Engineering Management
Field Project

Development and Implementation of a Statistical Risk Assessment Method for New Aircraft Performance at XYZ Corporation

By

Christopher R. Schroeder

Fall Semester, 2009

An EMGT Field Project report submitted to the Engineering Management Program
and the Faculty of the Graduate School of The University of Kansas
in partial fulfillment of the requirements for the degree of
Master's of Science

Tom Bowlin
Committee Chairperson

John Bricklemyer
Committee Member

Ray Dick
Committee Member

Date accepted: 11/9/09

Acknowledgments

First and foremost I would like to thank Troy, a valuable partner in the development of the statistical risk assessment method for new aircraft performance at XYZ Corporation. His expertise and guidance on the use of Design of Experiments, Response Surface Modeling, and the tools MiniTab and Crystal Ball were essential to the creation of this method.

I would also like to thank my wife Kristin for all the support and patience she has shown while I pursued my master's degree. Her love and encouragement during this time provided invaluable support and I could never adequately express my thanks.

To all the staff of Engineering Management at The University of Kansas, thank you for your time and dedication to students and higher learning.

Executive Summary

At XYZ Corporation (XYZ), the standard method of creating aircraft performance guarantees does not incorporate engineering risk or uncertainty. Since performance guarantees are a contractual agreement to the customer, the customer is not obligated to take delivery of the aircraft if the guarantees are not met. This results in considerable financial loss and immeasurable damage to the company's reputation. Because of this, the need for a simple method of evaluating the risk of a performance guarantee arose.

Although a myriad of risk assessment techniques exist in literature, a specific technique for performance guarantees was not available. This research develops a specific statistical risk assessment method (SRAM) that fits with XYZ's tools and culture. By implementing sensitivity analysis, design of experiments, response surface modeling, and Monte Carlo simulation, a Risk / Guarantee matrix can be developed. This matrix compares the level of risk associated with a particular performance value, allowing Management to select a guarantee based upon the amount of risk the company is willing to accept. In its initial implementation, SRAM successfully led Management to select a takeoff field length guarantee more conservative than initially desired due to the initial value's risk. Additionally, the other two desired performance values (range and maximum speed) were found to be within acceptable risk; therefore, they were selected for guarantees.

Potentially the greatest value of SRAM is its ability to evaluate risk in areas outside of performance guarantees. This research found that SRAM's potential for evaluating airplane configuration risk can greatly reduce the probability that a new product falls short of desired performance, ultimately reducing down stream engineering

cost. Identified as an extremely valuable tool, SRAM can potentially be applied in countless other aspects of design and engineering to better understand variation and assess risk.

Table of Contents

List of Figures

List of Tables

List of Symbols and Nomenclature

Symbols

α	Risk of Performance Shortfall
α_o	Risk of Data Outside of Confidence Interval
η	Performance Value
σ	Standard Deviation
C_n	Coefficient of the n^{th} Parameter
K	Constant of the Response Surface Equation
L	Number of Levels in an Experiment
P	Number of Parameters in an Experiment
x_n	n^{th} Parameter

Nomenclature

User	Refers to the One Implementing the Method of Interest
SRAM	Statistical Risk Assessment Method

Units

kts	Knots (nautical mile per hour)
nm	Nautical Mile

Chapter 1 – Introduction

XYZ Corporation (XYZ), a business aircraft manufacturer, has a reputation of building and selling some of the finest aircraft available. When a new aircraft is developed, it is typical practice in the industry to begin sales of the product before a flying prototype is built. To boost sales and quell uncertainty, the company always makes three guarantees to a customer as a "promise" of what they will receive when the aircraft is delivered. These three guarantees are: aircraft range, takeoff field length, and maximum speed.

Since a prototype of the new aircraft is not available when these guarantees are made, a heavy responsibility lies on Engineering to successfully predict these performance points. If a performance guarantee is not met when an aircraft is delivered, it is considered a breech of contract and the customer is no longer required to take delivery of that aircraft. This can have serious financial impact since the customer has not yet made their final payment and can seek legal action to recover their initial down payment. Furthermore, damage to XYZ's reputation causes immeasurable financial impact.

In its current state, the method for guarantee generation only requires Engineering to use nominal (most likely) input parameters when modeling aircraft performance. A nominal input results in a nominal output, thereby giving the user no gauge of the risk inherent in its value. Marketing and Sales apply a standard $\pm$ 3% to the guarantee to account for any variation in manufacturing, yet this does not cover engineering uncertainty since it is assumed engineering guarantees will be met. The goal of this research is to develop a method that takes a statistical approach to risk assessment so data-based-decisions can be made about aircraft performance. Risk assessment should be based on engineering uncertainty and allow the user to observe the final output and select a level of aircraft performance based upon a corresponding level of risk.

In addition, the potential for using the resulting statistical risk assessment method (SRAM) beyond risk assessment of performance guarantees will be discussed briefly in Chapter 4 – Results.

Chapter 2 – Literature Review

A literature review was conducted to identify any defined methods for statistical risk assessment and determine if they can successfully be applied to the stated problem. The literature search started with no specific focus given to the industry of concern. This was done with the intent of identifying different methods that could be applied to the aviation industry. Following the broad search, a more focused search of risk analysis in aviation was conducted. The last step was a final highly focused search on statistical risk assessment of aircraft performance. This final search yielded results in aircraft component risk but no results for overall aircraft performance risk.

Since no specifically defined method for statistical risk assessment of aircraft performance could be found, the need for such a method was affirmed. The following review addresses the strengths, weaknesses, and potential application of each text as it relates to risk assessment.

General Research

The first text reviewed was a paper entitled "Empirical Evaluation of the Association between Methodological Shortcomings and Estimates of Adverse Events" (Chou, Fu et al. 2006). This is a medical paper focusing on the risk of adverse effects of new prescription drugs. Although this paper presents strong use of statistical risk assessment, it is based on sampling. The main method for sampling is completed through clinical trials, in which the results of a drug are observed on a small group of volunteers. The results of multiple trials are then collaborated to represent the reaction of the entire public.

