EFFECTS OF UNCERTAINTY
ON REAL WORLD AEROSPACE MISSION DATA

THESIS

Steven J. Barosko, Captain, USAF

AFIT/GOR/ENS/04-01

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT/GOR/ENS/04-01

EFFECTS OF UNCERATINTY ON REAL WORLD AEROSPACE MISSION DATA

THESIS

Presented to the Faculty
Department of Operational Sciences
Graduate School of Engineering and Management
Air Force Institute of Technology
Air University
Air Education and Training Command
in Partial Fulfillment of the Requirements for the
Degree of Master of Science in Operations Research

Steven J. Barosko, B.S.
Captain, USAF

March 2004

AFIT/GOR/ENS/04-01

EFFECTS OF UNCERATINTY
ON REAL WORLD AEROSPACE MISSION DATA

Steven J. Barosko, B.S.

Captain, USAF

Approved:

//signed//	5 Mar 04
Dr. J. O. Miller Thesis Advisor	Date
//signed//	5 Mar 04
Dr. Kenneth W. Bauer Committee Member	Date

AFIT/GOR/ENS/04-01

Abstract

Uncertainty is an age old element of warfare. How does one measure success or failure on the battlefield? This thesis explores the effect of uncertainty on real world aerospace mission data from Operation Iraqi Freedom (OIF). Sources of uncertainty are identified during data preparation and the effects of uncertainty are investigated using multivariate analysis techniques. Two distinct cases are analyzed: one case with a low level of uncertainty in predicting whether or not a mission is pre-planned or alert generated, and another case with a high level of uncertainty in predicting whether or not a mission is a success or a failure. Three multivariate analysis techniques are used on both cases: Signal-to-Noise Ratio (SNR) saliency, Feed Forward Neural Network (FFNN) supervised training, and General Regression Neural Network (GRNN) supervised training. Measures of performance are gathered from each of these techniques and analyzed to determine the effect of uncertainty on the data.

Acknowledgements

I would like to express my sincere appreciation to my faculty advisor, Dr. J.O. Miller, for his guidance and feedback throughout my thesis effort. His tireless efforts to obtain real-world data from OIF made this research effort possible. Also, I would like to thank my reader, Dr. Kenneth Bauer, for his insights, comments, and inspiration of the multivariate analysis techniques utilized in this thesis. Additionally, I would like to thank the folks at Riverside Research Institute, especially Candace, Julie, and Scott, for supporting my classified processing efforts.

Most importantly, I would like to thank my family for their prayers and encouragement throughout my time at AFIT. Finally, I would like to express special thanks to my wife and son for their love, support, and understanding during the countless hours away from home working on this thesis. I will forever be in debt to their sacrifices.

Steven J. Barosko

Table of Contents

List of Figures

Figure		Page

List of Tables

Table		Page

EFFECTS OF UNCERATINTY
ON REAL WORLD AEROSPACE MISSION DATA

1. Introduction

"The best-laid schemes o' mice an' men (line 39) [3]" often go awry. This fact combined with the inherent uncertainties of any military operation have plagued military commanders throughout history. Clausewitz called this concept the "fog of war," with fog referring to the lack of clear information [16]. A good example of this concept is the battle of Gettysburg during the American Civil War. The Confederate commander, General Lee did not have any clear information about the enemy because his calvary, the "eyes and ears" of the army, was "joyriding" around the Union army and failed to provide any intelligence. The resulting uncertainties of enemy troop movements, strength, and fortifications, increased the "fog of war" for Lee. His chance to flank and rout the Union army at the Round Tops was lost due to uncertainties of the tactical situation. If Lee had a clearer picture of the battlefield, the carnage of Pickett's charge and the subsequent Confederate loss may have been avoided. The Joint Force Commanders (JFC) of today's military operations face the same uncertainties as Lee. Technology and operational doctrine have not yet conquered the difficulties of "fog of war."

1.1 Background

On today's modern battlefield, the Joint Force Air Component Commander (JFACC) is charged with the responsibility to exploit aerospace capabilities with a cohesive Air Operations Plan (AOP) using a responsive and integrated control system. The Aerospace Operations Center (AOC) is the centralized command center for the JFACC where planning, direction, control, and coordination of aerospace operations occur. An Air Tasking Order (ATO) is used to execute air operations on a daily basis in support of AOP and JFC objectives. After ATO execution, Mission Reports (MISREP) provide initial Battle Damage Assessment (BDA). [5]

Naturally, the JFACC would like to know, in real-time if possible, how well AOP and JFC objectives are being met. This said, modern warfare objectives are no longer measured by body counts and destruction of enemy war materiel, but with measures of effect derived from the tenets of Effects Based Operations (EBO). This adds another layer of uncertainty to the analysts in the AOC who are tasked with providing insight to the JFACC.

Finally, please note that throughout this thesis, AOC, ATO, and AOP will be used in general terms only. An AOC will become a joint AOC, or JAOC, when air operations are multi-service, or a combined AOC, or CAOC, when air operations are coalition. Similarly, the ATO and AOP will become a joint ATO and joint AOP respectively when air operations are multi-service.

