DETERMINATION OF MISSION SUCCESS PROBABILTY IN AIR FORCE FLYING MISSIONS BY BAYESIAN BELIEF NETWORKS AND OPERATIONAL RISK MATRICES

By

EMMANUEL BRIAN CAO

THESIS PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTERS OF SCIENCE

UNIVERSITY OF FLORIDA

2011

I would like to dedicate this thesis to Dr. Henry Pfister, who passed while I was a student at UF. Not only did his teachings about Bayesian Belief networks inspire the core of my thesis, but he also provided me encouragement and validation to pursue System Engineering studies.

ACKNOWLEDGEMENTS

I would like to thank Dr. Vladimir Boginski for his efforts in guiding me during my work for this thesis, as well as Judi Shivers who assisted me throughout my time at the UF REEF. I owe many thanks to my friends who guided me in my academic studies, particularly my friends in the 36th Electronic Warfare Squadron at Eglin AFB, without whom I would have never began my master's studies at UF. I would also like to show my gratitude to my family – my father Khai, my mother Bernadette, and my sister Alice- for their continual love and support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Abstract of the Thesis Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Master of Science

DETERMINATION OF MISSION SUCCESS PROBABILTY IN AIR FORCE FLYING MISSIONS BY BAYESIAN BELIEF NETWORKS AND OPERATIONAL RISK MATRICES

By

Emmanuel Brian Cao

December 2011

Chair: Vladimir Boginski
Major: Industrial and Systems Engineering

Continually changing technologies, especially in the field of software, allow us to enhance the manner of business and increase efficiency while decreasing risk. This the nature of Systems Engineering and in this thesis I will use two core programs, Netica™ and Design Expert®, to identify calculable risk for flying operations in an Air Force squadron. The current standard of risk identification in United States Air Force flying squadrons is with ORM, or Operational Risk Management, and essentially comes in the form of a worksheet where users identify the risky elements on a mission and produce an overall risk rating. I will use this tool in this thesis as the primary source of data to estimate a Mission Success Probability (MSP), a potentially invaluable deliverable to military leaders. If successful, the user will have a specific gauge as to how his mission will unfold with supporting data to make significant decisions, such as continuing or halting a mission with specific risks present. In this thesis, I will discuss the use of Bayesian mathematics, the basis for the Netica™ software that I used, and its utility to military decision making specifically in the realm of aviation. Like many Systems Engineering applications, the purpose of this idea is that it improves on existing procedures and technologies to deliver more specific results. It draws upon multiple sources to create automated capabilities

that were nonexistent before. Instead of only having an indicator of when ORM levels are "Low, Medium, High, or Extreme," one can now predict or have exact probabilities based on historical data. Based on this Mission Success Probability, or MSP, users can now become precisely aware of how certain risks affect their mission resulting in safer and more effective missions.

Upon completion of this case study, I validated the use of the Mission Success Probability calculator in Netica™ by measuring strong correlation effects in Design Expert® between predicted MSPs and actual Overall Mission Effectiveness (OME) percentages. In addition to quantifying the magnitude actual risks that occur during a typical flying mission, I confirmed multiple advantages my system has over the current risk assessment matrix.

CHAPTER 1
OPERATIONAL RISK MANAGEMENT (ORM)

1-1 ORM Background

Operational Risk Management is a widely used, regulatory enforced process used by all branches of the Armed Forces to specifically identify risks of a certain operation and to mitigate such risks. This is usually accomplished before an operation as the name implies, such as a combat mission or operational test. In the Air Force specifically, ORM is regulated by Air Force Instruction 90-902 and 90-9, Operational Risk Management. According to this AFI, ORM is defined as “a decision-making process to systematically evaluate possible courses of action, identify risks and benefits, and determine the best course of action for any given situation” [2].

The pertinent principles to this thesis are to “tailor ORM principles to meet the specific mission needs and operational requirements of each organization” and “that it provides the process and tools to understanding at risk behavior on both on and off duty”[2]. In other terms, the Air Force defines ORM as an all-inclusive effort that occurs before, during, and after any operation and functions to accomplish the mission in an effective and predictable manner, while keeping safety paramount. Air Force leaders must be able to achieve consistent and standardized results from the components they command [8]. The intended effect is for the ORM users to produce ways to mitigate or control the identified risk levels. These measures can be as simple as a discussion about safety awareness before an extended weekend, or complex as a worksheet detailing all the risky aspects that may adversely affect a flying mission. Air Force ORM is divided into 3 levels [10]:

- BASIC: Everyday, simple risk management measures such as wearing a safety belt

- OPERATIONAL: Job function related risk management procedures such as cancelling a risky mission.
- STRATEGIC: Staff or organization-level decision making managing risk, such as employing a safer tactic during a military campaign. An example of this occurred during Operation Storm when Air Force strike packages changed their attack profiles to a higher altitude due to unnecessary losses on low-level attack profiles.

1-2 ORM Utility to Air Force Flying Units

For the purposes of this thesis, we will mainly be dealing the Operational and Strategic level of ORM as it applies to flying units in the U. S. Air Force. All Air Force aviators are trained to counter operational risk with prescribed solutions which h ave proven effective. Figure 1-1 [15] is the 6 Step ORM Process by which all Air Force personnel are mandated to follow. These mitigation methods can often reduce risk and in many cases totally negate the risk.

For many flying squadrons, this process is manifested in a traditional ORM worksheet which captures all the possible risks on a mission and presents a score for the purposes of awareness. When the traditional ORM sheet is filled out, the user makes his or her inputs based on experience and garnered knowledge from fellow aviators or instructions. The user is always an experienced aviator, with at least 6 years of aviation and approximately 9-10 years being the mean [4]. This equates to at least 1,000 hours of flying experience per user and a worksheet that is often completed with a high degree of consistency. [1] Figure 1-2 is a snapshot of an actual ORM Matrix filled out for risks pertaining to “Terrain”.

There is much room for subjectivity in these matrices as flying experiences differ from user to user and the opinions or feelings concerning the risk of a mission may change accordingly, causing non-standardized results. For example, an aviator who had a near fatal experience with bird collisions in his or her flying history may consistently heighten the risk level of the BASH (Bird Aircraft Strike Hazard) as opposed to a user that has never had a bird incident and consistently gives a “Low” BASH rating. This is an inexact science, as aircrew members subjectively quantify the level of risk for a mission based only 4 levels of rating (“Low” to “Extreme”) and, although they have a higher awareness of risk factors, the final risk level determination is judged based on opinion and not factual evidence, leaving much to be desired.

In addition, a squadron may choose to set a threshold of say, “Extreme” levels of risk before it requires the flying crew to notify or ask permission from leadership to proceed with the mission. In terms of ORM deliverables, the process stops after the aircrew briefing but the principles are still practiced during the mission. In other words, once the aircrew determines the risk level, they discuss it as a group, weigh it against alternatives and mitigating factors, and then remain vigilant for risk-related factors throughout the mission.

At the very least, aircrews will be aware of the inherent risks of the mission and will be more vigilant than before. If this level were “High” or “Extreme”, leadership or higher levels in the military chain of command would be notified in order to make this determination for the crew. ORM Regulatory guidance prohibits acceptance of unnecessary risk, although in wartime conditions the tolerances may differ [13]. Colonel James Stanley from Headquarters Air Combat Command provides an excellent example of flying ORM from his excerpt as published in the *Combat Edge* magazine:

The ACC staff was reviewing a proposal for F-16s to pull some of our Iceland alert commitment. Historically, this alert had been tasked to F-4... and now F-15... aircraft, both of which are two engine aircraft. Now a proposal to use F-16s was up for consideration. During review of the proposal, the following issues were identified and researched: (1) the F-16 has only one engine, (2) prevailing weather and crosswind were valid concerns, and (3) alternate airfields would be difficult to reach due to the range of the F-16. It was the classic opportunity for ORM to work its magic! The hazards were identified as weather, range, and single engine operations. The risk was assessed for the F-16 to perform this mission in the demanding Iceland environment. Analysis revealed that no risk control measures were available or realistic. The decision was made not to task the F-16. [10]

1-3 Advantages and Disadvantages of Current ORM Methods

The data points used in this thesis will be recovered from ORM worksheets used by the MC-130P aircrews. The MC-130P aircrews are heavily tasked and combat tested aviators, and there are two main types of ORM worksheet used within this aircrew community on their training missions to satisfy the ORM requirements. One is the "9th SOS Operational Risk Management Form" (Appendix A-1) and the other is the "130th RQS Home Station Risk Assessment Form" (Appendix A-2). Both contain their share of advantages and disadvantages.

The "9th SOS Operational Risk Management Form" is divided into 5 subsections that identify major areas of risk to flying mission success, which are "Mission, Enemy, Environment, Troops, and Time". We see in this process that risk is identified by placing an "X" in the subsection that is affected, For example, if multiple crew members were experiencing personal issues before a flight, they would annotate this under "Human Factors" selection of the "Troops" section. The Aircraft Commander would then assess the severity of the risky factor and determine whether the subsection should be rated "Low, Medium, High, or Extreme". The overall risk assessment is then made based on the summation of the subsections and like the subsections is given an overall "Low", "Medium", "High", or "Extreme" rating. If the ORM level were elevated, the crew would make a determination to decide if the benefits of the mission outweigh the risks. For example, if a training mission was rated "High", the Aircraft Commander may elect to continue based on the number of training events he can get done. However, this may not be true in a combat mission, where a “High” rating may be too risky based on the increased number of threats and decreased amount of mitigating factors.

There are 3 main thresholds for the overall risk assessment. The first one requires Assistant Director of Operations (ADO) or Director of Operations (DO) approval if the mission is assessed as “Medium”, the second threshold requires Squadron Commander (in charge of multiple units of a single aircraft platform) approval if the risk assessment is an overall “High” and the third requires Group Commander (a leader in charge of multiple units of different aircraft platforms) approval if the risk assessment is “Extreme”. This is illustrated in Figure 1-3.

A positive aspect of this system is that it allows user oversight in instances where simply adding the risk scores would be inaccurate. An example would be a mission where all of the factors in the “Mission” subsection are rated as “Medium”, but the user may have and give the overall subsection a “High” rating if he feels the combined magnitude of the factors has a greater effect on an inexperienced aircrew. This would address shortfalls in standardized computations that do not recognize the intangible factors. Before the final risk assessment is made, the top risks of the mission are identified, as well as the means to mitigate them. For example, in the worksheet in figure, “Troops” are rated as medium because of the lack of proficiency in the crew members; however there are instructors on board to provide supervision. This, along with other mitigating factors help to lower the overall mission risk to “Low”.

However, this subjective system of scalar rating has a good number of drawbacks as well. (1) As previously stated, the overall results based on the same individual subjections and factors may vary from user to user and variance is undesirable in military systems which are constructed to have as much predictability as possible. (2) The ratings are not exact as the final ratings are measured on a scale. There is no exact science to this figure, and when there are means of delivering exact probabilities we should strive to achieve them.

The other method is the additive method, which is used for the "130th RQS Home Station Risk Assessment Form." The form is categorized in the same manner as with the difference being that each individual factor is given a numerical score with different factors given weights depending on their significance, and then collectively added to provide a final score. The final score is then used to determine a risk level and similar to the form the different score thresholds correlate to different approval levels. These are the main disadvantages of this system lies in the additive computation of the risk score because it disregards the conditional nature of many of the variables. In this worksheet, all of the variables are given a standard weight on a scoring scale. For example, "Illumination" is given 1 point and "12-14 Hour Crew Day" is given 2 points. In other words, this means that "12-14 Hour Crew Day", at its most extreme effect, holds twice or 100% more influence than "Illumination" which is not true in most cases. This relationship is not substantiated from any historical data but more of a subjective rating.

This leaves room to be desired for a modeling system that takes into account not only the individual factors' effect but also its effect on competing factors. In addition, there is a need for the standardization of factor weights, whereas in previous instances, users subjectively determined weights to assign to different sections and decided upon the risk level on their own. The flaw here is the standard of risk may change from one user to the next and may deliver inexact results. Therefore, this standardization should be based on actual data rather than individual opinion, with user oversight.

Figure 1-1. Adapted from "The 6 Step ORM Process" [15]

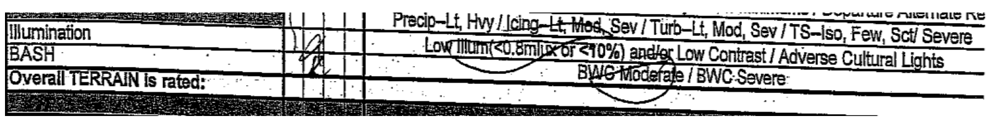


Figure 1-2. ORM Matrix Excerpt

Figure 1-3. ORM Approval Thresholds

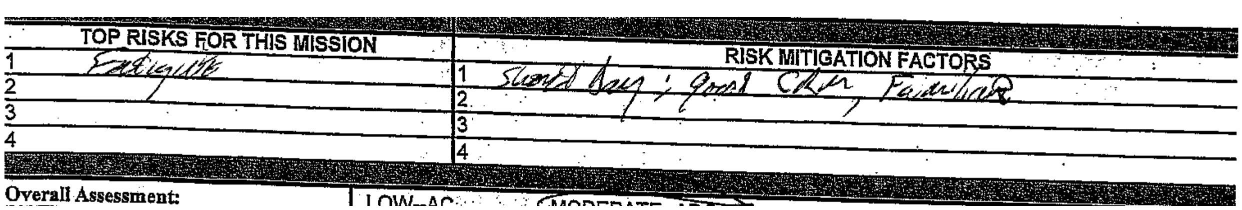


Figure 1-4. Mitigation Examples

CHAPTER 2
BAYESIAN BELIEF MODELS

2-1 A Bayesian Principles

A Bayesian Modeling of the ORM worksheet addresses all of these drawbacks addressed in the previous section and retains the benefits. In order to model relationships from concurrent events with probabilities based on historical data, Bayesian Belief Networks are the best tool to produce within these constraints. Bayesian Belief Networks were first introduced to me in Dr. Henry Pfister's Systems Architecture class (ESI 6552). Upon completion of sample model network exercises and research on the Netica™ tool, I saw the utility of a belief network for a Systems Engineer. In addition to modeling the components of a system in a graphical format, Bayesian Belief Networks proved the following tools for a Systems Engineer [14]. They are (1) the definition of variables, (2) the definition of causal relationships, (3) the identification of significant factors, (4) the evaluation of effect on other nodes, and (5) it incorporates historical data.