Although valid for evaluating the risk of a new drug, sampling is not a method that can be used for the stated problem. The goal is to accurately account for the risk before a prototype is developed. Even if sampling was completed on the prototype, this is only one sample point – far too small for a legitimate statistical sample. In addition, it would take years to gather enough data to create a viable model since only 20 to 30 aircraft are built a year. For these reasons, statistical sampling cannot be used to solve the proposed problem and as a result will not be considered for the problem solution.

A paper by Steve Verrill and Richard Johnson entitled "Confidence Bounds and Hypothesis Tests for Normal Distribution Coefficients of Variation" (Verrill and Johnson 2007) was the next to be reviewed. The paper was written for the United States Department of Agriculture with the intent of identifying the risk of chemical concentrations in particular environments. This paper is very similar to the medical example presented previously, due to its use of statistical sampling. To determine the risk of chemical concentrations, sampling is completed at areas of interest where the chemicals already exist. For the same reasons as in the medical text, the methods utilizing sampling for statistical risk assessment cannot be applied to the stated problem.

Degrees of Belief: Subjective Probability and Engineering Judgment (Vick 2002), a book on geotechnical engineering was the next item to be reviewed. The first statistical risk analysis method this book examines is the decision tree. A decision tree allows the user to compare multiple paths given the probabilities of potential events and their resulting consequences. An example decision tree is shown in Figure 1.

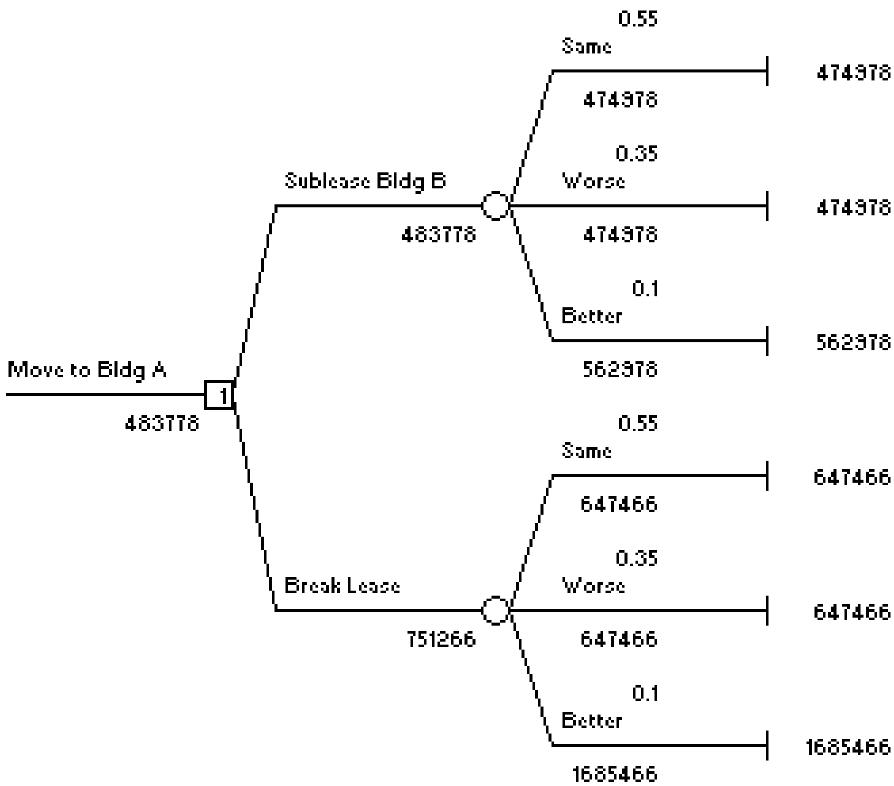


Figure 1: Example Decision Tree

Although the Decision Tree Method is successful at predicting the "best" outcome given the risk (probability) of each event, it requires the user to know (or have a good idea) of the probability of each event occurring. Knowing exact probabilities in XYZ Engineering is unlikely, however, the probability of variation is known. For example, an engineer may not be able to identify the probability a new engine will be delivered under design specification; however, they will be able to specify the variation bounds of an expected new engine. Since the decision tree requires exact probabilities, it is not considered a practical solution to the stated problem.

The decision tree method does have strength in enabling the user to see the probability of each defined outcome and compare each outcome's likelihood against another. This has value when identifying the risk of a given performance point.

The next method identified in Vick's book is the Likelihood / Consequence Matrix. In this method, the user creates a matrix where one axis is Likelihood and the other Consequence. Each axis has numeric scoring to identify the severity of either the likelihood or consequence of a given scenario. An example Likelihood / Consequence Matrix is shown in Figure 2.

	CONSEQUENCE				
LIKELIHOOD	Insignificant (1)	Minor (2)	Moderate (3)	Major (4)	Extreme (5)
Rare (1)	Low	Low	Low	Low	Low
Unlikely (2)	Low	Low	Low	Medium	Medium
Possible (3)	Low	Low	Medium	Medium	Medium
Likely (4)	Low	Medium	Medium	High	High
Almost certain (5)	Low	Medium	Medium	High	Extreme

Figure 2: Example Likelihood / Consequence Matrix

Although the Likelihood / Consequence Matrix is useful in determining a two-dimensional aspect of risk (likelihood and consequence), it cannot be applied to the stated problem. This is due to the fact that the stated problem is only concerned with one consequence, the failure to meet a guarantee. Also the Likelihood / Consequence Matrix is more qualitative than quantitative and does not meet the statistical assessment aspect of the problem statement.

The last subject of interest in this text is more of a general guideline for risk assessment. In the process of risk assessment it is important to use the input of Subject Matter Experts (SMEs). SMEs are used to identify the probability of events or the values of likelihood and consequence. The SME's value comes from their experience and something the author calls the "hard-easy". "Hard-easy" refers to a situation where a difficult question is asked of someone with little subject knowledge and they perceive the question as easy. The less the individual knows about the subject, the quicker and less educated the response. This appears to be a result of not knowing how little they know. Due to a SME's experience however, they know what they do not know and therefore give better estimates. SMEs are recognized as a valuable part of evaluating risk in guarantees and can play a key role in the solution to the stated problem.

The next book examined was *Operational Risk: Modeling Analytics* (Panjer 2006). This book was written with a focus on risk in banking and the insurance sector. The first point in Panjer's book explains that the definition of risk is something that results in financial loss. All other "risks" are ignored. This definition aligns with the stated problem since financial loss is the company's only concern as a result of a missed guarantee.