1.2 Problem Statement

One of the most important insights that analysts in the AOC can provide the JFACC is whether or not their aerospace missions are successful. However, how does one measure success? A direct hit on a target may not meet AOP and JFC objectives. For example, a building might have been abandoned by the enemy a week earlier and its destruction achieves no effect. Also, an incorrect or misidentified target might be struck with absolutely no effect. Even with accurate BDA, there is still a high level of uncertainty involved in determining mission success.

1.3 Research Objectives

The objective of this research is to investigate the effects of uncertainty in analyzing real-world aerospace mission data from OIF. This is accomplished by first identifying sources of uncertainty in the data and then using multivariate analysis techniques, at varying levels of uncertainty, to measure the impact of uncertainty in classification.

1.4 Research Scope

The real-world data is limited to ATO and MISREP databases compiled in the AOC during OIF and made available for use by the Warrior Prep Center located at Ramstein Air Base, Germany. Three consecutive ATO days worth of data near the middle of OIF were chosen for further analysis with multivariate techniques. The three techniques are as follows: SNR Saliency, FFNN supervised training, and GRNN supervised training. Finally, measures of performance are collected for each technique at two levels of uncertainty:

1. A high level of uncertainty in determining mission success.
2. A low level of uncertainty in determining if a mission was pre-planned or alert generated.

Determining whether or not a mission was pre-planned or alert generated has no real-world application. However, this case has a low level of uncertainty and serves to validate the analysis techniques in this thesis for future use with ATO and MISREP databases. The real-world application is determining whether or not a mission is a success. This case has a high level of uncertainty and uses the validated analysis techniques from the low uncertainty case to lay the groundwork for determining mission success.

1.5 Research Methodology

Research methods used in this thesis are data preparation, SNR saliency, FFNN supervised training, and GRNN supervised training. Data preparation involves data reduction and the "fusing" of ATO and MISREP databases into one usable database.

This database will contain selected statistical performance measures or features. The features with qualitative entries will be coded with numeric values. It is here that uncertainties in the data are identified. The one usable database is then used for multivariate analysis. SNR saliency uses FFNN supervised training with "noise" introduced as an extra input feature. The purpose of this is to reduce the number of input features without losing significant classification performance. Then, FFNN and GRNN supervised training is used to predict mission success for a high level of uncertainty case and mission generation for a low level of uncertainty case.

1.6 Thesis Organization

This thesis is organized into five chapters. This chapter includes a brief look at the background, problem, objective, scope, and methodology of the effects of uncertainty on real-world aerospace mission data. The second chapter contains a detailed literature review on topics relevant to the uncertainty in generating, collecting, processing, and analyzing aerospace mission data. These include the typical structure of the AOC, the ATO planning cycle, software systems and databases used in the AOC, EBO, and analysis operations in the AOC during OIF. Chapter three describes the methodology used for data preparation and the three techniques used for multivariate data analysis. Chapter four contains the numerical results from the data analysis and the last chapter includes conclusions and recommendations for future research.

2. Literature Review

This chapter provides a comprehensive review of literature relevant to this research. We begin by taking a look at the typical structure of the AOC, followed by a discussion of the six phases of the ATO planning cycle. Next, we review the numerous software systems and databases used in the AOC. Also, we will take a look at the philosophical concepts of EBO. Finally, we briefly discuss analysis operations in the AOC during OIF.

2.1 AOC Structure

Air Force Doctrine Document (AFDD) 2 provides guidance for the structure and primary functions of a typical AOC. The AOC is composed of four core divisions: Strategy, Combat Plans, Combat Operations, and Mobility (see Figure 2.1). We will take a closer look at the divisions that have a direct impact on the ATO and MISREP data used in this thesis, all except Mobility. [5]

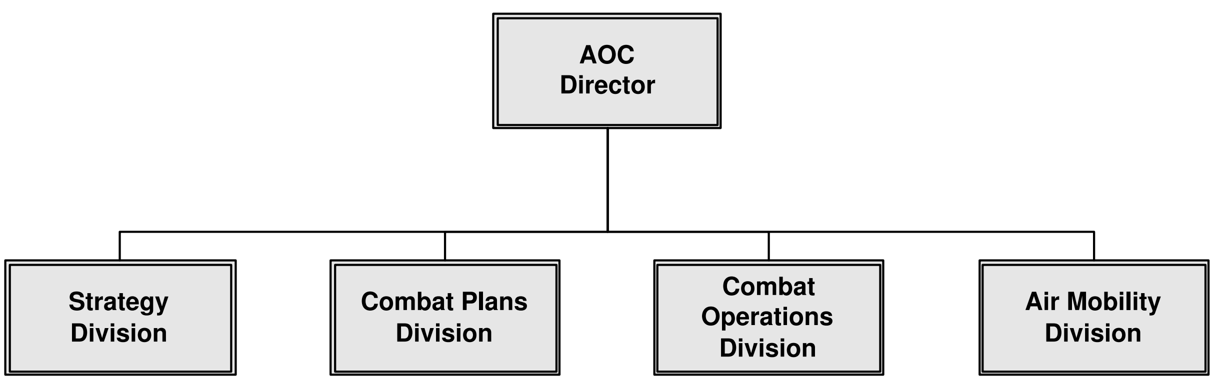

Figure 2.1 Typical AOC Structure

2.1.1 Strategy Division

The Strategy Division's (SD) primary function is long-term aerospace operations planning across the theater of operations. The SD is composed of two core teams, Strategy Plans and Operational Assessment, with their respective functionally-oriented specialty and support teams (see Figure 2.2). [5]