At the core of modeling algorithms is the conditional nature of the Bayesian calculation, which computes the probabilities of the multiple related events. The law of conditional probability, which is the basis for Bayesian software programs, is the work Rev. Thomas Bayes and more specifically his discovery in the 1700s of a basic law of probability [6]. This is expressed by

$$P(b|a) = \frac{P(a|b) \times P(b)}{P(a)}$$

where P(a) is the probability of even a happening and (a|b) is the probability of a occurring given that event b has occurred. For example, say that 50% of people in car accidents are seriously hurt. Also, 1 in 50,000 people get in accidents and 1 in 20 people have been seriously hurt. What is the probability that a person with a serious injury was involved in a car accident?

$$P(\text{Accident}|\text{Seriously Hurt Person}) = \frac{P(\text{Seriously Hurt Person}|\text{ Accident}) \text{ x } P(\text{Accidents})}{P(\text{Seriously Hurt Person})}$$

$$= \frac{0.5 \text{ x } (1/50{,}000)}{(1/20)}$$

$$=.0002$$

Meaning if someone has a serious injury, it had a .02% of being from a car accident. Another way to express this equation that better fits the context of this thesis is

$$P(H|E,\, c) \;=\; \frac{P(H|c) \times P(E|H,c)}{P(E|c)}$$

where the hypothesis H is updated based on the additional evidence E and past experience. This is called posterior probability expressed by P(H|E,c). The term (P(H|c) is the probability of H given c alone. P(E|H,c) is the probability of evidence assuming the hypothesis H and historical information is true. P(E|c) is the probability of evidence given c is true. This term is independent of H and can be regarded as a normalizing or scaling factor. This rule will be instrumental for the purposes of this thesis, as it will be used to compute multi-variable relationships in an efficient and accurate manner which would have been a long and arduous process if done by hand [19].

2-2 Bayesian Belief Networks

A Bayesian Belief Network models causes and effects in a particular system and delivers probabilities of system success based on the probabilities of its components. At the basis of the relationships between component probabilities and system calculations is Bayesian Mathematics as explained in the previous section. These types of networks are particularly useful in situations where all data is unavailable or decisions must be made with uncertainty. A Bayesian Belief Network graphically models the probabilistic relationships and presents the results in an intuitive manner [19]. We can see in the following sample diagram that each of the variables in the system are represented by a node, which internally contains the different states.

In Figure 2-1 the top nodes may just be dual states, where either the sky is sunny or cloudy. Raining is dependent on the Cloudy node, the grass can be wet or dry, and the sprinkler can be on or off. The nodes are connected with the arrows pointing in the line of influence, and in this example they are used to determine whether there will be wet grass or not. Each state inside of a node has conditional probabilities, which are determined by the parent nodes that connect to it. In this example, the model can be used to determine a case such as the lawn is wet, and if it was more likely to be caused by the rain or the sprinkler. It may be a result of the cloudiness which causes rain, or the sun, which may cause the homeowner to turn on his sprinkler. Historical information about the relationships between the nodes is used to determine whether the node is in one state or another are contained within each node in tables called Conditional Probability Tables (CPTs). These tables represent the probabilistic relationships between the states of the connected nodes. For this example, they would be used to answer questions like:

- If the lawn is wet, is more likely from the rain or sprinkler?
- How likely is it that I will water my lawn on a cloudy day?

2-3 Bayesian Application to ORM Modeling

This is the core of my thesis, as the goal is to make measurements about a system based on many sub factors which have varying degrees of effect based on other individual factors. For example, in the subsection labeled "Troops", "Inflight Fatigue" may be amplified if preflight fatigue is present, meaning that if someone is tired before a mission they will be exponentially more fatigued during the mission due to the increased number of external physical hindrances (i.e. reduction in oxygen, increased temperature). As we can see the current rating individualizes each category, for example, the "Inflight Fatigue" is rated solely based on aircraft temperatures and not in "Preflight Fatigue" or many other factors such as "Illumination (dimmer settings increase physical strain on night time visual flying). This current system places the responsibility on the user to draw relationships between the factors and accurately rate them.

A Bayesian Model would alleviate this responsibility, as long as the model correctly draws the relationships and uses an accurate weighting system. Using the previous example in the "Troops" category, we can enter different arguments to illustrate this example as seen in Figure 2-2. For "Preflight Fatigue", say the risk is increased from "Low" to "Medium". This, in turn, will decrease the likelihood of the "Inflight Fatigue" subcategory becoming a factor from 90% to 70% "Low" risk and will increase the overall "Troops" factor from 8.5% to 15.4%. Unlike the scalar and additive models, the Bayesian model incorporates all of the individual factors together instead of separately, which is a more accurate reflection of their relationships in actuality.

2-4 Netica™

Netica™ is the software program used to model the Bayesian Belief Networks. It functions in the same format as other probabilistic modeling software in that it graphically depicts the belief networks into junction trees for fast probabilistic reasoning. It functions by capturing the relationships among a set of joint probabilistic variables and displaying them in terms of nodes (the variables) and links. The links are causal uncertain influences that determine posterior probabilities based on prior distributions by using Bayesian Belief updates from the new findings. The Netica™ algorithm has the following characteristics.

- Assumes conditional probabilities as independent
- Prior distributions are dirichlet functions
- Requires a large number of cases
- Network starts in state of ignorance
- Nodes where case has a value and parental values begin with conditional probabilities if modified
- Learning is sequential and not iterative
- Automatic process that allows for missing values

In addition, it has a feature that allows the user to define the probabilistic relationships by equations. It can factor time delays, updating disconnected links and has facilities for easy discretization of continuous networks [14].

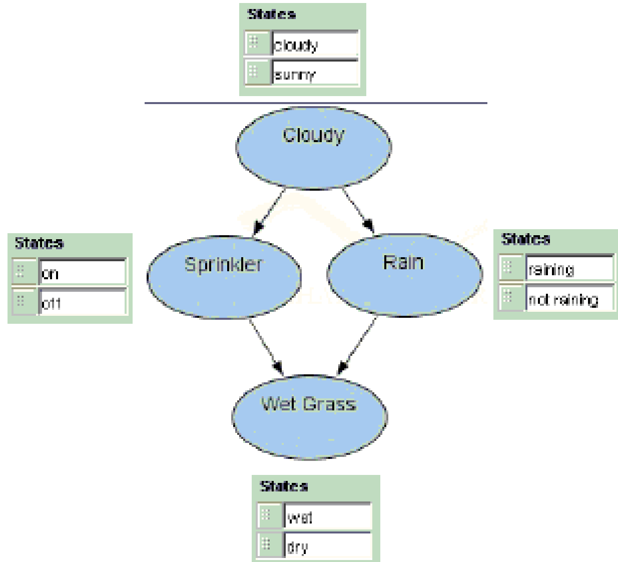


Figure 2-1. Sample Bayesian Belief Network

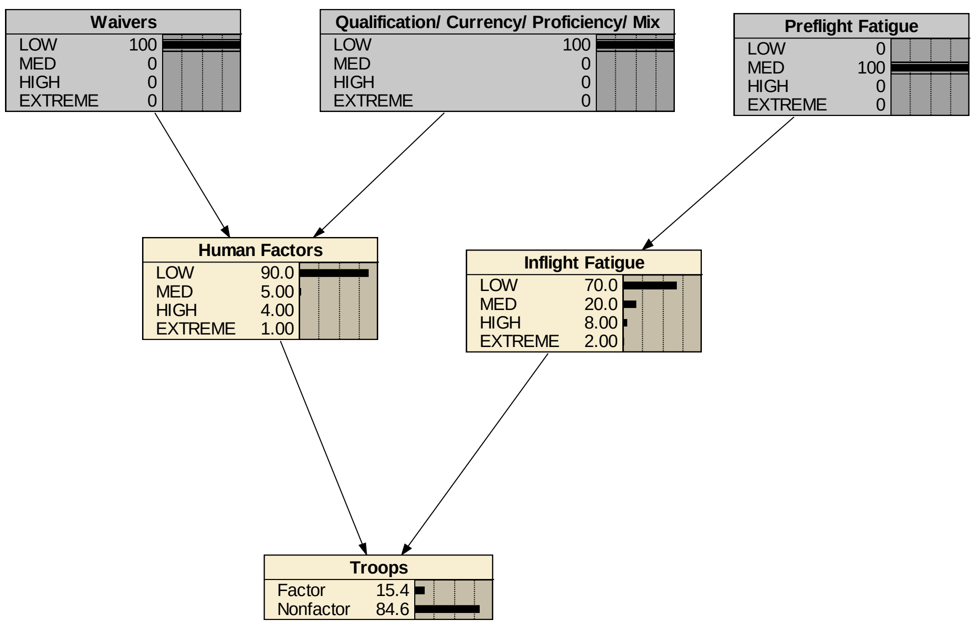


Figure 2-2. Sample Troops Belief Network

CHAPTER 3
CONVERSION OF ORM WORKSHEETS INTO MSP CALCULATOR

3-1 The Approach

As stated before, this thesis proposes to deliver a Mission Success Probability (MSP) level based on data from the ORM worksheets and calculated through Netica™. Combining these two tools, Netica™ will complement the vagueness and uncertainty the ORM worksheets presents. Netica™ will precisely gauge the odds of mission success based on historical data and will identify the main sources of risk. In order to determine the exact weights of all the factors used in the final MSP calculation, I will enlist the use of Design Expert® (DE), a Design of Experiments (DOE) software tool used measure statistical information about a test. The entering arguments to design a model in DE are the enumerated ratings in the ORM worksheet as well as the Overall Mission Effectiveness for the same mission, found in the Air Force Special Operations Command Form 40 (Figure A-3). The AFSOC Form 40 is essentially a mission summary that describes the final outcome of training missions and most importantly, it provides a final Overall Mission Effectiveness, or OME, at the bottom. This number quantifies the portion of the original goals that was actually accomplished during the mission and must be completed at the end of the mission by the same user that filled out the ORM worksheet at the beginning of the mission. The data set to design the model will be taken from 31 MC-130P training missions from Eglin AFB, FL over the period from May 2010 to June 2010 (Table B-1), and the data set to test the hypothesis will be taken from 30 flights from July 2010 to August 2010 by the same unit (Table B-2). This comparison will hopefully yield a successful correlation, and the end product will be a validated electronic version of the ORM worksheets, where a user would input the significant sources of risk in a mission and the final deliverable would the probability of mission success. Ideally, this number would be the used to justify the validity of the mission and a threshold would be set if there would be a need for further approval.

Once this final probability is calculated, I will perform an Analysis of Variance (ANOVA) using the Design Expert® software on the results and will be able to decipher the most significant factors affecting this score. In addition, the ANOVA will also give information about correlation factors and the accuracy of my hypothesis.

3-2 Design Expert®

Design Expert® is DOE based software that designs the structure of an experiment with an exponential amount of data points, and interprets the data in a variety of user friendly formats. D.E. was introduced to me the Design of Experiments class as a very valuable tool in designing lengthy experiments with multiple factors and testing for statistical significance of model factors.

Design Expert® Deliverables

Other than producing standard statistical data about an experiment, such as Sum of Squares and Degrees of Freedom, Design Expert® gives an in-depth analysis of process factors and mixture components, and also identifies critical factors and optimal conditions in a trial. This information is delivered in a variety of visuals, the most important one being the Analysis of Variance (ANOVA) Table [18]. ANOVA tables display the following key statistics about the correlation factor of a model:

- F VALUE. Test for comparing term variance with residual (error) variance. If the variances are close to the same, the ratio will be close to one and it is less likely that the term has a significant effect on the response. It is calculated by term Mean Square divided by Residual Mean Square.

- PROB > F. Probability of seeing the observed F value if the null hypothesis is true (there is no factor effect). Small probability values call for rejection of the null hypothesis. The probability equals the proportion of the area under the curve of the F-distribution that lies beyond the observed F value. The F distribution itself is determined by the degrees of freedom associated with the variances being compared. In plain English, if the Prob>F value is very small (less than 0.05) then the individual terms in the model have a significant effect on the response.)

- MEAN SQUARE. Estimation of model variance, calculated by the term sum of squares divided by term degrees of freedom

- ADJUSTED R-SQUARED**.** A measure of the amount of variation around the mean explained by the model, adjusted for the number of terms in the model. The adjusted R-squared decreases as the number of terms in the model increases if those additional terms don't add value to the model [18].

$$\text{Adj } R^2 = 1 - \left[\left(\frac{SS_{residual}}{df_{residual}}\right) \Big/ \left(\frac{SS_{residual} + SS_{model}}{df_{residual} + df_{model}}\right)\right] = 1 - \left[\left(\frac{SS_{residual}}{df_{residual}}\right) \Big/ \left(\frac{SS_{total} - SS_{curvature} - SS_{block}}{df_{total} - df_{curvature} - df_{block}}\right)\right]$$

- PREDICTED R-SQUARED. A measure of the amount of variation in new data explained by the model.

$$\text{Pred } R^2 = 1 - \left[\frac{PRESS}{SS_{residual} + SS_{model}}\right] = 1 - \left[\frac{PRESS}{SS_{total} - SS_{curvature} - SS_{block}}\right]$$

The predicted R-squared and the adjusted R-squared should be within 0.20 of each other. Otherwise there may be a problem with either the data or the model. Look for outliers, consider transformations, or consider a different order polynomial.

- ADEQUATE PRECISION**.** This is a signal-to-noise ratio. It compares the range of the predicted values at the design points to the average prediction error. Ratios greater than 4 indicate adequate model discrimination [18].

$$\left[\frac{\max(\hat{Y}) - \min(\hat{Y})}{\sqrt{\bar{V}(\hat{Y})}}\right] > 4 \quad \bar{V}(\hat{Y}) = \frac{1}{n}\sum_{i=1}^{n} V(\hat{Y}) = \frac{p\sigma^2}{n}$$

p = number of model parameters (including intercept (b 0) and any block coefficients)

σ^2= residual MS from ANOVA table

n = number of experiments [18]

These are key statistics in determining the correlation factor of a system, and I will primarily use them to determine the significance of model factors in this thesis as well as the accuracy of our calculated MSP.

Design Expert® Model Types

In addition to displaying all of the statistical information graphically in 2 and 3 D formats, D.E. also allows the user to adjust the factor settings to attain the optimal output. The following are the four types of design D.E. can model our test:

- FACTORIAL DESIGNS: Identifies vital factors affecting test; determines number of data points required for an accurate analysis but cannot analyze a prior test with limited data
- RESPONSE SURFACE MODEL (RSM): helps quantify the relationship between one or more measured responses and the vital input factors. This also allows for a historical model, meaning that it can be used for an accomplished test with a limited amount of data points
- MIXTURE DESIGN TECHNIQUES: Offers different configurations for optimal formulation
- COMBINED DESIGNS: Combines process variables, mixture components, and categorical factors in one design

The model that best fits our thesis design is the RSM Historical Model, because it allows for the use of data points from previously accomplished tests, while allowing up to 50 factors and assigning a numeric rating for each factor. In this thesis design, I scored a “Low” rating as 1, a “Medium” rating as 2, a “High” rating as 3, and an “Extreme” rating as 4. This was repeated for all 34 factors. This is illustrated in Table B-1, where the entire ORM worksheet in converted numerically on an Excel® spreadsheet.