To successfully identify risk, the author suggests creating a model of the system in question to explore potential outcomes. Since the creation of guarantees is completed through modeling (no aircraft exists yet, all predictions are made using computer models) this method is well received. The author advocates varying inputs to the model to understand their effect on the total outcome. In addition, the author suggests running the model multiple times with the same inputs to identify the variation in the model itself. The model used for performance predictions is based on the physics of flight; therefore, given constant inputs, the output will always be the same. For that reason, this research does not need to consider model variation. However, the point about input parameter variation is very valuable. The "real-world" variation of a parameter is what can cause a missed guarantee. The author also explains that by understanding the output variation, the user can utilize standard deviation and confidence intervals to understand the risk of the system. These methods appear to have great value in evaluating the risk of aircraft performance modeling.

Risk Modeling for Determining Value and Decision Making (Koller 2000), a general book on risk analysis was the next book to be reviewed. The main discussion in this book revolves around creating a model, much like what was suggested in the previous text. Before modeling is discussed however, more simple methods such as decision tree analysis and linear programming are discussed.

Mentioned previously, decision tree analysis does not work for the stated problem because of its requirement for event probabilities and the deterministic outcome. Linear programming offers more flexibility since the interaction of input parameters can be more clearly understood. However, since a working model of the aircraft already exists, linear programming is passed over to give more focus on the modeling section of Koller's text.

To develop a successful model, the author suggests identifying the input parameters that contain risk. Although it is important to understand all of the inputs required to create a successful model, by specifically noting the parameters with variation (risk), the user can begin to focus on what causes output variation. Once the risk parameters are defined, the user should identify the maximum, minimum, and most-likely value for each parameter. This process fits well with the engineering practices at XYZ. As stated previously, an engineer may not understand the probability that a parameter will not be accurate; however, SMEs can identify the maximum, minimum, and most-likely value of a given parameter. This is considered a strength of the modeling method.

Along with maximum, minimum, and most-likely value of a parameter, the author suggests assigning a peakedness value as well. The peakedness value is an assignment of likeliness, giving each parameter weight in the overall model. Peakedness is not necessary for the stated problem since the existing model already "assigns weights" to input parameters through the equations of the model (the model is based on physics of flight). Therefore this suggestion is not needed for the stated problem.

After the model is developed and the input parameter variation is known, the author recommends using Monte Carlo simulation to statistically vary the inputs to identify the overall variation (risk) in the output. There are seven statistical distributions that can be applied to input parameter variation. These distributions are: symmetrical, skewed, spike, flat, truncated, discrete, and bimodal. From the result of the Monte Carlo analysis, the user can identify confidence intervals on the outcome to identify the risk of a given output value. This technique appears to have excellent application for the proposed problem. Engineering can identify variation in input parameters and the user can then identify the risk of a resulting guarantee. A weakness of this method, however, is that the model used for performance predictions does not directly "link" to any commercial Monte Carlo simulation. Therefore Monte Carlo simulation is not a straightforward process.

Very similar to the two previously discussed texts, "Framework for Quantifying Uncertainty in Electric Ship Design" (Porche, Willis et al. 2004) uses modeling and Monte Carlo simulation to determine final system risk. This document was written specifically for risk assessment of different electric propulsion systems for Navy ships. The authors specify that the first step in risk analysis is to identify the input parameters that contain variation. After input variation is identified, the model is then linked to a Monte Carlo simulation that runs through the variation of input parameters to determine the resulting performance output variation. This report adds more validity to the use of modeling and Monte Carlo simulation by concurring with the other texts already reviewed. An apparently powerful method, the downfall for the proposed problem remains that the existing performance model cannot link to a Monte Carlo simulation.

Industry-focused Research

The next resources examined focus on the aviation industry. The first in this area is entitled *Risk Management and Error Reduction in Aviation Maintenance* (Patankar and Taylor 2004). This book concentrates on the application of risk management in aviation systems with the purpose of improving system safety to prevent incidents/accidents. The method promoted by this book is the SHELL model. Where SHELL represents controls capable of improving safety: S-software, H-hardware, E-environment, L-liveware (liveware = human in the loop). The discussion in Patankar and Taylor's book revolves around the interaction of each SHELL component and how they can be implemented to improve aircraft safety, thereby reducing risk. This type of risk analysis and mitigation does not apply to the stated problem; therefore, the methods of this text are not considered as a potential solution.

Introduction to Aviation Insurance and Risk Management (Wells and Chadbourne 1992) is a book that centers on risk assessment in aviation insurance. Similar to the previous banking and insurance text that was reviewed, this book specifies that risk is only defined as potential financial loss. No other type of risk should be considered. As mentioned before, this definition of risk does agree with the stated problem. XYZ's only concern about missing a guarantee is the financial loss that comes as a result.

A method for understanding potential risk in insurance is to evaluate past data for a similar situation (data sampling) and establish the standard deviation and confidence intervals for a given scenario. Since sampling is not a viable method for the stated problem, this approach cannot be used. The use of standard deviation and confidence intervals agrees with previously reviewed texts, so this piece of risk assessment still has value.

Wells and Chadbourne's text also discusses methods for risk control: inspection, safety programs, and training. Although these are valid methods that could be used to improve manufacturing, thereby minimizing the variation that exists in the model's input parameters, this is outside of the scope of the research. It is not the intent of this research to improve manufacturing; rather, the problem focuses on understanding the risk that currently exists prior to manufacturing and properly assessing the magnitude of risk. Therefore, these methods will be left for further discussion in Chapter 5.

The last text to be reviewed in this search was *Safety Management Systems in Aviation* (Stolzer, Halford et al. 2008). According to the authors, risk assessment begins with a subjective view using a Likelihood / Severity Matrix (see Figure 2). As described in a previous review, this matrix technique does not apply to the stated problem.

A more detailed approach to risk assessment is the use of a model to predict variation. The author suggests using probabilistic risk assessment such as Monte Carlo simulation to evaluate a system's risk. Appearing as a common theme in many texts, this method fits well with the engineering approach at XYZ because all of the required inputs are available, but the current model does not directly link to a commercial Monte Carlo simulation. This can be considered a shortfall for the stated problem.

