Woodward, William E., Captain, USAF

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(S)

Air Force Institute of Technology
Graduate School of Engineering and Management (AFIT/EN)
2950 Hobson Way, Building 640
WPAFB OH 45433-7765

8. PERFORMING ORGANIZATION REPORT NUMBER

AFIT/GOR/ENS/04-13

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

AF/XOR
Attn: Col Keith Martin
1480 HQ USAF phone: 703-695-8587
Pentagon DC 20330 e-mail: keith.martin@pentagon.af.mil

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The U.S. Air Force seeks to measure and prioritize risk as part of its Capabilities Review and Risk Assessment (CRRA) process. The goal of the CRRA is to identify capability shortfalls, and the risks associated with those shortfalls, to influence future systems acquisition. Many fields, including engineering, medicine and finance, seek to model and measure risks. This research utilizes various risk measurement approaches to propose appropriate risk measures for a military context. Specifically, risk is modeled as a non-negative random variable of severity. Four measures are examined: simple expectation, a risk-value measure, tail conditional expectation, and distorted expectation. Risk measures are subsequently used to weight the objective function coefficients in a system acquisition knapsack problem.

15. SUBJECT TERMS

Risk; Risk Analysis; Risk Measures

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
					Dr Richard F Deckro (ENS)
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER *(Include area code)*
U	U	U	UU	135	(937) 255-6565, ext 4325; e-mail: richard.deckro@afit.edu