On the first trial run, there were errors with the entire design as the ANOVA table returned a p-value of meaning that no factors were statistically significant. In addition, the model had a Predicted R-squared value of -0.0851. As stated above, this negative value implies that the overall mean was a better predictor of the response from the current model. This made sense, because in retrospect, there were many factors (34) without much variability in their individual ranges (1 and 2 scores only) to make any serious statistical difference. Simply put, the high level of variability with this model caused it to be ineffective. Therefore, I adjusted the model eliminating excess parent nodes and keeping only the key child nodes – "Mission", "Troops", and "Time" which directly determined the final MSP. In addition, for reasons of statistical significance which will be addressed later in the thesis, I added another factor "Maintenance" to account for mission cancellation due to the aircraft status. Now with only 4 factors, the ANOVA table returned much more viable results as seen below. The p-value was <.0001, meaning that the model terms were significant, and in this case, they were A("Mission) , C("Troops"), and D("Maintenance"). The F value was 157.01 which is strong and implies the model is significant. Finally, the Adequate Precision measurement was 31.424, which exceeds the standard of 4 for a good signal to noise ratio. All of these measurements confirm that the model we used followed a normal distribution with acceptable levels of variance, enough to make determinations about the model terms.

Figure 3-1 is an interaction graph of the significant factors in the model, with 12 of the data points located in the bottom left of the cube, corresponding to the "Low" setting (1) for three factors as well as a 100 % mean MSP. In regression analysis seen in Figure 3-1 below, Design Expert® computed the following regression equation to predict our model in coded terms and identifying the relative significance of the factors by comparing the factor coefficients.

I used these coefficients to determine each of the subcategories (Mission, Terrain, Troops, and Maintenance) relative impact on the MSP and produced the following Conditional Probability Table (Figure 3-2) to calculate the MSP:

3-3 Enumerating ORM Worksheets

This is the breakdown of the "9 SOS ORM Worksheet" and their relationships with other variables, if any.

A. Mission

This portion represents the risk presented by the mission profile and these are the individual factors explained:

- MISSION PRIORITY. The necessity of the mission based on the number of units supported and the command level from which the order originated
- DECONFLICTION. Connotes the number of aircraft that are a part of the same mission. This adds to the complexity if it causes a congested airspace.
- COMPLEXITY. The level of risk is increased due to the addition of more mission events, including Helicopter Air Refueling (HAR), airdrops, or nighttime low level navigation. Affected by Mission Priority and Deconfliction
- FAMILIARITY. Addresses the frequency of the mission events and many users are accustomed to the profile.
- PROFILE. Highlights the any extraordinary events that do not normally occur on a regular mission, with higher scores given to planned, complicated maneuvers such as High Altitude Airdrops. Affected by Deconfliction and Familiarity
- SUPPORTED FORCES. Highlights any risk that may arise from unfamiliarity with other units participating in the same mission.
- ENEMY. Another source of risk to mission completion is the presence of the enemy. Enemy is the opposing force to mission completion, and seeks to increase the factor of risk as much as possible. It is a major consideration in the MC-130P mission, and the mission profile is often shaped around enemy tactics. Because of

this, I have included it the “Mission” category. The factors pertaining to the “Enemy” include

- INTELLIGENCE. The strength of our intelligence and our own ability to locate the enemy and our knowledge of effective tactics to defeat the enemy

- PROBABILITY OF DETECTION. Enemy’s chance of detecting the user. Affected by the aircrew “Intelligence” factor

- PROBABILITY OF ENGAGEMENT. Enemy’s probability of attacking, especially in areas of high vulnerability. Affected by the Probability of detection.

- PROBABILITY OF DEFEATING THE THREAT. The user’s ability to counteract the threat and make it a non-factor. Affected by the Probability of engagement

B. Terrain

Terrain denotes the external factors that the user generally does not have any control over. This category divides the mission temporally by the order in which the events occur.

- DEGRADED AIRCRAFT. The condition of the aircraft

- START/TAXI/TAKEOFF. Highlights any possible issues during the very start of the mission, such as taking off with a heavy aircraft in a short runway. Affected by Degraded Aircraft

- ENROUTE. Any mechanical risks encountered on the high level portions of the flight, such as the inability to pressurize the aircraft. Affected by Performance

- OBJECTIVE. Any risks present in the area of focus for the mission. Affected by Enroute

- RECOVERY/DIVERT. Any mechanical hazards anticipated in the terminal portion of the mission, such as a gear malfunction. Affected by Objective

- WAIVERS. Risk inherent in higher level approval for non-standard practices

- PERFORMANCE. The level of aircraft reliability. This is dependent on condition in which Maintenance troops can maintain the Aircraft (FMC-Fully Mission Capable, PMC- Partially Mission Capable, NMC-Not Mission Capable). Affected by Waivers

- ILLUMINATION. Level of brightness. At night, this plays a role in Low Level Navigation and Terrain Following.

- WEATHER. Any risk based on forecasted adverse weather. Affected by Illumination

- BASH- Bird Aircraft Strike Hazard –amount of birds in the air

C. Troops

This is the most controllable category of all. It pertains to the factors concerning the human conditions of the user. The aircrew is usually a mix of 8 personnel comprised of 5 positions, so there a variety of personal issues that may affect the mission. This category has a significance factor of 28% from the regression model, and once again it contains the "Time" subcategory as well.

- PREFLIGHT FATIGUE. Rates the users physical state and rest quality

- INFLIGHT FATIGUE. Rates the environmental conditions that may wear on the user. Affected by Preflight Fatigue

- WAIVERS. Again, any risk inherent to higher level approval for non-standard procedures pertaining to physical conditions (waiving authorized rest time to launch a crew early)

- HUMAN FACTORS. Risk from personal distractions. Affected by Waivers

- TIME: This is another manageable factor as the user controls the timeline for the mission and the majority of the time. In the case that real world conditions call for a compressed timeline (i.e. a priority alert mission or consecutive flying schedules), this may be a factor, as limited time available results in a greater chance of routine flying procedures being violated. Compressed timelines, although not always causal, have been usually found to be influencing factors in major aircraft accidents, as evidenced by the SOHO Spacecraft Accident [7]. Because of this, I included Time as a subcategory under the "Troops" category. Execution and Preparation are parent nodes.

- PLANNING. Time available in the planning process. The greater the amount of planning time, the higher probability for mission success

- PREPARATION: The level of user preparedness prior to mission execution. Affected by Planning

- EXECUTION: The balance between the time available and the amount of mission

3-4 Use of Netica™ to build ORM Model

These factors listed above were modeled in the same way they are presented in the ORM worksheets, with the individual factors as parent nodes for the individual category they belong to, now listed as child nodes. In some cases, these individual factors are parent nodes for other individual factors as described in Section. For example, “Supported Troops” and “Complexity” are parent nodes for “Profile”, which is in turn a parent node for the “Mission” subcategory node. Finally, these Categorical child nodes become the parent nodes for the Mission Success Probability (MSP) child node. The MSP node contains the conditional probability table from figure, which is used in calculating the MSP. The initial Netica™ model is Figure A-3 and the final Netica™ model can be seen in Figure A-4.

3-5 Data Collection

The experiment began as previously stated by transferring the prescribed settings from the ORM worksheets onto the Netica™ belief network in the parent nodes. After a few trials, it was necessary to make a standardized rule concerning the nodes in order to maximize the use of the Bayesian Calculations

- All Parent Nodes should be filled out first
- Child nodes should be filled out only if there is a rating for that node other than “Low”. Otherwise, the node calculation should be decided by the node’s Conditional Probability Table. This allows for the user to see the effect of the parent node on the child node, with the option to override the belief network in areas the user sees fit.

This addresses a drawback of the network that was addressed earlier in Section "1-3 Advantages and Disadvantages of Current ORM Methods"

- Categorical Nodes and, obviously, the Mission Success Probability Node, should not be changed and the Bayesian calculation should be used.

It also became obvious in the trial runs that there needed to be a change in the Netica™ model with respect to the weighting scale given to the maintenance factor, which includes the "Degraded Aircraft" and "Start/Taxi/Takeoff" nodes. Maintenance errors account for 25.8% of mission failure, and in a calculation where the maintenance factor is normalized along with the "Terrain" subsection; the resulting mission success percentage only shows a difference of 2.2% at the most extreme levels. The most significant source of error occurred in situations where there were high and extreme levels of risk in the maintenance subsection. This has such a poor correlation factor that Design Expert® determines that there are no significant terms in this model and that the overall mean is a better determinant of the score. In other words, the error caused by the missions due to maintenance is large enough to render the Netica™ model statistically useless. In order to correct for this, I separated the maintenance subsection from the rest of the "Terrain" subsection which returned much improved results. Instead of having a statistically insignificant model, the correlation factor now had an F-Test value of 157.01 which shows that there is strong correlation and that we can use this Netica™ design to model risk. Therefore, we will use 4 overall categories (Troops, Maintenance, Terrain, and Mission) to determine mission success as opposed to the 5 categories presented in the earlier part of the thesis.

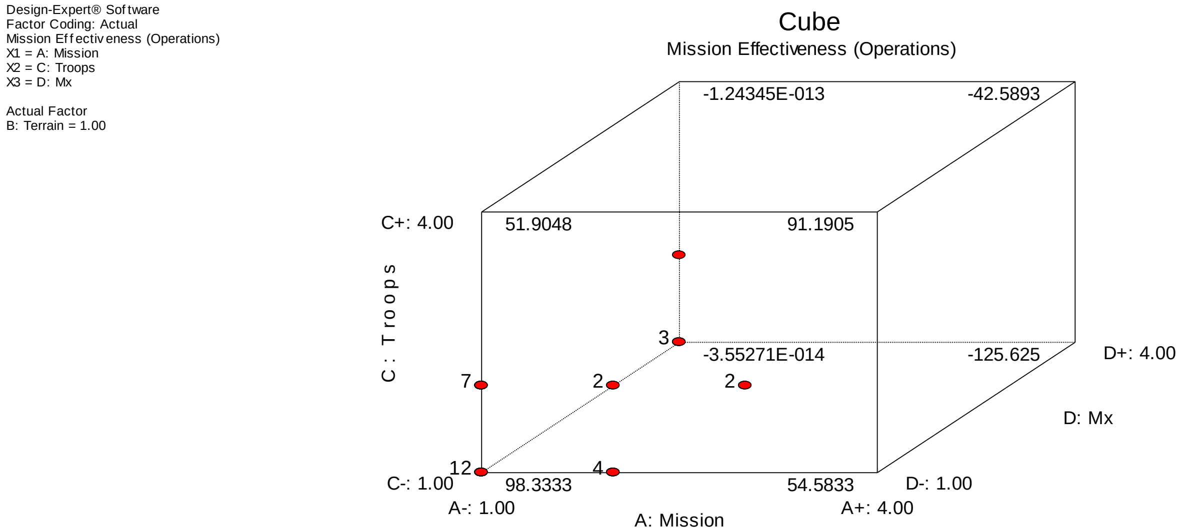
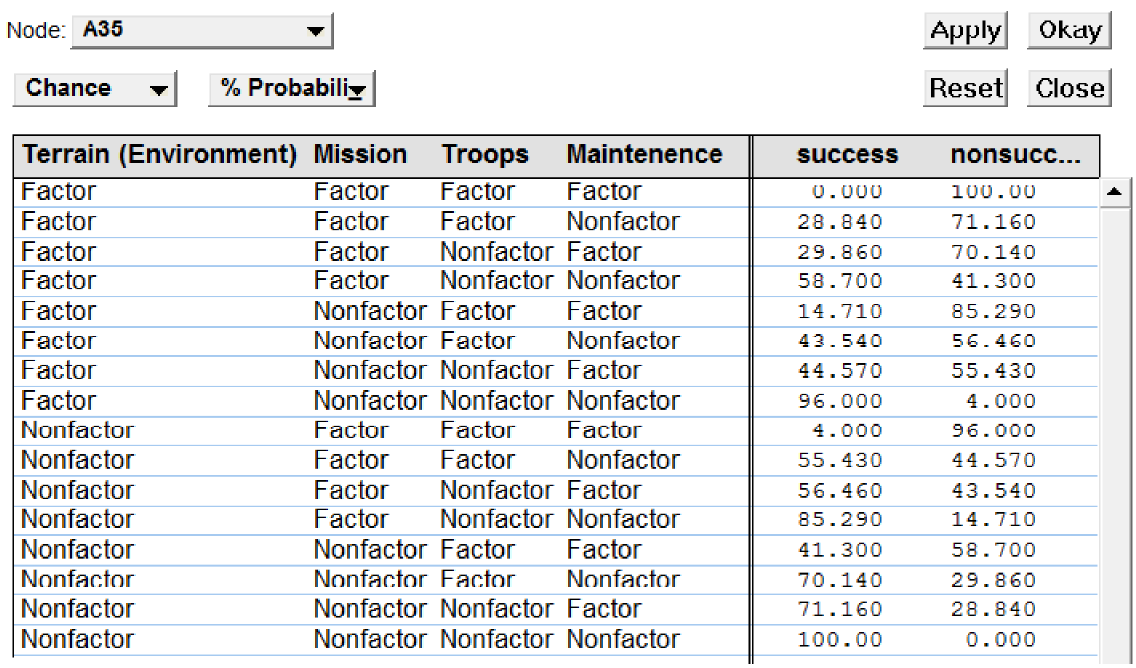


Figure 3-1. Design Interaction Plot



Terrain (Environment)	Mission	Troops	Maintenence	success	nonsucc...
Factor	Factor	Factor	Factor	0.000	100.00
Factor	Factor	Factor	Nonfactor	28.840	71.160
Factor	Factor	Nonfactor	Factor	29.860	70.140
Factor	Factor	Nonfactor	Nonfactor	58.700	41.300
Factor	Nonfactor	Factor	Factor	14.710	85.290
Factor	Nonfactor	Factor	Nonfactor	43.540	56.460
Factor	Nonfactor	Nonfactor	Factor	44.570	55.430
Factor	Nonfactor	Nonfactor	Nonfactor	96.000	4.000
Nonfactor	Factor	Factor	Factor	4.000	96.000
Nonfactor	Factor	Factor	Nonfactor	55.430	44.570
Nonfactor	Factor	Nonfactor	Factor	56.460	43.540
Nonfactor	Factor	Nonfactor	Nonfactor	85.290	14.710
Nonfactor	Nonfactor	Factor	Factor	41.300	58.700
Nonfactor	Nonfactor	Factor	Nonfactor	70.140	29.860
Nonfactor	Nonfactor	Nonfactor	Factor	71.160	28.840
Nonfactor	Nonfactor	Nonfactor	Nonfactor	100.00	0.000

Figure 3-2. Design Conditional Probability Table

Mission Effectiveness (Operations) =
+166.45833
-14.71230 * Mission
-29.86111 * Troops
-28.83929 * Maintenance
+9.22619 * Mission Troops
-9.09722 * Mission * Maintenance
+5.15873 * Troops * Maintenance

Figure 3-3. Design Regression Analysis taken from Anova table

CHAPTER4
RESULTS

4-1 Models and Computational Results

Once these adjustments are made, the results improved to their best possible correlation factor. In Table B-1 is the list of all the individual factors and their rating for each run, along with the MSP and actual Overall Mission Effectiveness (OME). According to the ANOVA table (Table B-4), the F-value of 219.18 implies that the model is significant, and the p-value is < .0001, meaning that there is less than a 0.01% chance the F value was the result of noise. The "Lack of Fit" value of The Adequate Precision Ratio is 34.429. This implies there is good correlation, because APRs of less than 4 imply there is too much noise or variance. In addition, the normal probability plot in Figure 4-2 shows there is a normal distribution of the MAP with no significant outliers. Overall, these results show that the model built in Netica™ shows successful correlation to actual mission results.