Aircraft Performance Risk Research

A final search was conducted focused on statistical risk assessment of aircraft performance. It yielded a paper entitled "Risk-Based Probabilistic Approach to Aeropropulsion System Assessment" (Tong 2002). This paper outlines a method to determine the probable variation of an aircraft engine during development. The document, written for NASA, assumes that a model has already been developed to predict the performance of an engine. XYZ's existing performance model allows for such a process. With an established model, the inputs and the input variation are then identified so they can be implemented in what the author calls "the fast probability integration technique (FPIT)". FPIT is a NASA built, Monte Carlo-style tool that links to the model to simulate thousands of input variations. The simulation output can then be studied to determine the overall system variation. Mentioned in many of the reviewed texts, the concepts defined by this author appear to have vast potential for the stated problem. The only major deficiency that still remains is the inability to link XYZ's performance model to a Monte Carlo simulation.

Key Findings

Substantial value was found in the literature review as common themes for risk assessment emerged. The method of modeling the system matches well with the stated problem since the process of creating guarantees already requires a model. By identifying the required input parameters and understanding those that contain variation (risk), the variation (risk) of the output can be understood. Also, SMEs can be utilized to assign variation to parameters, which ultimately bound the problem. Since a statistical risk assessment method as it applies to total aircraft performance was not found in current literature, the need for such a method was reaffirmed.

Chapter 3 – Research Procedure

The guidance from the literature review was used to develop a research procedure to create a method that is capable of assessing aircraft performance risk. The intent of this research was to create a method that integrates well with XYZ processes and engineering tools. The various stages of the research process are outlined below.

The first stage of the research procedure was to identify the performance programs that are available within XYZ. The program selected was then used to model and predict the aircraft performance. Selection was based on which program provided the greatest ease of use and functionality for statistical risk assessment of performance guarantees.

Once the performance program was selected, the input parameters that cause performance variation were identified. It is important to identify every parameter that can cause performance variation. If all variation sources are not recognized, the statistical assessment will be an incomplete and inaccurate representation of risk.

With the input parameters identified, SME expertise was utilized to provide the expected variation of each parameter. All input parameters can be categorized under an engineering discipline such as aerodynamics, structures, weights … etc. For each discipline, there is at least one SME available to indicate the expected variation of a given input parameter. The variation specified by the SME should account for all the expected engineering variation in a given parameter; however, the estimate should not be overly conservative or contain excessive variation.

Next, the tools within XYZ that are capable of Monte Carlo simulation were identified and evaluated. The Monte Carlo tools were assessed on their simulation speed and their versatility for modeling different variation distributions. Additionally, the tool's ability to communicate with the selected performance program was reviewed and weighed into the selection of the Monte Carlo tool used.

Once the Monte Carlo tool was selected, different methods of linking the performance program and the Monte Carlo simulation were explored. The methods available within XYZ capable of connecting the performance program and Monte Carlo simulation were first identified and then evaluated. Ease of implementation and ability to provide an interface between the performance program and Monte Carlo simulation was assessed. The best option was selected for the SRAM method.

The final step was to identify a presentation method to summarize the Monte Carlo output and communicate the risk of a given performance value in a straightforward manner. By reviewing the output data from the Monte Carlo simulation and the presentation styles of the output, an easy to understand form was created to express the risk in aircraft performance values.

Chapter 4 – Results

Following the steps outlined in Chapter 3, a method capable of assessing the statistical risk of an aircraft performance guarantee was developed. The first step in developing SRAM was the identification and assessment of XYZ's tools that are capable of modeling and predicting aircraft performance.

Overview of Existing Performance Model

Numerous computer programs capable of modeling aircraft performance are available at XYZ. The performance program selected is the same tool that is used to create the Pilot's Operating Manual (POM) for delivered airplanes. By using this same performance program for SRAM, it ensures that the SRAM performance values are compatible with information published in the POM. The tool is known as the Computerized Aircraft Model (CAM) and was developed within XYZ. CAM takes user-defined inputs for aircraft characteristics and uses physics equations to calculate aircraft performance. For instance:

Takeoff Field Length = *f* (weight, thrust, drag, etc…)

The form of the CAM model does not change for different aircraft; it remains constant and the input values are adjusted to represent different aircraft. At the highest level CAM acts as a simple Input/Output Code – supply the required inputs, specify the desired performance calculation, and the performance value is calculated for the user. CAM is a powerful tool and is able to calculate many aspects of airplane performance.

Identification of Input Parameters

Identifying all the potential inputs that cause performance variation can appear impractical due to the perceived enormity of the task. However, the task is made relatively simple since the CAM software does not change from model to model. The input parameters that cause variation are simply the inputs required in CAM for a particular performance calculation. For example, the only inputs required to calculate maximum speed are: weight, thrust, and drag. Therefore, only these parameters need to be considered when determining input variation for the Monte Carlo simulation.

Furthermore, only parameters that cause significant performance variation need to be considered. By completing a relatively simple sensitivity study, the parameters with the most significance can be identified. To do this, CAM is run through a series of cases where each input parameter is varied one at a time. The resulting change in the output helps to identify the parameters with very little or no impact. These insignificant parameters are not considered to have variation, which helps to simplify the Monte Carlo simulation.

Monte Carlo Simulation

Two different tools available within XYZ are capable of Monte Carlo simulation and possess linking capability. The tool that provided the best functionality and speed was Crystal Ball, developed by Decisioneering Inc. Crystal Ball allows the user to define the maximum and minimum variation and expected distribution (normal, flat, skewed, etc…) of input parameters; it then uses the probable variation of the inputs to calculate the output variation for numerous cases. This yields a statistical distribution of the output and allows the user to observe variation that occurs due to input uncertainty.

The ideal implementation of a Monte Carlo simulation is to directly link to a model and allow the simulation to run thousands of cases. Although Crystal Ball has linking capability, it is incapable of linking to CAM. Additionally, some calculations in CAM are computationally intensive and take several minutes to obtain a result. This is not conducive to a 1000+ case simulation. Both of these disadvantages are shortfalls for the selected tools, so further attention is given to solving these issues.