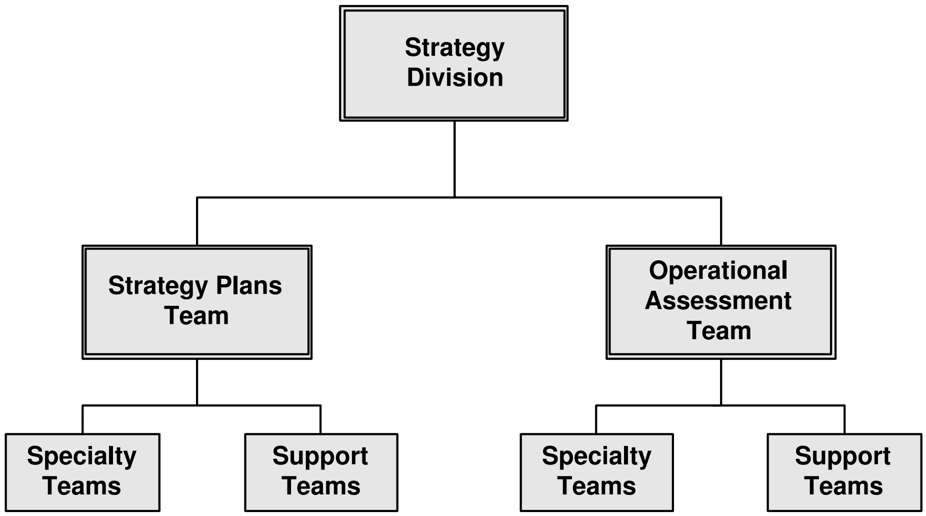

Figure 2.2 AOC Strategy Division

The Strategy Plans Team is responsible for developing the JFACC Estimate of the Situation, proposed aerospace strategy, and AOP to ensure continuity from strategy to execution. This team is staffed with representatives of the aerospace forces and capabilities under JFACC control. [5]

The Combat Assessment Team is responsible for assessing aerospace operations at the operational or campaign level. Products at each phase of air operations are evaluated to determine whether desired aerospace objectives are being met. Areas of interest include BDA, munitions effectiveness assessment, re-strike options, and overall mission assessment [5]. This is the team where analysts in the AOC are supposed to be placed according to the Unit Manning Document (UMD) [11]. We will see in the last section of this chapter that this changed during the course of OIF.

2.1.2 Combat Plans Division

The Combat Plans Division's (CPD) primary function is near-term aerospace operations planning and force allocation in accordance with JFACC guidance. The

CPD is composed of two core teams, Master Air Attack Plan (MAAP) and ATO/Aerospace Control Order (ACO) Production, with their respective functionally-oriented specialty and support teams (see Figure 2.3). [5]

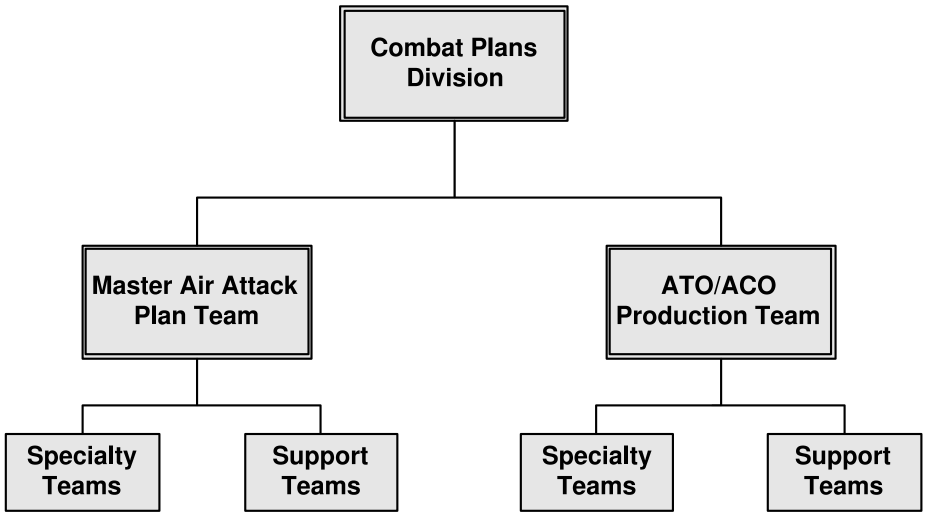

Figure 2.3 AOC Combat Plans Division

The MAAP Team is responsible for developing a daily employment plan for available assets, the MAAP, using the Joint Integrated Prioritized Target List (JIPTL) and guidance from the JFC and JFACC. The MAAP contains optimized target, platform, weapon, and timing information for missions to be included in the ATO. [5]