4-2 Significant Factors

Here is the list of significant factors ordered by most significant to the least significant, measured by the difference of individual factors' magnitude on the MSP at its full effect ("Extreme) and its null effect ("Low"). Figure 4-3 represents the magnitude order prior to allocating a separate section for maintenance, and Figure 4-4 represents the order after the change.

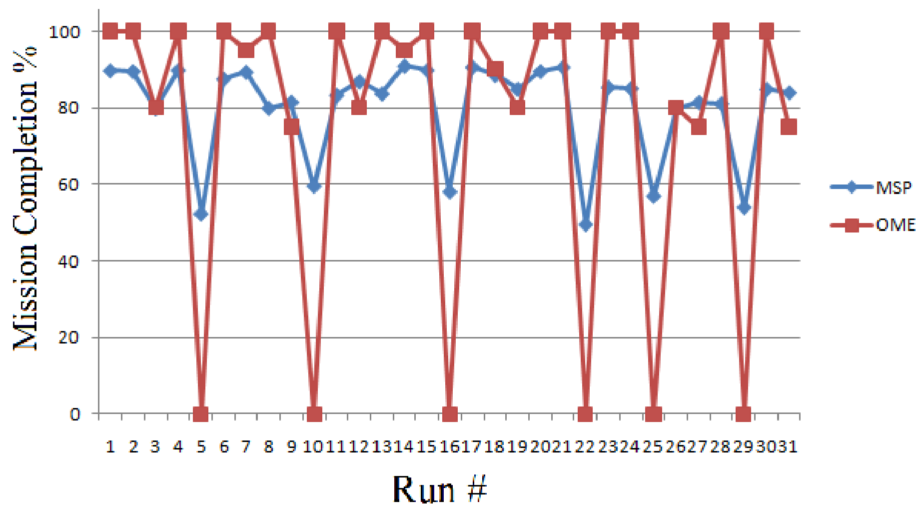


Figure 4-1. Predicted Mission Success Probability compared with Overall Mission Effectiveness

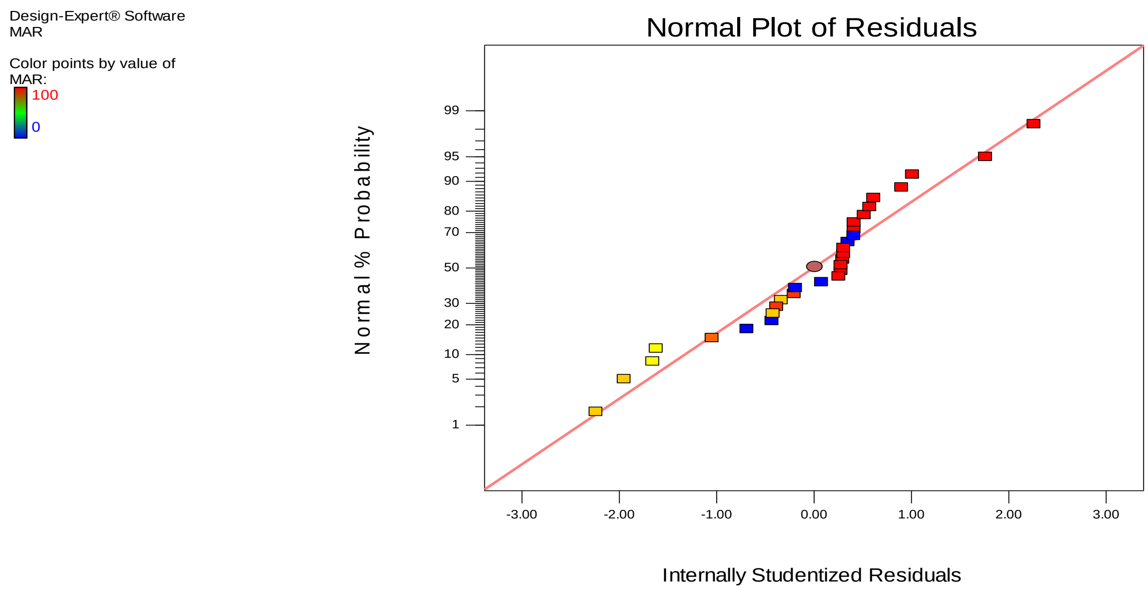


Figure 4-2. Normal Probability Plot of Model Results

1. Troops………….22.00%
2. Mission………..21.00%
3.Inflight Fatigue………… …………………………19.20%
4. Human Factors……… ………………………..15.00%
5. Enemy…………12.10%
6. Time……………11.10%
7. Preflight Fatigue……. ………………………..11.00%
8. Profile……………9.20%
9. Waivers (Personnel)………..8.00%
10. Qualification/ Proficiency……….6.60%
11. Prob. Of Defeating Threat……………..6.30%
12. Terrain……….5.60%
13. Prob. Of Engagement…….5.60%
14. Priority……….3.70%
15. Supporting/ Supported Forces…….. ………………………..3.00%
16. Weather………2.90%
17. Recovery/ Divert………………..2.20%
18. Complexity….1.80%
19. Preparation….1.70%
20. Objective……..1.30%
T-21. BASH…………1.00%
T-21. Execution… 1.00%
23. Enroute……….0.70%
24. Familiarity…..0.60%
25. Intelligence….0.50%
T-26. Deconfliction…….. ……………………………0.40%
T-26. Type………….0.40%
T-28. Illumination…………….0.30%
T-28. Prob. Of Detection….0.30%
T-30. Performance…………..0.20%
T-30. Start/Taxi/TO…………0.20%
32. Waivers (AC)…………… 0.10%
33. Degraded Aircraft………0.10%
34. Planning……………………0.02%

Figure 4-3. List of Factors with their Magnitude Effect on MSP, no separate Maintenance Effect

1. Troops…………………… 34.1%
2. Maintenance…………32.9%
3. Start/Taxi/Takeoff….29.6%
4. Inflight Fatigue……….27.4%
5. Degraded Aircraft…..26.2%
6. Human Factors……….21.4%
7. Mission…………………18.9%
8. Time………………………17.4%
9. Preflight Fatigue…….13.7%
10. Terrain…………………11.4%
11. Waivers (Troops)….10.9%
12. Qualification/ Proficiency……………………9.0%
13. Profile………………7.6%
14. Weather…………..5.2%
15. Enemy………………4.6%
16. Preparation………3.8%
17. Objective…………3.6%
T-18. Enroute…………3.2%
T-18. Priority…………3.2%
T-18. Recovery/ Divert……………………..3.2%
21. Performance…….3.1%
22. Supporting/ Supported Forces….2.6%
23. Waivers (Terrain)………………..2.5%
24. Complexity……….2.4%
25. BASH……………………..1.7%
26. Execution………………1.4%
27. Familiarity……………..0.8%
T-28.Deconfliction……….0.6%
T-28. Illumination………..0.6%
T-28. Type…………………...0.6%
31. Probability of Defeating Threat…………………………0.4%
32. Intelligence…...........0.3%
T-33. Planning………………0.2%
T-33. Probability of Engagement………………..0.2%
35. Probability of Detection………………………0.3%

Figure 4-4. List of Factors with their Magnitude Effect on MSP, Maintenance Effect included

CHAPTER 5
DISCUSSION

5-1 Lessons Learned

Maintenance

The "Maintenance" factor requires its own category. The original proposal called for the maintenance factor to be a subset of the terrain category. However, upon the initial trial, I found that the most significant source of error was the "Maintenance" node in terms of frequency and magnitude, so much so that it rendered the Netica™ model statistically insignificant. As detailed, I made adjustments and corrected the discrepancy allowing the model to have significant factors with an F-value of 157.01 whereas there were no significant factors before.

The effect of Human Factors

From Figure 4-4, we see that the human element, or "Troops", has the most significant effect on our MSP with a factor of 34.1%. There are several reasons to explain this. As stated previously, this is the most controllable factor in aviation and has the most variability in system effect. "Machines are always much more reliable because they can be redesigned. Human nature is not so easily changed" [3]. The ability to accept and mitigate a certain amount of risk while successfully completing a mission lies with the user, so all of the elements affecting this ability, such as physical condition, personal distractions, or overall proficiency, should have the most bearing. In other words, "Maintenance", "Terrain", or "Mission" can be all rated as "Extreme", but the decision and ability to complete a mission in spite of these risks lies within the "Troops" category. This relationship is well documented in civilian aviation as well in terms of accident rates. As we can see in Table 5-1, the human element is causal for 71.9% of all US aircraft accidents in 2007.

The effect of random events

Most statistical test models are rarely ever given a 100% reliability rate. Although, ultra-reliability and 6-sigma methodology can bring us close, eventually there will always be an outlier data point that decreases this reliability rate. These random events are certainly a factor in the MSP Netica™ model, as none of the nodes in their natural state ever have a 100% probability rating. This presents an issue because by nature Bayesian calculations will continually decrease a system's probability with more iteration if the factor probabilities are fewer than 100%, which they all are. We see this in our Netica™ design as selecting a "Low" setting in the factors has less of an impact the more displaced the node is from the actual MSP node. This results in a MSP of 88.5% with all nodes are set to their natural state, and if all our controllable nodes are set to "Low", the best MSP is 91.0%. Although, this presents a disparity with missions with actual OMEs of 100%, which occur quite often (48.4% of data set missions), it is telling of the effect of random events not covered in the ORM worksheet. This means that random events not accounted for when the ORM worksheet is completed have an 11.5% (100% - 88.5%) average effect on all missions covered in this data set, and 9.0% (100% - 91%) when all controllable factors are minimized. This is the result of unforeseen factors appearing during the course of the flight. For example, the user may mark a "Low" setting for "Profile", but during the course of the mission, an unforeseen tasking may be issued to the crew which would normally have raised the ORM level to "Medium" prior to takeoff. This tasking may have affected the OME, and then factors into the overall effect of all of these unforeseen random events. Users of the Netica™ worksheet should recognize this when looking at the MSP and understand that the MSP is limited to fewer than 100% due to the overall average effect of random events in the conditional probability calculation.

5-2 Recommendations

For these reasons, I support my original thesis to make use of the Bayesian Belief Network approaches in order to determine risk level of a flying mission over the traditional ORM Matrix. To recap, these are the following improvements on the standard ORM Matrix

Maintenance

The influence of the maintenance factor has been well documented in this thesis. The bottom line is that it affects 25.8% of the missions in the sample group and is the direct cause of 100% of the cancelled missions (0% Overall Mission Effectiveness), resulting in an overall 32.9% significance factor from Figure 4-4. In this thesis, the magnitude was recognized and addressed by dedicating a separate category “Maintenance”. Not only did this decision result in the Netica™ model being functional with a strong correlation factor, it also drastically changed the order of individual factor significance in the model, as seen in the difference between Tables 4-1 and 4-2. Figure 4-3 represents the Netica™ model without a separate “Maintenance” category and many of the “Maintenance”-related factors are prioritized in the bottom-third of all factors, but in the current usable model (Figure 4-4) they are in the top-third with the “Maintenance” node given the second highest rating (32.9%) in table.

The traditional ORM worksheet only provides a subjective estimate of the “Maintenance” factor and categorizes it as only a partition of the “Terrain” category. The flaw here is that the Netica™ model combines elements from other “Terrain” factors to affect and diminishes the significance of the “Maintenance” factor. For example, in the traditional matrix the user may possibly encounter an “Extreme” risk rating in the “Start/Taxi/Takeoff (Maintenance)” factor, meaning that a cataclysmic mechanical incident, such as a gear malfunction, will probably happen in the initial portion of the flight. Coupled with “Low” ratings for the remaining items of the “Terrain” category, we will only see a 0.1% maximum change (Figure 4-3). This would be erroneous in 25.8% of the cases in this study that contain an “Extreme” risk for Maintenance issues and have ultimately resulted complete mission cancellation. The Netica™ model with a separate “Maintenance” category would be able to accurately model the “Extreme” rating, decreasing the MSP by 29.6% (Figure 4-4) in the same example, alerting the user that the probability of mission failure is very high.

Identification of the most significant factors

In the traditional ORM matrix, the user must subjectively identify the top risks for the mission. Although this is usually done easily in less risky missions in scenarios where risk is more frequent, this is less easily quantifiable. In these missions where there are multiple significant factors, it is important to prioritize the top risks as there is often only a limited amount of time that can be dedicated to mitigation and only the most important factors can be addressed. If we follow the order of factor significance shown in table, we will have a general guideline as to the order in which should prioritize present risks. An example would be occurs quite often in the combat theater for missions that involve casualty evacuation. Due to the urgent nature of the mission, aircrews will encounter “Time”, “Human Factors”, “Preflight Fatigue”, and “Preparedness” risk factors. If they only have time for address a certain amount of these risks, the order in which they should emphasize their mitigation techniques according to Figure 4-4 is “Human Factors, “Time”, “Preflight Fatigue”, and then “Preparedness”.

Thresholds

Much like the 9th SOS Operational Risk Management worksheet, this design should contain thresholds as indicators to notify leadership when the MSP reaches a certain level. From the model analysis in ANOVA (Table B-4), the mean is 75.00 with a standard deviation of 8.1.