Response Surface Modeling

By using a response surface representation of a performance calculation, both the computational time and direct-link to Crystal Ball issues can be resolved. A response surface is an equation that represents the relationship of known inputs and outputs. It can be likened to a trend-line equation that is fit to a set of data. A linear relationship was found to be adequate for cases used to develop guarantees. It is recognized that some scenarios may be complex enough to require a higher-order representation. Since a linear equation was acceptable, it was used for this research due to its simplicity. The equation developed is of the form:

$$\eta = C_1 \cdot x_1 + C_2 \cdot x_2 + C_3 \cdot x_1 \cdot x_2 + \ldots + C_n \cdot x_n + K$$

where: η is the performance value (range, maximum speed, etc...)
C is the coefficient of x
x is an input parameter (drag, weight, thrust, etc…)
and K is a constant

This equation can be used directly with Crystal Ball. Plus, the utilization of a single equation allows for thousands of case iterations in a short amount of time.

The method for obtaining a response surface is relatively simple and for this research was accomplished by using MiniTab 15, a product of MiniTab Inc. All that is required to create a response surface in MiniTab is the minimum and maximum values that bound the variation of input parameters and the resulting outputs for combinations of inputs. This information defines the relationship between inputs and outputs. MiniTab then uses this information to define the K and C terms of the response surface equation. It is important to note, that by selecting the minimum and maximum value of input parameters, the response surface is limited to evaluating input variation that lies within the bounds used to create the surface. Extrapolation outside of the input extremes can lead to inaccurate results.

The most robust method for creating a response surface is known as a full factorial experiment. This is completed by using CAM to calculate the performance output for every permutation of maximum and minimum inputs. Calculations are completed for range, takeoff field length, and maximum speed, yielding three different response surface equations. Unfortunately, to do a full factorial representation of some performance values, 64 separate cases would be required in CAM. With computational time of 10 to 20 minutes for some cases, a full factorial representation of the performance calculation was far too prohibitive

Design of Experiments

To accelerate the development of the response surface, Design of Experiments (DOE) was used to reduce the number of CAM runs required to create a response surface. Table 1 shows a representation of a full factorial experiment with two parameters.

Table 1: Example Full Factorial Experiment

Experiment #	x_1	x_2	Performance Output
1	Max	Max	η_1
2	Max	Min	η_2
3	Min	Max	η_3
4	Min	Min	η_4

It is shown that each parameter varies through its maximum and minimum value for every possible permutation of inputs x_n. This yields every possible maximum and minimum permutation of the performance output η_n. The number of experiments in a full factorial depends on the number of parameters and the number of levels for each parameter (e.g. 2; "max" & "min"). The number of experiments required can be found using the following equation:

Number-of-Experiments = L^P

where: L is the number of levels
P is the number of parameters

For example, with as few as two levels and 10 parameters, the number of cases required for a full factorial is 1024.

By implementing a fractional design, the number of runs required to develop a response surface can be greatly reduced. Table 2 shows the number of experiments required for a full, half, and quarter factorial.

Table 2: Cases Required for Fractional Factorial

	Full Factorial	DOE ½ Factorial	DOE ¼ Factorial
4 parameters, 2 levels	16	8	N/A
6 parameters, 2 levels	64	32	16
8 parameters, 2 levels	256	128	64

Some fidelity of the resulting response surface is lost when a fractional factorial is implemented. The two values highlighted in Table 2 represent the limits of a fractional design, where the information excluded from the design starts to adversely affect the accuracy of the response surface. For the purpose of this research, if a fractional design from the highlighted area was used to develop a response surface, the equation was evaluated at multiple points and checked against actual CAM outputs to verify its accuracy. In general, this is good practice when implementing marginal DOEs.

Example Method Use

Having established the methods required to allow Monte Carlo simulation of the performance calculations, the method was implemented on a current development project. A new XYZ product, X-03, was used as the basis for development and implementation of SRAM. The first step of the method is to define the input parameters required to calculate each performance output. This information is taken directly from the CAM user interface. Then, detailed discussions with SMEs are held to establish the maximum and minimum value for each input parameter.

The maximum and minimum specified are then used as the maximum and minimum of a BetaPERT distribution (one of the possible distribution shapes used by Crystal Ball). When the distribution is considered symmetrical, the maximum and minimum value represent +/- three standard deviations ($\pm 3\sigma$) from the mean. $\pm 3\sigma$ represents 99.7% of the variation expected in that parameter. In other words, 99.7% of the data falls between the $\pm 3\sigma$ bounds; this is known as a confidence interval (CI). More attention will be given to CIs later in this section

Once the parameters and their maximum and minimum values are identified for each performance calculation, the sensitivity of each parameter is determined by running CAM at each input's maximum one at a time. When completed, this data is collected in a table for each performance calculation. Table 3 shows the parameters and sensitivities of range for X-03. Parameter names and units have been removed and values adjusted to maintain company confidentiality.

Table 3: Parameters and Sensitivity of Range

Parameter	Minimum	Nominal	Maximum	Sensitivity (nm/unit)
x_1	8960	9117	9198	0.1
x_2	1451	1481	1510	22.6
x_3	3631	3646	3652	0.68
x_4	5880	6000	6120	20.8
x_5	0.0294	0.0300	0.0309	4.6
x_6	0.0379	0.0389	0.0416	5.4

When observing the sensitivity, it is important to recognize that it is nautical miles (nm) per unit of change. Though parameter x_1 and x_3 have relatively small sensitivities, the unit change that occurs from their maximum and minimum is still considered significant. The total variation of x_3 results in a range spread of only 14 nm, yet it was still considered significant enough to remain in the DOE. Therefore, all the parameters shown in Table 3 were considered for the creation of the response surface.

A full factorial for 6 parameters (x_1 through x_6) and 2 levels (max & min) requires 64 cases, and range calculations in CAM can take upwards of 20 minutes for one case. Therefore, at a total time exceeding 21 hours, it was determined that a fractional factorial be used. A quarter factorial was finally selected due to the time investment still required for a half factorial. Since this design is at the limit of the DOE (see Table 2), it was checked against CAM calculations to verify its accuracy.

Next, takeoff field length was evaluated. Table 4 shows the parameters and sensitivities associated with takeoff field length.