The ATO/ACO Production Team is responsible for producing and distributing a timely and executable ATO, ACO, and Special Instructions (SPINS) on a daily basis. This team is staffed by fully-qualified operational experts from each aircraft or weapon system which may be tasked by the JFACC. Planning is done by two specialty teams, fighter/bomber and support, and two support teams, airspace management and communication. [5]

2.1.3 Combat Operations Division

The Combat Operations Division's (COD) primary function is executing the current ATO through constant monitoring of all aerospace missions. The COD is composed of two core teams, Offensive Operations and Defensive Operations, with their respective functionally-oriented specialty and support teams (see Figure 2.4). [5]

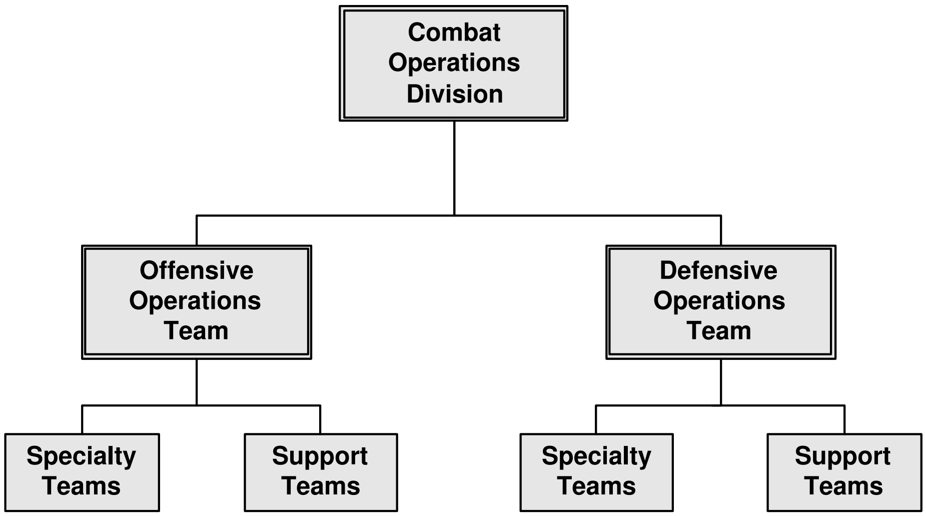

Figure 2.4 AOC Combat Operations Division

The Offensive Operations Team is responsible for monitoring current offensive aerospace operations, dynamic mission requirements, and resource status. Time critical changes to the ATO based on invalid targets, new high priority targets, and enemy actions are coordinated and published by this team. [5]

The Defensive Operations Team is responsible for monitoring current defensive operations and the enemy air threat. Team members assess fighter and air defense missile assets and coordinate to provide sufficient air defense assets. [5]

2.2 ATO Planning Cycle

Joint Publication 3-30 provides guidance for the execution of the ATO planning cycle. The cycle was developed to provide the JFC an effective and efficient process to plan and execute air operations. The cycle is a repetitive process where three ATOs are in existence at any given time. The first ATO is being executed as today's plan, the second is being produced as tomorrow's plan, and the third is being planned as the day after tomorrow's plan. Each ATO typically covers a 24-hour period of air operations. The six phases (see Figure 2.5) of the ATO cycle are as follows: JFC/Component Coordination, Target Development, Weaponeering/Allocation, ATO Development, Force Execution, and Combat Assessment. [7]