Using this, we determine the following threshold much in the same manner a grading scale is made.

1. >83.1%: No notification necessary

2. 75.0-83.1%: Assistant Director of Operations/Director of Operations Approval

3. 66.8-74.9%: Squadron Commander Approval

4. <66.8%: Group Commander Approval

As a quick reference, this scale has good fidelity as 100% of the failed missions have MSPs in the 4th category and 86.7% (13/15) of the missions with a 100% OME have MSPs in the first category.

User Control of the worksheet still remains the same as a traditional ORM Matrix

It is an often repeated mantra of that the responsibility of the outcome of the mission falls solely upon the operator. There is a limit to the role of technology in a mission, and the user ultimately accepts the risk. This is also the case with the Netica™ Belief Network, where the user has the final decision to determine the risk, and the tool functions only as a reference tool. The Netica™ Belief network allows users to adjust any node as they see fit, including the child nodes if they do not agree with the calculated results. For example, the user rates “Preflight Fatigue”, “Human Factors”, and “Inflight Fatigue” all as “Medium” in the parent nodes, but the child node “Troops” only gives a 15.0 % effect. The user may feel that this may be too low of an effect since all of these factors compounded warrant a “High” rating, and can elect to change the child node “Troops” to reflect that. However, if these child nodes are changed they will be annotated as such when the child node itself turns to a purple color (Figure 5-1). Doing so would not be generally advisable unless the conditions warrant, because it would negate the Bayesian calculation from the parent nodes.

5-3 Areas for Continued Research

Maintenance

In this thesis, I propose that a separate probability matrix or belief network to be dedicated to the “Maintenance” factor and completed by maintenance personnel, with operational review. This includes direct input from maintenance personnel and possibly an entire belief network dedicated to the maintenance process. For example, a “Maintenance” belief network would be constructed with the same framework of the final one are presented in this thesis, with the main categories being the major components of the aircraft, being “Engines and Communications”, and “Navigation”, “Environmental and Electronics”, “Hydraulics” and so forth. Also, “Troop Condition” would also be a category as it was in the mission success network. This would function much the same way the “Troops” subsection factors in the original ORM worksheet as an indicator of Maintenance Troops physical well-being, which as stated before has a significant bearing on task completion rate. The goal would be for the Maintenance Bayesian Model to furnish an overall capability rating based on the confidence levels of the subcategories, and this rating would then be used in the Mission Bayesian Model as the maintenance factor. An example Netica™ network is seen in Figure 5-2. This would better complete the strength of the overall mission success probability, by ensuring higher fidelity in the most significant variable. A study should be conducted much like this one, into the effectiveness and accuracy of the network. This would better explain the most significant factor is mission completion rate.

Some hypotheses to test that may improve maintenance effectiveness would be

- How much does human fatigue/maintenance rest affect job completion rate?
- How much do external factors (i.e. time lines, operational pressure, leadership pressure) affect job completion rate?
- How much does experience affect the job completion rate of time?

These are just some of the issues the Netica™ network would provide evidence for possible improvements in the maintenance process.

Mitigation Factor

In Section "5-1 Lessons Learned" we saw that even at the lowest levels of risk in the Netica™ design, the MSP at its highest was 91.0% which presents a 9.0% average disparity with 15 missions in our data set with a 100% OME. I believe that, for the missions in which the aircrew completed the mission with a higher OME than the Netica™-calculated MSP, the aircrew practiced good Risk Management techniques as prescribed in Section "1-2 ORM Utility to Air Force Flying Units" and negated the amount of risk present inherent in the mission. This is a good point for research, as to the quantification of the Risk Mitigation techniques. For example, the researcher may determine that if the "Inflight Fatigue" factor is at an "Extreme" rating, reducing the overall MSP by 27.4% with all other factors constant (Figure 4-4), the factor can be reduced down to 15.0% if there is cold water and extra crew members on board to relieve the fatigued crew. This data can be used to qualitatively rate different mitigation techniques for corresponding risks.

The effect of "Extreme" data points

A basic text of the Bayesian Belief network is it produces outcome likelihoods based on incomplete data. This is evident in the "High" and "Extreme" rating levels, situations that occur with little to no frequency. Because of this infrequency, there is a lack of data to support many of the “High” and “Extreme” risk scenarios, and so belief probabilities must be used to best estimate the effect. For example, there are no “Extreme” BASH risk scenarios in any of the sample, nor have I ever seen this type of risk in my flying experience. An “Extreme” BASH rating would probably constitute a sky full of birds essentially forming an impassable layer in the sky. Therefore, we have to best guess the effect of this type of setting to calculate the MSP. Perhaps research into the safety reports of aircraft mishaps may reveal more negative effects of “high” or “extreme” settings. A good example of this type of data source is the NALL Report [3] which reports annual aircraft accident data, categorized by the source of risk leading to the accident. Figure 5-3 is a breakdown of aircrafts mishaps by different portions of a flight containing varying levels of risk.

In addition, the lack of data for “High” and “Extreme” scenarios also presents a flaw in this thesis in terms of the “Enemy” category. There is little data of the effect of a “High” or “Extreme” risk level for “Enemy” because of the lack of actual experiences with the enemy so much of this network is determined by belief. To increase fidelity in this area, more research is necessary into incidents where actual enemy contact has occurred or scenarios with a “High/Extreme” rating. These scenarios can also be performed and gauged in a simulator as well. Of course, much of this information would be controlled and classified, so this would be a research topic to be implemented only in a closed or wartime environment, withholding results to the public. The research would involve much historical extrapolation possibly from previous wars/conflicts where actual enemy contact has occurred. Figure 5-4 gives us a comparison of aircraft loss rates during the Vietnam War, which is classified as a “High” threat environment, as compared to aircraft lost in Operations Iraqi and Enduring Freedom, which are classified and

“Low/Medium”. We see that the loss rates are 7 times more frequent in Vietnam as it is in OIF/OEF, and although there are numerous other variables that would play into the final “Enemy” calculation, this gives us an idea of the type of historical data we can refer to in differing threat environments.

Table 5-1. Adapted from “General Aviation Accidents 2007” [3]

Major Cause	All Accidents	Fatal Accidents
Mechanical	219 (15.8%)	19 (7.5%)
Other or Unknown	170 (12.3%)	42 (16.7%)
Pilot-related	996 (71.9%)	191 (75.8%)

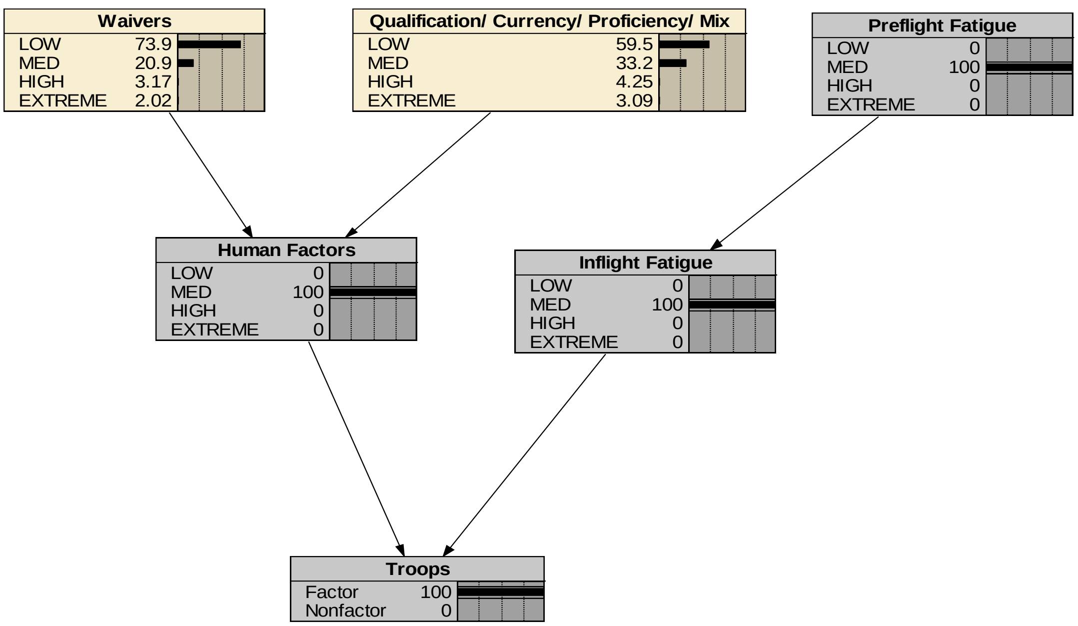


Figure 5-1. Altering child node example

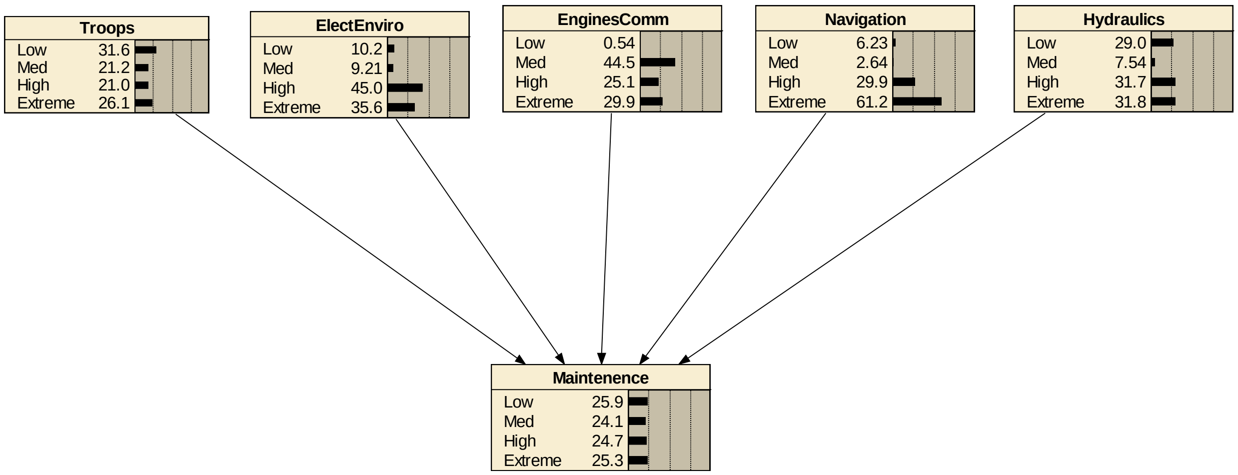


Figure 5-2. Example Maintenance Netica™ Network

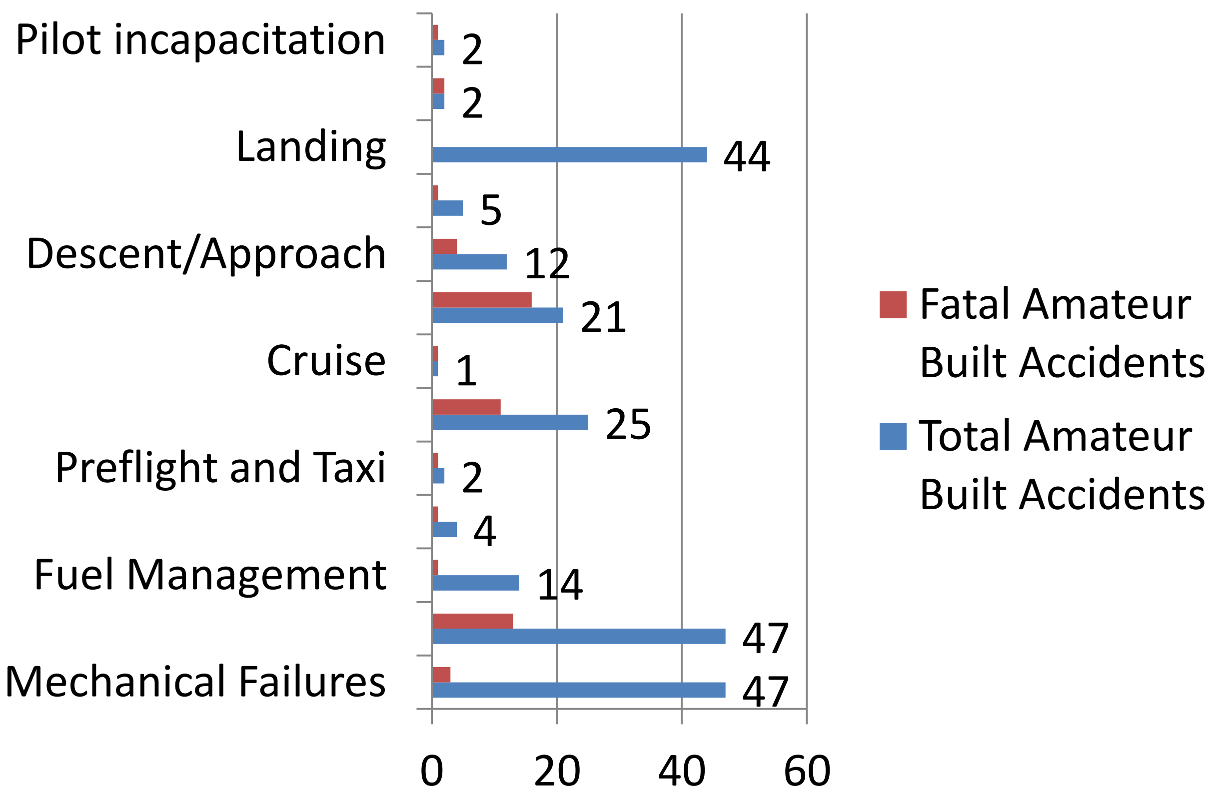


Figure 5-3. Adapted from "Aircraft Mishaps in 2007 Categorized by Phase of Flying" [3]

Table 5-4. Adapted from "Comparison of OEF/OIF Rotorcraft Hostile Combat Action Losses to Vietnam Conflict" [12]

	Vietnam (Attack and Observation)	OEF/OIF	Vietnam (Cargo and Utility)	OEF/OIF	Vietnam (Total)	OEF/OIF
Losses	757	35	1,309	35	2,066	70
Fatalities	644	33	2,421	112	3,065	145
Fatality/Loss Ratio	0.85	0.94	1.85	3.2	1.48	2.07
Flight Hours	2,927,130	1,310,619	9,777,753	1,705,654	12,704,883	3,026,483
Combat Loss Rate	25.86	2.67	13.39	2.05	16.26	2.31
Combat Fatality Rate	22	2.52	24.76	6.57	24.12	4.79

CHAPTER 6
CONCLUSION

In this thesis, I have created a mission success probability calculator based on a Bayesian Belief Network that incorporates mission factors from a traditional U.S. Air Force ORM (Operational Risk Management) Worksheet. This worksheet coupled with their associated mission completion summaries provided me with enough historical data from a 2 month flying period to complete the probability tables for each factor. I began the thesis by explaining my reasoning behind the use of ORM worksheets, the Netica™ program, and my development of the Belief Network. After a few trial runs, I found flaws in the network and made adjustments until it produced a satisfactory result. It was necessary to attribute an entire category to “Maintenance” in order to make the model statistically significant. I used Design Expert® software to determine the weighting system for each factor in the model based on historical data from the initial data set, and also to determine the validity of the results. The results showed a mean 75.0% with 8.1% standard deviation and F-value of 219.18 correlation factor. These results not only validated my original thesis but multiple learning points surface as well, the most important one being the significance of the human and maintenance factor. From these learning points as well as other questions I encountered during the experimental process, I found areas for continual research that would provide insight into the nature of probability analysis and its application to flying operations. This thesis provided a contribution towards the improvements to an operational system in the realm of military aviation and validated the results with System Engineering Tools.