Table 4: Parameters and Sensitivity of Takeoff Field Length

Parameter	Minimum	Nominal	Maximum	Sensitivity (ft/unit)
x_1	8960	9117	9198	0.5
x_2	93	95	97	0
x_3	90	92	94	0
x_4	1.035	1.054	1.073	83
x_5	0.475	0.5	0.525	6
x_6	0.0351	0.0355	0.0359	0.6
x_7	13.1	13.2	13.3	0
x_8	5880	6000	6120	36
x_9	0.0294	0.0300	0.0309	0.8
x_{10}	0.0379	0.0389	0.0416	4

Although the takeoff field length performance calculation starts off with 10 parameters, which would result in 1024 cases for a full factorial, sensitivity analysis identifies 6 insignificant parameters that can be removed from the design. The parameters removed from consideration are highlighted in gray in Table 4. By removing unimportant parameters, a full factorial can be completed with just 16 cases.

Finally the maximum speed parameters were considered. Their values and sensitivities are shown in Table 5.

Table 5: Parameters and Sensitivity of Maximum Speed

Parameter	Minimum	Nominal	Maximum	Sensitivity (kts/unit)
x_1	8960	9117	9198	0.002
x_2	5880	6000	6120	1.26
x_3	0.0294	0.0300	0.0309	0.34
x_4	0.0379	0.0389	0.0416	0.04

Initially with four factors, sensitivity analysis suggests a two parameter DOE. This results in only four required cases for a full factorial. The parameters removed from the DOE are once again highlighted in gray.

Having the risk parameters defined for each performance calculation and the optimal factorial design selected, the DOEs for generating the three response surfaces were created. By specifying the number of parameters, levels, and factorial type (full, half, or quarter), MiniTab 15 will generate a corresponding DOE outline. The DOE-outlines for range, takeoff field length, and maximum speed are shown in Tables 6 through 8.

Table 6: Quarter Factorial DOE for Range

Case #	x_1	x_2	x_3	x_4	x_5	x_6	Range
1	Min	Min	Min	Min	Min	Min	η_1
2	Max	Min	Min	Min	Max	Min	η_2
3	Min	Max	Min	Min	Max	Max	η_3
4	Max	Max	Min	Min	Min	Max	η_4
5	Min	Min	Max	Min	Max	Max	η_5
6	Max	Min	Max	Min	Min	Max	η_6
7	Min	Max	Max	Min	Min	Min	η_7
8	Max	Max	Max	Min	Max	Min	η_8
9	Min	Min	Min	Max	Min	Max	η_9
10	Max	Min	Min	Max	Max	Max	η_{10}
11	Min	Max	Min	Max	Max	Min	η_{11}
12	Max	Max	Min	Max	Min	Min	η_{12}
13	Min	Min	Max	Max	Max	Min	η_{13}
14	Max	Min	Max	Max	Min	Min	η_{14}
15	Min	Max	Max	Max	Min	Max	η_{15}
16	Max	Max	Max	Max	Max	Max	η_{16}

Table 7: Full Factorial DOE for Takeoff Field Length

Case #	x_1	x_2	x_3	x_4	Takeoff Field Length
1	Min	Min	Min	Min	η_1
2	Max	Min	Min	Min	η_2
3	Min	Max	Min	Min	η_3
4	Max	Max	Min	Min	η_4
5	Min	Min	Max	Min	η_5
6	Max	Min	Max	Min	η_6
7	Min	Max	Max	Min	η_7
8	Max	Max	Max	Min	η_8
9	Min	Min	Min	Max	η_9
10	Max	Min	Min	Max	η_{10}
11	Min	Max	Min	Max	η_{11}
12	Max	Max	Min	Max	η_{12}
13	Min	Min	Max	Max	η_{13}
14	Max	Min	Max	Max	η_{14}
15	Min	Max	Max	Max	η_{15}
16	Max	Max	Max	Max	η_{16}

Table 8: Full Factorial DOE of Maximum Speed

Case #	x_1	x_2	Maximum Speed
1	Min	Min	η_1
2	Max	Min	η_2
3	Min	Max	η_3
4	Max	Max	η_4

Each experiment for the three performance calculations was run in CAM and the resulting performance (η_n) recorded. Once all the experiments were completed, Tables 6 through 8 (with η_n) were put into MiniTab 15 where the equations for the three response surfaces were generated. The equations were then transferred to Crystal Ball to be utilized by the Monte Carlo simulation.

In Crystal Ball, the input parameters are defined by their maximum, minimum, most likely value, and desired distribution shape. This information is available in Tables 3 through 5 and a BetaPERT distribution was selected for all parameters. Once the parameter information and response surface equation were properly setup in Crystal Ball, 50000 trials were run to develop an output variation of each performance guarantee.

The CIs of the resulting statistical distribution were then used to identify the risk of a corresponding performance value. Since a CI defines what amount of data is known to be within a specific number of standard deviations of the mean, it also defines what is known to be outside of that number of standard deviations. Table 9 shows the CIs for different standard deviations (Solutions 2008).

Table 9: Standard Deviations and Corresponding Confidence Intervals

Standard Deviation	Confidence Interval
1σ	68.3%
2σ	95.5%
3σ	99.7%

Risk in general is then defined as the amount of data not accounted for in the CI.