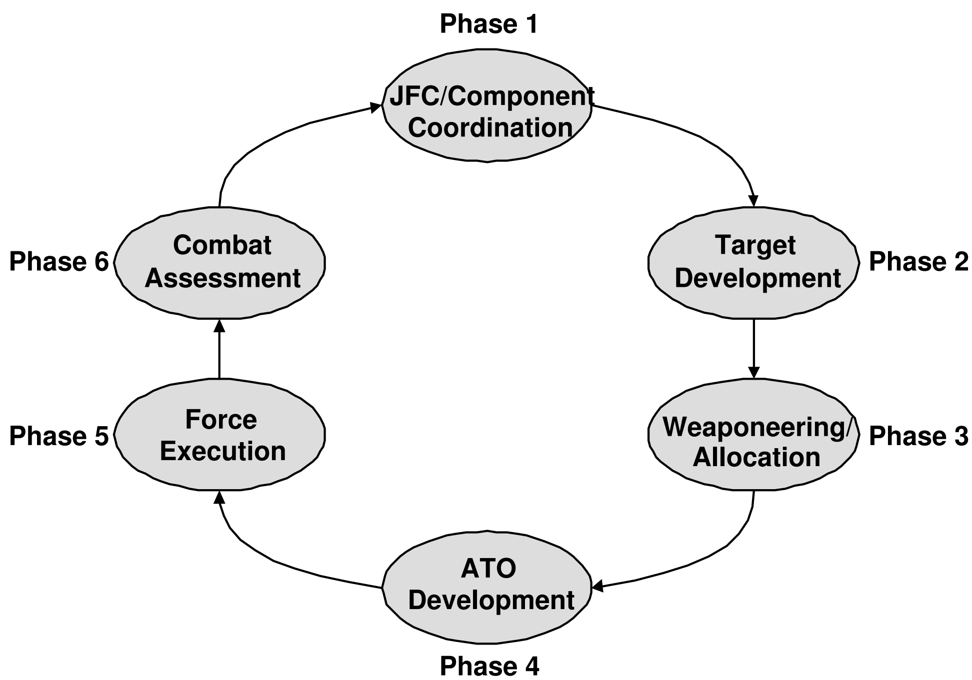

Figure 2.5 ATO Planning Cycle

The JFC coordinates with component commanders to solicit support requirements, recommendations, and assess the results of the war fighting effort. This helps the SD to identify target priorities, planning guidance, procedures, rules of engagement (ROE), and other planning measures [7]. Also, the air effort is apportioned by priority or percentage of effort to geographic areas, mission-type, and/or significant categories for the campaign [6]. As a result of this phase, JFC guidance and objectives are defined for target development and air apportionment [7].

The JFC guidance and objectives from Phase 1 are used to focus target development. The CPD nominates targets that meet these requirements. As a result of this phase, a prioritized list of targets, the JIPTL is developed. [7]

The JIPTL from Phase 2 is used to develop weapon and sortie allocations for each target to be included in the MAAP. Weapon assignment is made after a detailed look at expected results for lethal and non-lethal weapons against each prioritized target. Sortie allocation is based on the total available sorties measured by aircraft type, capability, and the mission requirements. As a result of this phase, the MAAP is available as a foundation for the ATO. [7]

The MAAP from Phase 3, in addition to JFC and JFACC guidance, target worksheets, and component requirements, is used to complete work on the ATO. The ATO, together with SPINS and the ACO, provides enough operational and tactical detail for components to plan and execute all of the missions that are tasked by the ATO. A final ATO will integrate airspace, air defense, tanker plan, comm plan, electronic warfare, SPINS, ROE, and suppression of enemy air defense (SEAD) for each specific mission. As a result of this phase, the ATO is ready for execution under the direction of the JFACC. [7]

The final ATO from Phase 4 is executed in a 24-hour period by the COD of the AOC. The COD has offensive, defensive, and specialty operations teams to support the execution of the ATO. The teams monitor execution of the ATO, integrate information and resources, analyze and prioritize options, provide options and recommendations to the JFACC to redirect assets, and direct any changes to the ATO. As a result of this phase, combat operations results are available for assessment. [7]

Combat operations results from Phase 5 are used to assess the effectiveness in achieving JFC objectives. However, combat assessment occurs at the other phases of the ATO cycle using formal and informal processes. This allows the JFC/JFACC staff to continuously evaluate the results of the current campaign and adjust objectives and strategies dynamically. As a result of this phase, the ATO cycle is complete and combat assessment recommendations are used in development of subsequent ATOs. [7]

2.3 AOC Software Systems and Databases

The AOC is populated with many software systems and databases, each having a critical contribution to the overall war effort. Information passes from these systems and databases in a sequential fashion as the ATO planning cycle is executed. However, each phase of the ATO planning cycle is dependent on data from a previous process before they can begin. The following discussion looks at those software systems and databases typically found in the AOC and their interrelationships. [12]

First, during the JFC/Component Coordination phase, SD uses Microsoft Excel to develop a prioritized list of Tactical Tasks. When complete, the Excel spreadsheet is passed to the Interim Targeting Solution (ITS) software system. ITS uses current ATO Day, Objective Label, and Priority from the Excel spreadsheet to import data needed for future planning into its own proprietary Microsoft SQL 2000 database. [12]

Second, during the Target Development phase, nominated targets and objectives within the ITS database are used by the Joint Guidance Apportionment and Targeting (JGAT) group to develop the JIPTL. The JIPTL is stored in both its internal database and on the Theater Battle Management Core System (TBMCS) software system Intelligence Shared Data Server (ISDS) database. [12]

Next, during the Weaponeering/Allocation phase, the CPD uses the MAAP Toolkit software system to create the MAAP. The MAAP Toolkit uses the JIPTL and other information from the TBMCS ISDS database. Outputs from the MAAP toolkit, Target Planning Worksheets (TPW), are stored in its own proprietary Oracle database and the Theater Air Planner (TAP) database. [12]

Additionally, during the ATO Development phase, the CPD uses the TAP software system to build products for ATO production. TAP builds these products using TPWs from its database and JIPTLs from the TBMCS ISDS database and stores them in the TBMCS Air Operations (AO) database. The TBMCS AO database is used to produce the actual ATOs. [12]

Finally, during the Force Execution, the COD uses the Automated Deep Operations Coordination System (ADOCS) software system to coordinate time sensitive targeting, correlate real-time feeds, and display real time battlefield information. MISREPS from the operational units are stored in Microsoft Excel spreadsheets. [13]