APPENDIX A
FIGURES

9 SOS Operational Risk Management Program (Aug 2006)

DATE: S/Umy CALL SIGN: Slash 91 Aircraft Commander:

MC-130P	L	M	H	E	SPECIFIC CONSIDERATIONS (Circle Applicable Items)
MISSION					
Type					Training/FCF/Exercise/Contingency/Combat
Complexity					Low Level / AR / ~~HAR / Form~~ / NVG Land / AD
Priority					Bilateral / ~~Multi-Lat~~ / HHQ Directed
Deconfliction					Large Air Component / Dissimilar Formation
Familiarity					Unfamiliar Routes / DZ's / LZ's / Events not often accomplished
Profile					Planned evasive maneuvers / Low Flap setting drops / Checkrides /High Altitude Drops
					Simulated EP's / Confidence Maneuvers / Seat Swaps / DVs / FARP / AERPS
Supporting/Supported Forces					Unfamiliar SOF-CSAR User, Personnel / Equipment Infil/Exfil / Non-SOF-CSAR user
Overall MISSION is rated:					
ENEMY					
Intelligence					Threat Locations / Known threat w/ no known locations
Probability of Detection					Possible / Mobile Threat / Probable / Daytime 100% Detection
Probability of Engagement					Possible / Hostile enemy posture / Threats on DZ/LZ leg
Probability of Defeating Threat					Likely / Unlikely
Overall ENEMY is rated:					
TERRAIN (ENVIRONMENT)					
Degraded Aircraft					~~No Spare~~ / Missing Non-MESL / No MX Avail /Missing MESL Equip
Start/Taxi/Takeoff					Delayed takeoff / Extended ETIC / Unfamiliar airport
Enroute					Significant Charted Hazards / Uncharted Flt Area / IMC in Non-Radar Environment (OEF/O
					Desert / Over Water / Low Contrast / Mountains / High DA
Objective					Blind DZ / ALZ / Unsurveyed / Uncontrolled / Mountainous / Unfamiliar DZ Markings
Recovery/Divert					Complex Approach Procedures / No procedures published / Non-Radar
Performance					Marginally above requirements / Barely meets requirements
Waivers					Performance, Suitability OG/CC and below / Above OG/CC required
Weather					Ceiling or Visibility Marginal / Ceiling or Visibility Below Minimums / Departure Alternate Re
					Precip--Lt, Hvy / ~~Icing--Lt, Mod,~~ Sev / Turb--Lt, Mod, Sev / TS--Iso, Few, Sct/ Severe
Illumination					Low Illum(<0.8mlux or <10%) and/or Low Contrast / Adverse Cultural Lights
BASH					BWC Moderate / BWC Severe
Overall TERRAIN is rated:					
TROOPS					
Preflight Fatigue					O <0800L / Circadian Rhythm Chg / Marginal Crew Rest / Chronic Fatigue / Double-Triple
In-flight Fatigue					Land > 0100 / > 10 Hrs Day / > 12 Hrs Day / > 6 Hrs NVG / > 8 Hrs NVG
					Prolonged Hot (>85); Cold (<32) Exposure / Extreme Hot (>95); Cold(<10)
Human Factors					Minor Personal Distractions / Major Life Events / Senior Supervisor Flying / DV On Boar
Qualification/Currency/Proficiency/Mix					Non-current / Non-proficient / Inexperienced / Non-hard Crew / Unqualified Member
Waivers					Duty day, Crew rest, etc. / OG/CC level or below / Above OG/CC level required
Overall TROOPS is rated:					
TIME					
Planning					Day Prior Planning / Day of Planning / Compressed Brief / Alert Launch / Airborne Taskir
Preparation					General Plan / Guidelines Only / Changing Criteria
Execution					Multiple Events, TOTs / Multiple Customers / Local Area Sortie / Off-station Sortie
Overall TIME is rated:					

TOP RISKS FOR THIS MISSION	RISK MITIGATION FACTORS
1 Fatigue	1 Short Day ; Good CRM, Familiar
2	2
3	3
4	4

Overall Assessment: LOW--AC MODERATE--ADO/DO HIGH--DO/CC EXTREME--OG/C

Initial by appropriate position:

AC ____ CP ____ NAV KK ____ NAV ____ FE RMC ____ ACES ____ LM ____ LM ____

Additional crewmembers: ____ ____ ____ ____ ____ ____

*AC will brief the appropriate supervisor based upon overall assessment prior to flight. If assessed to be "Extreme" contact DO or CC for course of actio

Figure A-1. 9th SOS Operational Risk Management Form

CALL SIGN: **KING 58**

130 RQS Home Station Risk Assessment

DATE: 11 APRIL 2011

TRAINING — X number of people it applies to		Sched	Actual
Event Re-Currency	1 x		
Event Re-Qualification	2 x		
Special Mission Initial Qualification	2 x		
Mission Re-Qualification	2 x		
Initial Qual/Cert	4 x		
Crewmember first Basic Sortie in > 14 days	1 x		
Crewmember third flight in 3 days	1 x		
Mulitiple seat/crew swaps planned	1 x		
Evaluations	3 x		
0		0	

Scheduling		Sched	Actual
No Navigator* - Non-Tactical WX < 3000/3	3		
No Navigator* - Out of Area	2		
No AMSS ON NCSS*	1		
Crew Member Double Turn	2		
Engine Running Onload/Offload	1		
TOTAL		0	

CREW REST		Sched	Actual
Less than 12 hours Crew Rest	3		
Flight Prior to or just after major Holiday	1		
Fatigue - per crewmember affected	1		
TOTAL		0	

CREW DUTY DAY - Apply once per Sortie		Sched	Actual
10-12 hrs	1		
12-14 hrs	2		
> 14 hrs	5		
Flight Planned to terminate after 2300L	3		
TOTAL		0	

FAMILIARIZATION SORTIES		Sched	Actual
Fam for new Crewmember	1		
Fam or incentive passenger onboard	1		
High Visibility Mission (airshow/demo/inspection/DV)	2		
TOTAL		0	

AIRCRAFT		Sched	Actual
Maintenance Delay	2		
Aircraft has not flown in over 30 Days	1		
Aircraft has notable MX problem (MOC View)	2		
Same day change of configuration on aircraft	2		
First Flight with Benson Tank Installed	5		
Flight above 140K weight	1		
Spare Aircraft (After Crew Show)	4		
TOTAL		0	

BIRD AVOIDANCE		Sched	Actual
MOFFETT FAF			
BWC - **SEVERE** (BRIEF)	NO GO		
BWC - **MODERATE** (BRIEF)	3		
BWC - **LOW**	0		
TRAINING AREA			
BWC - **SEVERE** (BRIEF)	5		
BWC - **MODERATE**	3		
BWC - **LOW**	0		
TOTAL		0	

PLANNING		Sched	Actual
Changes to Route	1		
Changes to Sequence	1		
Changes to Events	2		
Compressed Planning Time	4		
Late Crewmember	1		
TOTAL		0	

WEATHER		Sched	Actual
VFR			
Day > 3 miles visibility	0		
Day < 3 miles visibility	2		
Night > 3 miles visibility	1		
Night < Marginal visibility	4		
IFR			
Day - No Alternate Required	1		
Day - Alternate Required	2		
Night - No Alternate Required	2		
Night - Alternate Required	3		
Forecast Adverse Weather			
Low Level Winds > 25 knots (actual or forecast)	2		
Moderate or greater Turbulence (actual or forecast)	2		2
Rain/Snow Showers on route	2		
Ground Fog	3		
Icing	3		
Thunderstorms	3		
TOTAL		0	2

TERRAIN		Sched	Actual
Unfamiliar Area	1		
Overwater > 500' AWL	1		
Overwater < 500' AWL	4		
TOTAL		0	

ILLUMINATION - Tactical Events Only		Sched	Actual
Twilight Transistion	2		
> 60%	2		
10% - 60%	1		
< 10%** (Requires mitigation and OG Notification)	3		
TOTAL		0	

TEMPERATURE		Sched	Actual
> 90°F (32°C)	1		
40°F - 90°F (4°C - 16°C)	0		
15°F - 40°F (-10°C - 4°C)	1		
< 15°F (<-10°C)	2		
TOTAL		0	

Notifications - Initials OG/CC or Rep ______________

ALL APPLICABLE ITEMS OVER 3 POINTS WILL BE BRIEFED AND MITIGATED IF POSSIBLE

* No Nav/AMSS Requires OG/CC Notification

MISSION		Sched	Actual
Off Station / Cross Country	2		
FCF/One Time Flight for MX	5		
Emergency Procedures	1		
Day Low Level	1		
NVG Modified Contour	2		
Unfamiliar Route	2		
HAR - Day	1		
HAR - NVG	2		
HAR - Unaided Night	3		
Airdrop Equipment SATB/Bundles	1		
Airdrop Equipment - Actual	3		
Airdrop Live Personnel	4		
INFIL/EXFIL w/ User	2		
Actual Assault Zone - Day	1		
Actual Assault Zone - NVG/Unaided nt	3		
Actual Pyrotechnics	2		
AERPS	3		
Insufficient Map Coverage	1		
Unfamiliar/Out of Ordinary Profile	1		
Night Water Ops	4		
FARP/HOTGAS	4		
TOTAL		0	

EXERCISE		Sched	Actual
Administrative Airlift	1		
Bilateral US Forces (i.e. JA/AATT)	2		
Bilateral Allied Forces	3		
HHQ Exercise	4		
TOTAL		0	

OTHER		Sched	Actual
Approved Pharmaceutical Use	1		
Aircrew Arming	2		
TOTAL		0	

RISK ASSESSMENT TOTAL 0 /

		Initials
Low Risk	0-25	
Medium Risk - Notify Ops Sup	26-45	
High Risk - Notify SQ/CC	46-65	
Extremely High Risk - Notify OG/CC	66+	

Mitigating Actions Taken -

No, KIO

Aircraft Commanders Signature ______________

C:\Users\christopher.nance\Desktop\Copy of 130RQS ORM-APR 10

Figure A-2. 130th RQS Home Station Risk Assessment Form

Air Force Special Operations Command
Mission Summary Report

Day/Date	Show Time	Fuel Load	Call Sign
05/13/2010	12:00(L)	42K	Shado91
Sortie	Brief	Mission Configuration	Mission Type
0941	13:00(L)	TAC	T3XA
Aircraft Commander	Load Time	Aircraft Configuration	Mission Number
McWilliams			5EK0xxxxxxx
Tail Number	TakeOff Time	Mission Load	Mission Symbol
695822	16:00		T3XA
Tail Swap To	Land Time / Duration		Authorization Number
	22:00(L) / 4		A10-xxxx

Range MSN #	Range Name	Range Profile	Range Time	Remarks
	W50			Oceana Multilat

Flight Log

T/O ICAO	DATE	T/O (z)	LAND (z)	Hours
KNTU	13 APR 10	2012	2238	2.4
			Total Hours:	**2.4**

Aircraft Commander's Comments

Cancel	☐	**Maintenance**	☐
Late T/O	☐	**OPS**	☐
Ground Abort	☐	**Weather**	☒
Air Abort	☐	**Other**	☐

MX Debriefer's Initials

Additional Comments (Use back if needed):

User WX CX airdrop x 2. RTB VPS uneventful.

Training

Trainee	Event (T-Ride)	Status (I/C)

Risk Assessment

8	16	24	32	40
Low Risk	**Low Caution**	**High Caution**	**High Risk**	

Mission Events

Scheduled:						
	LL	**DF**	**LF**	**AR**	**PRO**	**ECM**
Accomplished:						

Other Events

MX Effectiveness
100
Training Effectiveness
100
Overall Mission Effectiveness
100