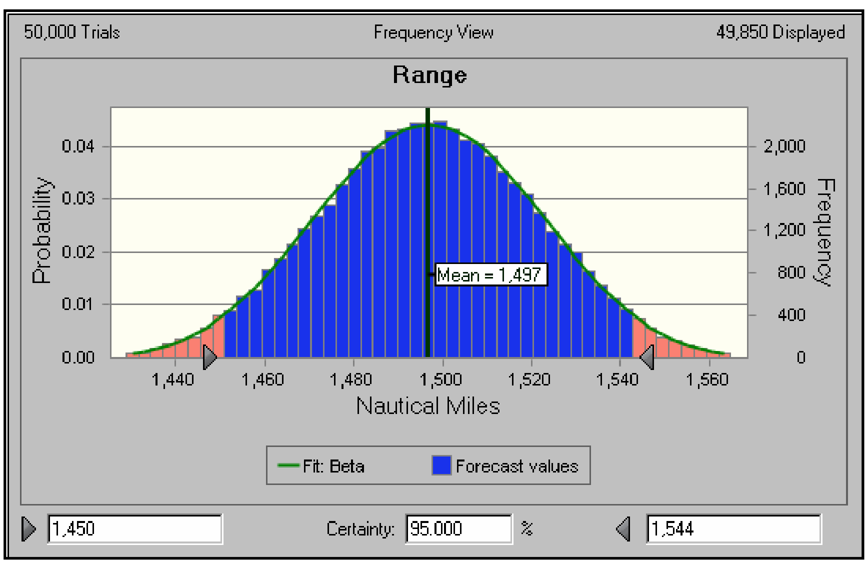
$\alpha_o = 1 - CI$

If the methods of the Insurance and Banking Industry are followed, risk should only be defined as something that poses a financial threat. In that light, the above definition of risk is redefined. For a guarantee such as range, the customer is only concerned if the resulting range is less than the guarantee. There is no financial risk if the range is better than the guarantee. The above equation for risk considers whether performance will be higher or lower than the guarantee. Since financial risk only exists if airplane performance is deficient of the guarantee, risk is redefined as:

$\alpha = (1 - CI)/2$

Crystal Ball conveniently allows the user to input the CI they desire and the upper and lower bounds of the distribution for that CI are displayed. Figure 3 shows the Crystal Ball output for X-03 range with a 95% CI.

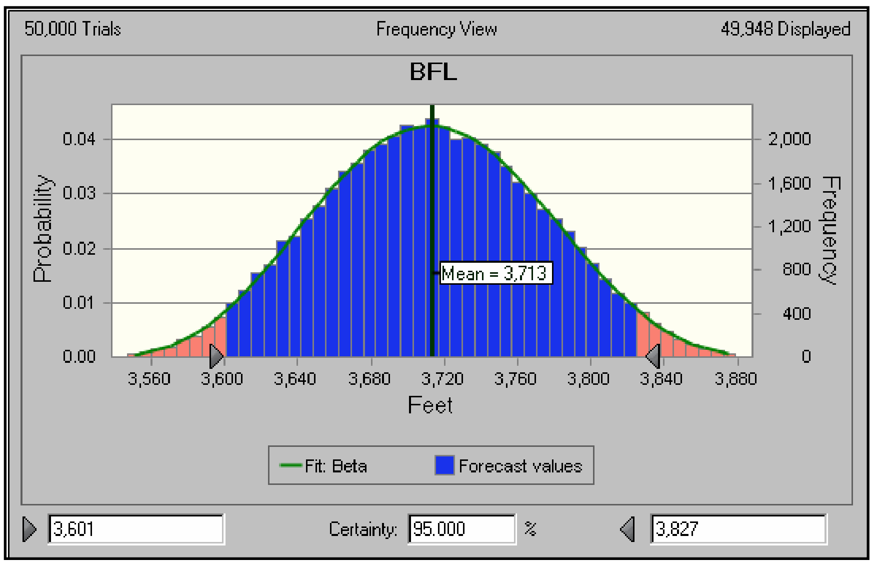
Figure 3: Crystal Ball Output for Range



As seen in Figure 3, with a 95% CI the range varies from 1450 nm to 1544 nm. At the 95% CI there is a 2.5% risk that the production aircraft's range will fall short of 1450 nm.

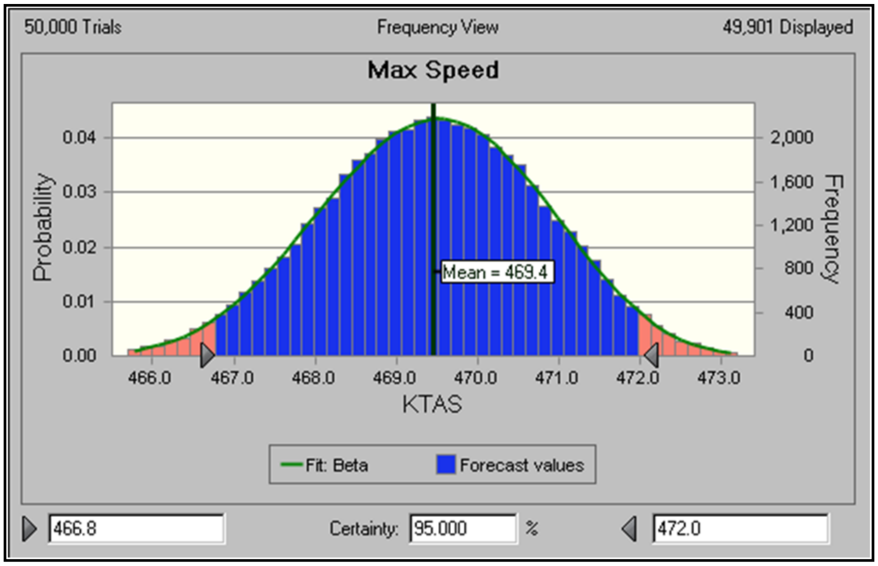
Figure 4 shows the Crystal Ball output for the takeoff field length. The takeoff field length at 95% CI is 3827 ft. Unlike Range, the upper boundary for takeoff field length is used because it is the more restrictive value (a lower value for takeoff field length is better).

Figure 4: Crystal Ball Output for Takeoff Field Length



The Crystal Ball output for maximum speed is shown in Figure 5. With a 95% confidence interval, the maximum speed is 467 kts, ensuring that 97.5% of the expected variation results in speeds faster than 467 kts.

Figure 5: Crystal Ball Output for Maximum Speed



Once Crystal Ball is run and the variation plots are obtained, the user can create a Risk / Guarantee matrix. By adjusting the CI interval input in Crystal Ball, the upper and lower boundaries of the distribution plot shift to incorporate or exclude data. The user records the CI and the boundary that corresponds to the guarantee (lower boundary for range and speed, and the upper boundary for field length). The data is then compiled into a Risk / Guarantee matrix. The Risk / Guarantee matrix for range and takeoff field length is shown in Figures 6 and 7 respectively. The maximum speed guarantee matrix was not included because at 95% CI (2.5% risk), the speed of 467 exceeds the maximum operating speed of X-03. Therefore, it is recognized that the variation required to reduce the maximum speed to less than the maximum operating speed poses no risk and the guarantee can be set at the maximum operating speed limit of the airplane.