As you can see, there are at least seven disparate databases used to store data from the first five phases of the ATO planning cycle: Microsoft Excel spreadsheets, ITS Microsoft SQL 2000, JIPTL, TBMCS ISDS, MAAP Toolkit, TPS, and TBMCS AO. The information in these databases is what the analysts in the AOC need to conduct the last phase, Combat Assessment. One challenge is getting access to each of these databases and merging them into a correlated and aggregated view. Another challenge is dealing with the inconsistency of data formatting in each database (e.g. one database may use f-16 while another uses f16). Both of these challenges create uncertainty in performing data analysis. [12]

2.4 Effects-Based Operations

In addition to uncertainty in data, a high level of uncertainty lies in measuring the operational success of EBO. To better understand EBO, we investigate Myers and his six basic tenets. The focus of EBO is to persuade enemy leaders that they have lost the conflict and cannot achieve their objectives. EBO attempts to accomplish this without large-scale engagements and widespread destruction. His six basic tenets of EBO are attrition only when necessary, people not things, systems approach, linkage, universal application, and knowledge. [14]

Conflicts of the past were attrition-based, where success was measured in body counts and ground gained. The winner was often determined by which side was willing to commit more lives and material to the conflict. Enemy leaders of today have little concern for loss of life or material until they are directly threatened. Ultimately, the focus of EBO is the enemy leader. [14]

EBO is based on the fact that war is waged between people and not machines. There are characteristics of the human psyche that can be manipulated directly and indirectly. Accordingly, EBO focuses on psychological and sociological versus physical effects. [14]

We must understand that the enemy is a complex system of political, cultural, technological, military, and economic components. These components have key nodes and links which can be manipulated by EBO. As a result, we can achieve political, economic, and social effects not possible by military force alone. [14]

Enemy targets and objectives are attacked because they produce effects linked to strategic and operational objectives. EBO has the effect of linking geopolitical, economic, communications, and military objectives into a wide variety of targets. These objectives can be translated into desired effects before any military operations begin. [14]

The power of EBO is that it is not limited to military operations. It has a universal application across the spectrum of government operations because of its focus on human behavior. Therefore, EBO can handle conventional military operations, guerilla warfare, diplomatic action, economic crisis, etc. [14]

Knowledge is the key to EBO. We must rapidly obtain and disseminate information in order for our desired effects to be successful. If we can leverage our technology to obtain information on our enemy, we can use EBO to create lasting effects with taking unnecessary risks with our people and resources. [14]

2.5 The AOC during OIF

We have discussed all the theory behind the AOC, ATOs, data, and EBO, now we take a look at a first-hand account of analysis operations in the AOC during OIF. When OIF began, analysts were placed in the Operational Assessment team of the Strategy Division, according to UMD, and also in the JIPTL Tracking Cell. However, leadership soon realized that an independent analysis cell was needed. Therefore, during OIF a new Analysis and Assessment Division was stood up. The result from lessons learned from OIF may lead to a different role and placement for analysts in future AOCs. The major software tools and systems used were Microsoft Excel and Access, TBMCS, ADOCS, MAAP Toolkit, and ITS. SAS was also used but not approved for use in the AOC. A major source of uncertainty in the ATO planning cycle was the lack of BDA. The entire strategy was based on a Strategy to Task development and was effects based. However, measuring how well effects were being accomplished proved a difficult task. [11]

2.6 Summary

The literature review provides important background on topics relevant to this research. The typical structure of the AOC, the ATO planning cycle, and the numerous software systems and databases used in the AOC all influence how data is generated, collected, processed, and analyzed. These influences create varying levels of uncertainty in the aerospace mission data used in this thesis. Ideally, we would like to measure results at the operational level based on EBO objectives. However, this adds additional uncertainty to the data. All of these sources of uncertainty were present in the AOC during OIF.

3. Methodology

In this chapter, the methodology used for data preparation and multivariate analysis is described. The data preparation section reviews the process used to create a usable database from the "raw" ATO and MISREP data. The multivariate section takes a look at common practices, architecture, and mathematical models used for each analysis technique.

3.1 Data Preparation

The source data for the ATO and MISREP databases used in this thesis was generated and collected in the AOC during OIF. This data provides only a tactical level look at aerospace missions, since the data contains no connection between mission results and EBO objectives. The data has been processed from its "raw" form into a set of Microsoft Access databases. Thus, some of the uncertainties that existed in the original ATOs and MISREP have been compensated for. However, there still exists plenty of uncertainty in the data. The data was made available for use by the Warrior Prep Center located at Ramstein Air Base, Germany.

The objective of data preparation is to reduce the Microsoft Access databases to a manageable number of data exemplars. The first step was to reduce the data to three consecutive ATO days. These days were chosen near the middle of OIF in order to reduce the bias created by operations tempo at the beginning and end of the conflict. Next, 15 features were chosen and the set of Microsoft Access databases "fused" into one usable Microsoft Excel spreadsheet. The challenges of this task were deciding what features to keep, coding the qualitative data, and dealing with uncertainty in the data. There are a total of 638 data exemplars in the final useable database (see Appendix B for a database sample).

3.1.1 Features

A detailed explanation of each feature, the coding scheme used, and uncertainty involved is as follows:

MSN Type: Denotes the type of mission flown. Only those missions in which ordnance was dropped or weapons fired were kept in the database. Categories include, but are not limited to, SEAD and close air support (CAS). This feature is coded 0 through 7. There is a low level of uncertainty in this data.