AFSOC FORM 40 **Air Force Special Operations Command Form 40**

Figure A-3. AFSOC Form 40

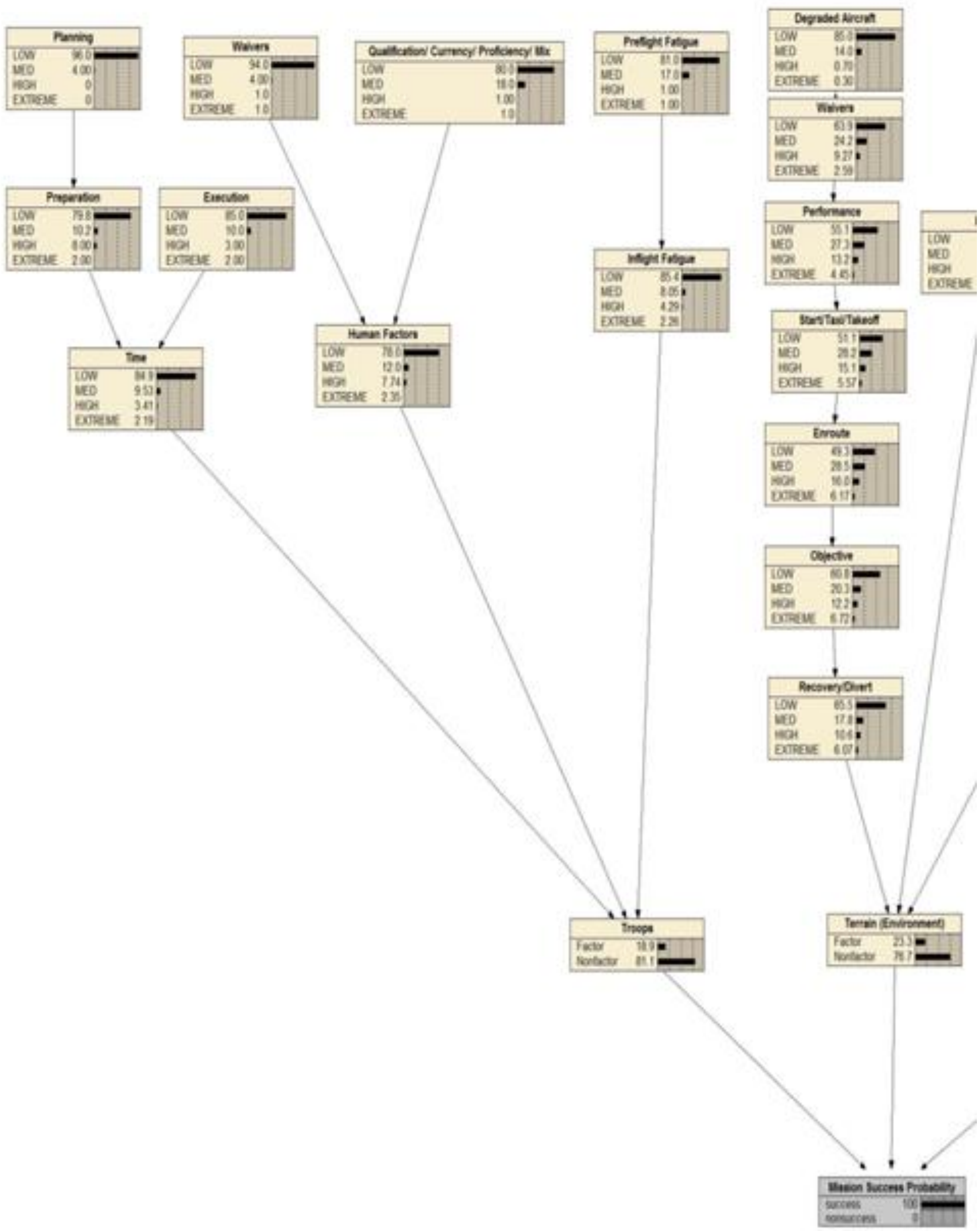


Figure A-4. Netica™ Model with Maintenance Category not separated

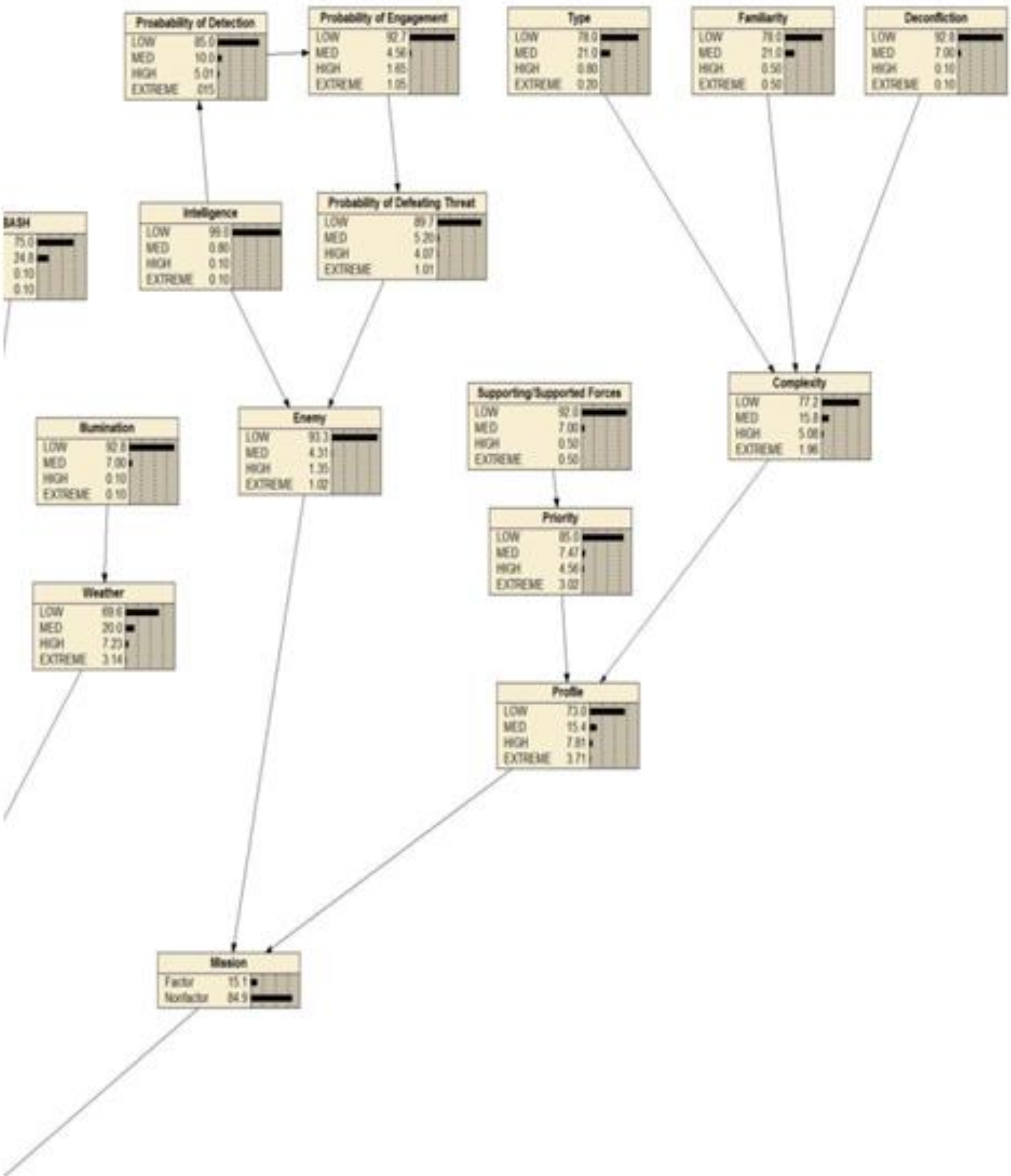


Figure A-4. Continued

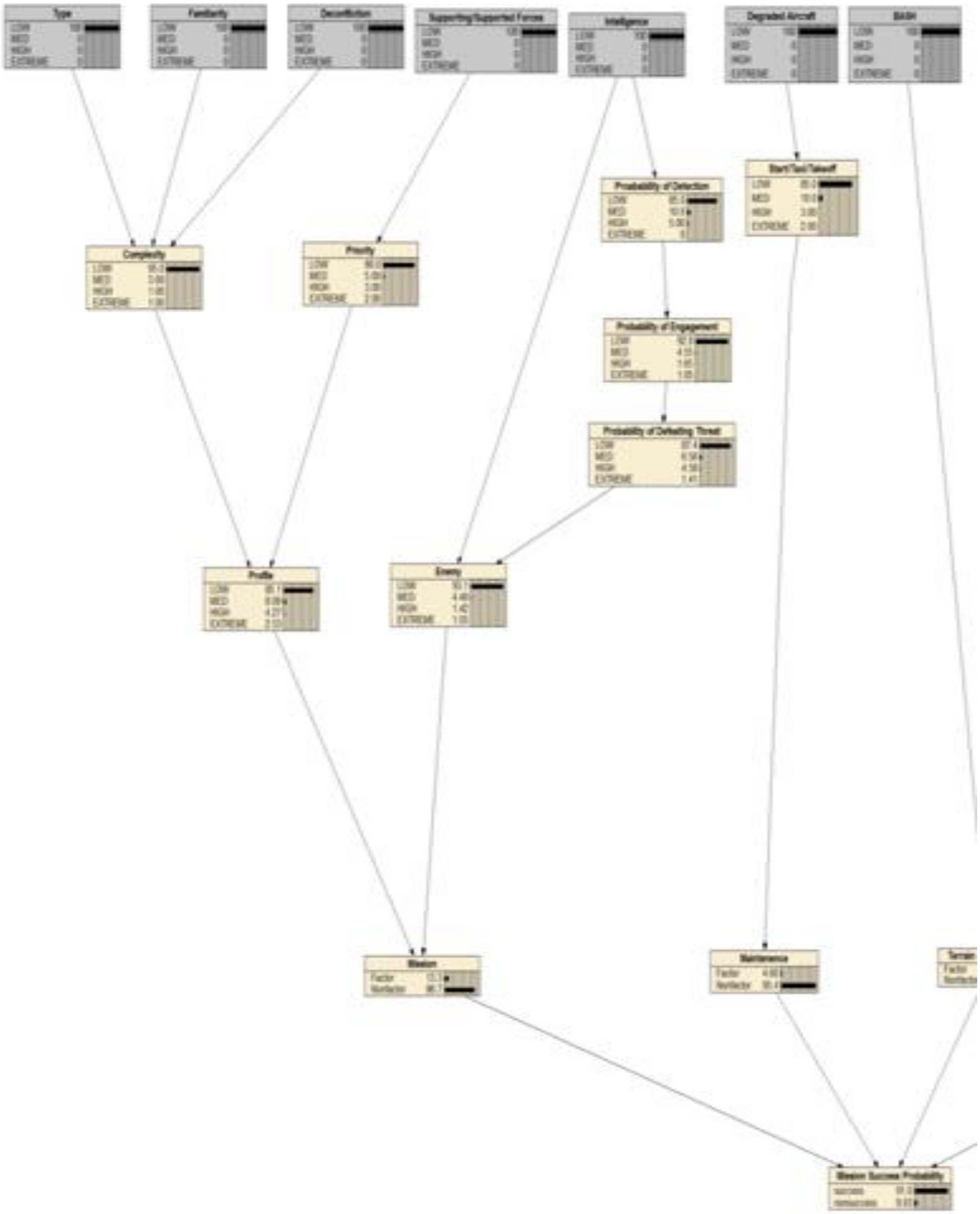


Figure A-5. Netica™ Model with separate Maintenance Category

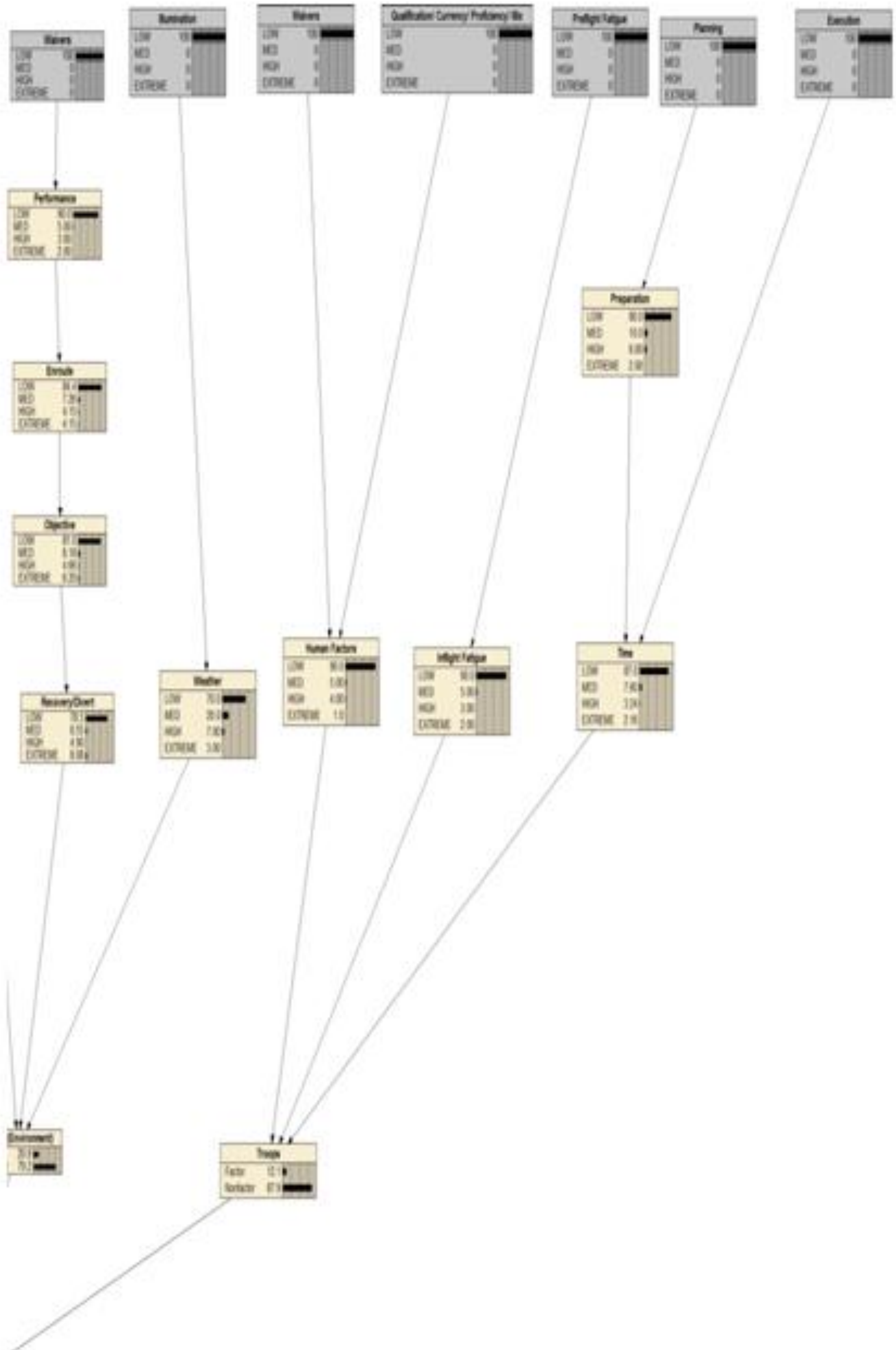


Figure A-5. Continued

APPENDIX B
TABLES

Table B-1. ORM Data with Overall Mission Effectiveness, May –June 2010

Msn#	Type	Complex	Priority	Deconflict	Familiarit	Profile	Supp Forc	Mission	Intelligen	Prob Det	Prob Eng	Prob DT	Enemy	Deg AC	ST/TO	en route	Objective	Rec./Div
1	2	2	1	1	2	2	1	2	1	1	1	1	1	1	1	1	1	1
2	1	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1
3	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1
4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
5	2	1	1	1	1	2	1	2	1	1	1	1	1	2	1	1	1	1
6	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1
7	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1	1	1	1
8	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
9	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
10	2	1	1	1	1	2	1	2	1	1	1	1	1	2	1	1	1	1
11	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
12	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
13	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
14	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
15	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
16	1	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1
17	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1
18	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
19	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
20	1	2	1	1	2	2	1	2	1	1	1	1	1	1	1	1	1	1
21	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
22	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1
23	2	2	1	1	2	2	1	2	1	1	1	1	1	1	1	1	1	1
24	1	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1
25	1	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1
26	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
27	1	2	1	2	2	2	2	2	1	1	1	1	1	2	1	1	1	1
28	1	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1
29	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
30	1	2	1	2	2	2	2	2	1	1	1	1	1	1	1	1	1	1
31	1	2	1	1	1	2	1	2	1	1	1	1	1	1	1	1	1	1