Figure 6: Risk / Guarantee Matrix for Range

CI (risk)	Guarantee
95% (2.5%)	1450 nm
90% (5.0%)	1458 nm
85% (7.5%)	1463 nm
80% (10.0%)	1467 nm

Figure 7: Risk / Guarantee Matrix for Takeoff Field Length

CI (risk)	Guarantee
95% (2.5%)	3827 ft
90% (5.0%)	3810 ft
85% (7.5%)	3799 ft
80% (10.0%)	3788 ft

The presentation form, CI (α):η, where η is the performance value associated with a level of risk α, was selected as the standard method for summarizing and presenting the results of the Monte Carlo simulation. The matrices are provided to Management instead of single guarantee values. The concept is that the lower the CI selected, the more optimistic the resulting performance, but the less sure XYZ will be that the guaranteed performance will be met. With a standard risk level defined by the airplane program or potentially the entire company, the performance guarantee is a simple lookup from the matrix. This provides Management with a straightforward process for guarantee selection and also supplies data to support the selection of the guarantee value.

Example Method Outcome

During SRAM's development and first implementation, it was well received by both Engineering and Management. It had a noteworthy impact on the performance guarantees selected for X-03. Upon review of the Guarantee / Risk matrix with Management and Marketing, it was determined that the marketing-desired takeoff field length carried far too much risk to be used as a guarantee. Therefore, based on the risk Management was willing to accept, a new takeoff field length was selected for the published guarantees. It was also determined that the range and maximum speed values desired by Marketing had acceptable risk; therefore, the desired performance was used as the guarantee points. In its first implementation, SRAM proved valuable by providing data that supported two of the desired guarantee points and data that would lead management to select a less risky guarantee for the third. It was viewed as such a valuable method that it has been made the standard method for development of future airplane guarantees.

The SRAM method provides a straightforward method to help evaluate risk in performance guarantees. It also provides the ability to identify how much risk must be accepted if a guarantee be improved by *x* amount. It allows Management to select a guarantee that has an increased level of risk, where the increased risk might be offset by the increased number of sales generated due to the improved guarantee. Ultimately the Risk / Guarantee matrix allows management to select the guarantee that best fits XYZ's goals and risk plan.

SRAM not only lessens the risk of a guarantee, but it also shifts the burden of guarantee selection from Engineering to Management. Arguably, guarantee selection should be Management's responsibility; however, with the past process, the guarantee was a single point resulting from an assessment of nominal inputs. This caused the guarantee to essentially be set by Engineering. By placing the responsibility of guarantee selection on Management, it ensures the selection of guarantees that best suit the company.

SRAM also had another large impact in the final wing configuration of X-03. Two configurations were under consideration, a wing with winglets and a wing without. The marketing-desired range was achievable with both configurations; however, the configuration with winglets had less risk and showed much greater potential of achieving the desired range. Based on this information, a wing with winglets was selected for the final configuration.

The functionality of SRAM outside of performance guarantees is considered a significant strength of the method. As brought to light in the winglet study of X-03, SRAM is a valuable tool for risk assessment of configuration design. For example, during configuration design (well before guarantee selection) the size of the wing, tail, or thrust of the engine is selected to provide the marketing-desired performance of a new product. Much like the guarantees, the performance evaluation of these components is based on nominal expected values, which ultimately results in a design that nominally meets its performance goals. However, when variation is introduced in the components, the nominally ideal configuration may no longer meet the desired goals. By using SRAM to account for variation in components, a more robust configuration can be designed, resulting in a higher quality product.

Chapter 5 – Suggestions for Additional Work

Although SRAM has proven to be a powerful method for risk assessment and is potentially adaptable to many different applications, it is recognized that the process can be streamlined and its usability improved. If CAM were made to integrate with a tool such as Crystal Ball, the need for response surfaces and DOE would be greatly reduced. With these processes removed from SRAM, the process time would diminish significantly along with the initial learning curve.

As mentioned at the end of Chapter 4, SRAM has great potential beyond risk assessment of performance guarantees. It is expected that a large amount of value lies in SRAM's application outside of performance guarantees and configuration design. Further implementation of SRAM should be explored and evaluated.

Beyond SRAM in engineering, there is always the potential to focus on and improve manufacturing processes. This lies outside of SRAM's utilization, but the sensitivity analysis implemented within SRAM can be used to identify areas that pose the largest threat of variation. In general, by improving processes, training, and equipment, the overall variation in a product is reduced. This directly drives the variation of input parameters in SRAM, which results in less performance risk, and thus, less financial risk to the company.

References / Bibliography

Chou, R., R. Fu, et al. (2006). Empirical Evaluation of the Association Between Methodological Shortcomings and Estimates of Adverse Events. U. S. D. o. H. a. H. Services. Portland, OR, Oregon Evidence-based Practice Center.

Koller, G. (2000). Risk Modeling for Determining Value and Decision Making. Boca Raton, FL, CRC Press LLC.

Panjer, H. H. (2006). Operational Risk: Modeling Analytics. Hoboken, NJ, John Wiley & Sons Inc.

Patankar, M. S. and J. C. Taylor (2004). Risk Management and Error Reduction in Aviation Maintenance. Burlington, VT, Ashgate Publishing Company.

Porche, I., H. Willis, et al. (2004). Framework for Quantifying Uncertainty in Electric Ship Design. O. o. N. Research, RAND.

Solutions, S. S. (2008). 6σ Green Belt Training, Six Sigma Solutions.

Stolzer, A. J., C. D. Halford, et al. (2008). Safety Management Systems in Aviation. Burlington, VT, Ashgate Publishing Company.

Tong, M. T. (2002). Risk-Based Probabilistic Approach to Aeropropulsion System Assessment. NASA. Cleveland, OH, Glenn Research Center.

Verrill, S. P. and R. A. Johnson (2007). Confidence Bounds and Hypothesis Tests for Normal Distribution Coefficients of Variation. U. S. D. o. Agriculture. Madison, WI, Forest Products Laboratory.

Vick, S. G. (2002). Degrees of Belief: Subjective Probability and Engineering Judgment, American Society of Civil Engineers.

Wells, A. T. and B. D. Chadbourne (1992). Introduction to Aviation Insurance and Risk Management. Malabar, FL, Krieger Publishing Company.