 # AC: Denotes the number of aircraft in the mission. There is no coding and a low level of uncertainty in this data.

MSN Time: Denotes the total mission time calculated from the take-off of the first aircraft in the mission to the landing time of the last aircraft. There is no coding and a low level of uncertainty in this data.

TOS: Denotes the total time on station spent in the target area. There is no coding and a low level of uncertainty in this data.

AC Type: Denotes the type of aircraft used in the mission. The aircraft include Air Force, Navy, Marine, and British variants. This feature is coded 0 through 19. There is a low level of uncertainty in this data.

 # ESM: Denotes the number of electronic surveillance measures employed against the aircraft during the mission. There is no coding and a low level of uncertainty in this data.

 # SAF: Denotes the number of surface-to-air firings against the aircraft during the mission. There is no coding and a low level of uncertainty in this data.

SAF Type: Denotes the type of surface-to-air fire encountered. Categories include, but are not limited to, surface-to-air missiles (SAM) and anti-aircraft artillery (AAA). This feature is coded 0 through 4. There is a moderate level of uncertainty in this data due to the various text-based inputs in the "raw" data (e.g. "AAA"", "30mm", and "guns" are entries that may or may not be grouped together as AAA).

STGS: Denotes the number of unexpected enemy sightings during the mission. There is no coding and a low level of uncertainty in this data.

 # ORD: Denotes the number of ordnance dropped and/or fired during the mission. There is no coding and a low level of uncertainty in this data.

ORD Type: Denotes the type of ordnance dropped and/or fired during the mission. Categories include, but are not limited to, AGM-164A, GBU-31, and 30mm. This feature is coded 0-34 for each individual target and an average value is used at the mission level. There is a low level of uncertainty in this data.

ALT: Denotes the aircraft altitude when the ordnance was dropped and/or fired during the mission. There is no coding and an average value is used at the mission level. There is a low level of uncertainty in this data.

TGT Type: Denotes the type of target tasked by the mission. Categories include, but are not limited to, tanks, buildings, and bridges. This feature is coded 0 through 11 for each individual target and an average value is used at the mission level. There is a moderate level of uncertainty in this data due to the various text-based inputs in the "raw" data (e.g. "tank", "T-72", and "armor"" are entries that may or may not be grouped together as tanks).

Alert: Denotes whether a mission was pre-planned or generated from air or ground alert aircraft. This feature is coded 0 for a pre-planned mission and 1 for an alert mission. There is a low level of uncertainty in this data.

Effect: Denotes a measure of performance for mission success. This measure is based on the average individual effect for all targets attacked during the mission (see Table 3.1 for sample calculation). The following heuristic was used to code the effect for each individual target:

1. The weapon hit the target and went high order
(the primary charge detonated): coded 1.
2. The weapon missed the target or hung: coded 0.

3. The weapon has uncertain results: coded .5.

Table 3.1 Sample Effect Calculation

Target	Effect
# 1:	.5
# 2:	0
# 3:	1
Mission Effect:	$(.5+0+1)/3 = .5$

There is a high level of uncertainty in this data due to the various text-based inputs in the "raw" data (e.g. "success", "successful weapon release", and "weapon hit target and went high order" are entries that may or may not be grouped together and coded 1).

3.1.2 Data Cases

Now that the raw data is reduced into one usable database and coded, two main data cases are chosen for multivariate analysis. The first case uses Alert as an output feature, leaving 14 input features. This case has a low level of uncertainty for the output feature with all levels of uncertainty for the input features. The second case uses Effect as an output feature, leaving 14 input features. This case has a high level of uncertainty for the output feature with moderate and low levels of uncertainty for the input features. Overall, the choice of cases serves to compare and contrast the effects of uncertainty on the data.

The objective of the Alert case is to predict whether a mission was pre-planned or alert generated. This will be accomplished by using multivariate analysis classification techniques. Some additional coding (see Table 3.2) and data reduction is needed before starting analysis. Targets (output feature identifiers) are coded with a {.1,.9} values instead of typical {0,1} unipolar values since the infinite nature of unipolar sigmoids (used in the FFNN) near {0,1} [10]. Using targets between {.1,.9} produces finite sigmoid outputs. Figure 3.1 shows how the unipolar sigmoid approaches 0 at $-\infty$ and 1 at ∞. The 638 data exemplars are reduced down to 455 since only 3 of the 8 mission types have both pre-planned and alert generated missions (see Table 3.3 for exemplar breakdown).