Table B-1. Continued

Maintena	Performar	Waivers	Weather	Illuminati	BASH	Terrain	Pref. Fatig	Inflight Fa	Human Fa	Qual/Prof	Waivers	Troops	Planning	Prep	Execution	Time	OME	MSP
1	1	1	1	1	1	1	2	2	1	1	1	1	1	1	1	1	80	84.2
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	100	88.4
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	100	88.8
4	1	1	2	1	2	1	1	2	2	2	1	2	2	1	1	1	0	57.2
1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	100	87.6
1	1	1	2	1	1	1	1	2	1	2	1	2	1	1	1	1	100	84.4
1	1	1	1	1	1	1	2	1	1	2	1	2	1	1	1	1	100	84.7
1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	95	89.2
1	1	1	1	1	1	1	1	1	1	2	1	1	1	1	1	1	100	87.9
4	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	0	59.9
1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	100	88.6
1	1	1	1	1	1	1	1	1	2	1	1	1	1	1	1	1	90	87
1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	100	87.2
1	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	100	87.1
1	1	1	1	2	2	1	2	2	2	1	1	2	1	1	1	1	80	75.6
1	1	1	1	1	2	1	1	1	2	1	1	1	1	1	1	1	100	83.5
4	1	1	2	1	1	1	1	2	1	2	1	2	1	1	1	1	0	63
1	1	1	1	1	2	1	2	1	1	1	1	1	1	1	1	1	100	81.7
4	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	0	59.6
1	1	1	1	1	1	1	1	2	2	2	1	2	1	1	1	1	80	84.9
1	1	1	2	1	2	1	1	2	2	2	1	2	2	1	1	1	75	85
1	1	1	1	1	1	1	1	2	2	2	1	2	1	1	1	1	100	85.9
4	1	1	1	1	1	1	2	2	1	1	1	1	1	1	1	1	0	58.5
1	1	1	1	1	1	1	1	2	1	1	1	1	1	1	1	1	100	83.4
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	95	87.5
1	1	1	1	1	2	1	2	2	2	1	1	2	1	1	1	1	100	76.5
1	1	1	1	2	2	1	1	1	1	1	1	1	1	2	2	1	75	74.6
4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	48
1	1	1	1	1	1	1	1	2	1	2	2	2	1	1	1	1	100	82.5
1	1	1	1	1	1	1	1	1	2	1	1	1	1	2	1	1	80	88.1
1	1	1	2	1	1	1	1	2	1	2	1	2	1	1	1	1	75	83.7

Table B-2. ORM Data with Mission Success Probability and Overall Mission Effectiveness, July –August 2010

Msn #	Type	Complexit	Priority	Deconflict	Familiarity	Profile	Supp Force	Mission	Intelligenc	Prob Det	Prob Eng	Prob DT	Enemy	Deg AC	ST/TO	Maintenan	en route	Objective
1	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
3	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
4	1	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1
5m	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	4	1	1
6	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
7	1	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1
8	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
9	1	2	1	2	2	2	2	2	1	1	1	1	1	2	1	1	1	1
10m	1	1	1	1	1	2	1	1	1	1	1	1	1	1	1	4	1	1
11	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
12	1	2	1	2	2	2	2	2	1	1	1	1	1	1	1	1	1	1
13	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1	1	1	1
14	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
15	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
16m	2	1	1	1	1	2	1	2	1	1	1	1	1	2	1	4	1	1
17	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
18	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
19	2	2	1	1	2	2	1	2	1	1	1	1	1	1	1	1	1	1
20	1	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1
21	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1
22m	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	4	1	1
23	2	1	1	1	1	2	1	2	1	1	1	1	1	2	1	1	1	1
24	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1
25m	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	4	1	1
27	1	2	1	1	2	2	1	2	1	1	1	1	1	1	1	1	1	1
27	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
28	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1
29m	2	2	1	1	2	2	1	2	1	1	1	1	1	1	1	4	1	1
30	1	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1

Table B-2. Continued

Rec./Div	Performan	Waivers	Weather	Illuminatio	BASH	Terrain	Pref. Fatig	Inflight Fat	Human Fa	Qual/Prof	Waivers	Troops	Planning	Prep	Execution	Time	MSP	OME
1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	89.8	100
1	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	89.6	100
1	1	1	1	2	2	1	2	2	2	1	1	2	1	1	1	1	79.8	80
1	1	1	1	1	2	1	1	1	2	1	1	1	1	1	1	1	89.9	100
1	1	1	2	1	1	1	1	2	1	2	1	2	1	1	1	1	52.1	0
1	1	1	1	1	2	1	2	1	1	1	1	1	1	1	1	1	87.6	100
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	89.4	95
1	1	1	1	1	2	1	2	2	2	1	1	2	1	1	1	1	79.9	100
1	1	1	1	2	2	1	1	1	1	1	1	1	1	2	2	1	81.5	75
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	59.4	0
1	1	1	1	1	1	1	1	2	1	2	2	2	1	1	1	1	83.3	100
1	1	1	1	1	1	1	1	1	2	1	1	1	1	2	1	1	86.9	80
1	1	1	1	1	1	1	2	1	1	2	1	2	1	1	1	1	83.7	100
1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	91	95
1	1	1	1	1	1	1	1	1	1	2	1	1	1	1	1	1	89.9	100
1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	57.9	0
1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	90.7	100
1	1	1	1	1	1	1	1	1	2	1	1	1	1	1	1	1	88.6	90
1	1	1	1	1	1	1	2	2	1	1	1	1	1	1	1	1	84.9	80
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	89.6	100
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	90.7	100
1	1	1	2	1	2	1	1	2	2	2	1	2	2	1	1	1	49.3	0
1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	85.4	100
1	1	1	2	1	1	1	1	2	1	2	1	2	1	1	1	1	85.1	100
1	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	56.8	0
1	1	1	1	1	1	1	1	2	2	2	1	2	1	1	1	1	80	80
1	1	1	2	1	2	1	1	2	2	2	1	2	2	1	1	1	82.4	75
1	1	1	1	1	1	1	1	2	2	2	1	2	1	1	1	1	82.1	100
1	1	1	1	1	1	1	2	2	1	1	1	1	1	1	1	1	53.8	0
1	1	1	1	1	1	1	1	2	1	1	1	1	1	1	1	1	84.9	100

Table B-3. ANOVA Table for Netica™ Model

Source	Sum of Squares	df	Mean Square	Value	p-value (Prob > F)	
Model	40852.8	6	6808.8	157.01	< 0.0001	significant
A-Mission	*90.62*	*1*	*90.62*	*2.09*	*0.1612*	
C-Troops	*16.57*	*1*	*16.57*	*0.38*	*0.5424*	
D-Mx	*1676.31*	*1*	*1676.31*	*38.66*	*< 0.0001*	
AC	*87.2*	*1*	*87.2*	*2.01*	*0.169*	
AD	*353.98*	*1*	*353.98*	*8.16*	*0.0087*	
CD	*153.58*	*1*	*153.58*	*3.54*	*0.072*	
Pure Error	1040.77	24	43.37			
Cor Total	41893.6	30				
Std. Dev.	6.59		**R-Squared**	0.9752		
Mean	72.58		**Adj R-Squared**	0.9689		
C.V. %	9.07		**Pred R-Squared**	N/A		
PRESS	N/A		**Adeq Precision**	31.424		
Case(s) with leverage of 1.0000: Pred R-Squared and PRESS statistic not defined						
Factor	**Estimate**	**Coefficient df**	**Error**	**Standard Low**	**95% CI High**	**95% CI VIF**
Intercept	15.97	1	15.29	-15.58	47.53	
A-Mission	-21.58	1	14.93	-52.4	9.23	13.56
C-Troops	9.15	1	14.81	-21.41	39.71	15.22
D-Mx	-58.03	1	9.33	-77.29	-38.77	31.91
AC	20.76	1	14.64	-9.46	50.97	18.88
AD	-20.47	1	7.16	-35.26	-5.68	17.29
CD	11.61	1	6.17	-1.12	24.34	11.47

Table B-3. Continued

Final Equation in Terms of Coded Factors:						
	Mission Effectivenes s (Operations)	=				
	15.97					
	-21.58	* A				
	9.15	* C				
	-58.03	* D				
	20.76	* A * C				
	-20.47	* A * D				
	11.61	* C * D				
Final Equation in Terms of Actual Factors:						
		Mission Effectivenes s (Operations)	=			
		166.458				
		-14.712	* Mission			
		-29.861	* Troops			
		-28.839	* Mx			
		9.22619	* Mission * Troops			
		-9.0972	* Mission * Mx			
		5.15873	* Troops * Mx			

Table B-4. ANOVA Table for Thesis Results with Maintenance Category Added

Source	Sum of Squares	df	Mean Square	F Value	p-value (Prob>F)	
Model	42566.8	3	14188.9	219.18	0.0001	significant
A-A	*115.53*	*1*	*115.53*	*1.78*	*0.1932*	
A2	1042.05	*1*	*1042.05*	*16.1*	*0.0005*	
A3	*1295.41*	*1*	*1295.41*	*20.01*	*0.0001*	
Residual	1683.17	26	64.74			
Lack of Fit	*1483.17*	*22*	*67.42*	*1.35*	*0.4262*	
Pure Error	*200*	*4*	*50*			
Cor Total	44250	29				
Std. Dev.	8.05			**R-Squared**	0.962	
Mean	75			**Adj R-Squared**	0.9576	
C.V. %	10.73			**Pred R-Squared**	0.9535	
PRESS	2059.12			**Adeq Precision**	34.429	
Coefficient Factor	**Estimate**	**Standard df**	**95% CI Error**	**95% CI Low**	**High**	**VIF**
Intercept	0.39	1	5.85	-11.63	12.41	
A-A	-97.86	1	73.26	-248.45	52.72	168.6
A2	749.85	1	186.9	365.67	1134.03	830.61
A3	-594.57	1	132.92	-867.79	-321.36	283.7
Final Equation in Terms of Coded Factors:						
OME	=					
		0.39				
		-97.86	* A			
		749.85	* A_2			
		-594.57	* A_3			
	Final Equation in Terms of Actual Factors:					
OME	=					
		1442.682				
		-67.6258	* A			
		1.01343	* A_2			
		-4.76E-03	* A_3			

LIST OF REFERENCES

[1] AIR FORCE INSTRUCTION 11-2HC-130 VOLUME 1. “HC-130 Aircrew Training.” 2 May 2007

[2] AIR FORCE INSTRUCTION 90-901. “OPERATIONAL RISK MANAGEMENT”. 1 APRIL 2000

[3] “AOPA AIR SAFETY REPORT”. *2008 NALL REPORT*. 25 Mar 2009

[4] D. B Pannell. “AIR FORCE SPECIAL OPERATIONS COMMAND RATED EXPERIENCE DECLINE: EVOLUTION AND SOLUTIONS.” 22 November 2006

[5] D. C. Montgomery, *Design and Analysis of Experiments*, 7th edition. 2009

[6] H.L. Pfister. Bayesian Belief Networks, Netica™ Software. 11 Nov 2008

[7] K. A. Weiss. “An Analysis of Causation in Aerospace Accidents”. 22 June 2001

[8] J.C Cantrell and H.L Andrews. “WHERE DOES THE AIR FORCE NEED OFFICERS”. *Airpower Journal*. Winter 1993

[9] K.D. Anthony. “Introduction to Causal Modeling, Bayesian Theory and Major Bayesian Modeling Tools for the Intelligence Analyst.” 27 Nov 2006

[10] J.L. Stanley. “ORM 101: A Basic Course in Common Sense”. *Combat Edge*. May 1999

[11] M. W. Maier and Eberhardt Rechtin. “The Art of Systems Architecting.” 2 Sept 2001

[12] M. Couch and D. Lindell. “Study on Rotorcraft Survivability”. *Aircraft Survivability*. 26 July 2010

[13] NAVAIR. “What is Operational Risk Management?” 4 Sep 2003

[14] Netica™ Application. User’s Guide. Version 1.05. March 15, 1997

[15] ORM University. “The 6-Step ORM Process”. http://www.seco.noaa.gov/Safety/ORM/ORMUCBT%201_0/fundamentals/chapter2/chapter.html

[16] R.L. Featherstone.” DETERMINATION OF CRITICAL FACTORS IN UNMANNED CASUALTY EVACUATION IN THE DISTRIBUTED ENVIRONMENT.” June 2009

[17] S.S. Therrien."A BAYESIAN MODEL TO INCORPORATE HUMAN FACTORS IN COMMANDERS' DECISION MAKING." June 2002

[18] Stat-Ease. *Software Overview with Masthead Software Overview.* 10 Sep 2010

[19] Bnet.builder. *About Bayesian Belief Networks.*2004

BIOGRAPHICAL SKETCH

Emmanuel Brian Cao was born in 1981 to Khai Cao and Bernadette An in Tustin, CA. He spent his early childhood in Santa Ana, CA before moving in 1988 to Oceanside, CA, a suburb north of San Diego. Raised by Vietnamese immigrants, Emmanuel learned at a young age the value of preparing and sacrificing for the future. Emmanuel then attended St. Augustine High School in San Diego where he played tennis, participated in many leadership clubs, and graduated with Magna cum Laude Honors in 1999. After high school, Emmanuel chose the University of California, Berkeley as his college. Pursuing a medical school scholarship, Emmanuel joined the Air Force Reserve Officer Training Corps but, like many college students, Emmanuel's changed his career choice midway through his collegiate studies. Inspired by his fellow AFROTC cadets into a love for aerospace and prompted by the Sept. 11, 2001 tragedy, Emmanuel decided to take a more proactive role in the military as an aviator. Emmanuel was chosen to be an U.S. Air Force Navigator and after graduating with a degree in Physical Sciences and earning his officer's commission in 2003, Emmanuel began his training which he completed in 2005. His first operational assignment was to the 9th Special Operations Squadron in Eglin Air Force Base, Florida as a Navigator on the MC-130P Combat Shadow Aircraft. There, Emmanuel fulfilled one of his childhood dreams to defend America's freedoms as an aviator, flying over 140 combat missions in Operations Iraqi Freedom and Enduring Freedom and accruing over 1800 hours on the aircraft. While at Eglin AFB, Emmanuel was inspired by his friends to continue his education, this time pursuing a Master Degree in Systems Engineering at the University of Florida. Emmanuel is currently stationed at the 130th Rescue Squadron, Moffett Federal Airfield in Mountain View, California as a Rescue Aviator. Emmanuel's hobbies include running, golf, traveling, and he has one sibling, Alice.