Table 3.2 Alert Case Targets

Class	Target 1	Target 2
Alert	.9	.1
Pre-Planned	.1	.9

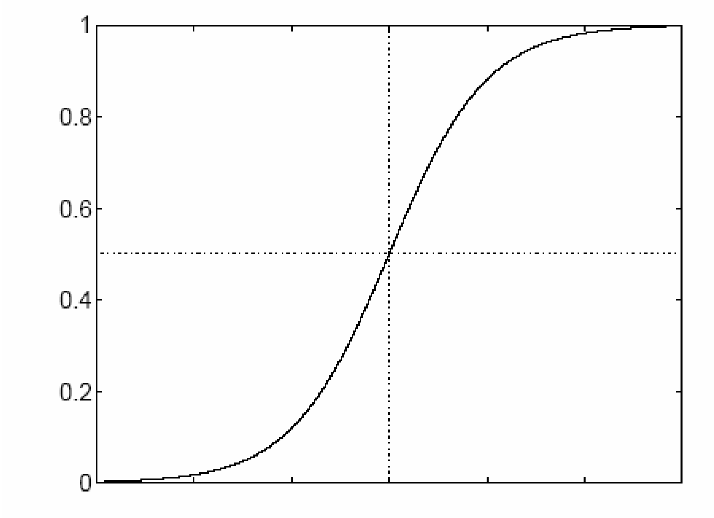

Figure 3.1 Unipolar Sigmoid

Table 3.3 Alert Case Exemplar Breakdown

Class	# Exemplars	% Total
Alert	208	45.7
Pre-Planned	247	54.3

The objective of the Effect case is to predict whether or not a mission was a success. This will be accomplished by using multivariate analysis classification techniques. Some additional coding (see Table 3.4) is needed before starting analysis. Targets are coded with {.1,.9} values, with the same reasoning as the Alert case. The problem here is that the Effect feature is coded at values ranging from 0 to 1. A threshold value must be selected to separate mission success from failure. Any Effect value greater than or equal to the threshold is coded a success, while any Effect value less than the threshold is coded a failure. Three different thresholds were chosen for investigation: .85, .75, and .65. (see Table 3.5 for exemplar breakdown). The thresholds were chosen to give a broad range of mission success from the low 50% range to the low 70% range. In reality, one would expect that the mission success rate would be greater than 90% for US aerospace forces. However, this was not practical due to adverse effects on measures of performance (see Chapter 4 for details). The three sub-cases will be referred to as Effect .85, Effect .75, and Effect .65 throughout the rest of this thesis.

Table 3.4 Effect Case Targets

Class	Target 1	Target 2
Success	.9	.1
Failure	.1	.9

Table 3.5 Effect Case Exemplar Breakdown

Threshold	Class	# Exemplars	% Total
.85	Success	328	51.4
.85	Failure	310	48.6
.75	Success	422	66.1
.75	Failure	216	33.9
.65	Success	458	71.8
.65	Failure	180	28.2

3.2 Multivariate Analysis Techniques

The objective of multivariate analysis is to take each of the four cases and measure the impact of uncertainty in predicting either mission success or the source of mission generation. First, SNR saliency is used to reduce the number of input features. Finally, FFNN and GRNN classifiers are used on the reduced feature data for output feature prediction. A MATLAB® m-file was created to execute each of these techniques (see Appendix C).

3.2.1 Common Practices

There are several procedural and statistical practices common to all three multivariate techniques:

1. A total of 10 training runs are performed at each design point. A design point is defined as a specific case (Alert, Effect .85, Effect .75, or Effect .65) executed with a specific multivariate analysis technique (SNR saliency, FFNN, or GRNN).

2. Exemplars from the data case are randomly split into training and validation sets. The split is accomplished by adding a random feature [Uniform(0,1)] to each exemplar and then sorting all of the exemplars in ascending order based on random feature values. The first 80% of exemplars are placed in the training set and the remaining 20% in the validation set. This random data split is repeated before each training run.

3. Data from each case is standardized so that all features have a zero mean and unit variance [2]. Standardization of the data ($\mathbf{X}$) is obtained using Equation 3.1, where $\mathbf{X}_s$ is a matrix of standardized data, $\mathbf{X}_d$ is a matrix of mean corrected data (see Equation 3.2), and $\mathbf{D}$ (see Equation 3.3) is a diagonal matrix composed of variances (see Equation 3.4) for each feature in $\mathbf{X}$ (also the diagonal of the covariance matrix of $\mathbf{X}$) [8].

$$\mathbf{X}_s = \mathbf{X}_d \cdot \mathbf{D}^{-1/2} \tag{3.1}$$

$$\mathbf{X}_d = \mathbf{1} \cdot \overline{\mathbf{x}}' \tag{3.2}$$

$$\mathbf{D} = \begin{bmatrix} s_1^2 & 0 & \dots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \dots & \dots & s_n^2 \end{bmatrix} \tag{3.3}$$

$$s^2 = \frac{1}{n-1} \mathbf{x}_d' \cdot \mathbf{x}_d \tag{3.4}$$

4. Rotation, with Principal Component Analysis (PCA) component coefficients, is used to transform data from each case into principal components of common factors for potentially improved classification. Data ($\mathbf{X}$) is rotated using Equation 3.5, where $\mathbf{X}_r$ is a matrix of rotated data and $\mathbf{L}$ is a matrix of PCA component coefficients. The component coefficient matrix ($\mathbf{L}$) is calculated (see Equation 3.6)

from a matrix of eigenvectors ($\mathbf{A}$) and a diagonal matrix of eigenvalues ($\mathbf{\Lambda}$) from the sample correlation matrix ($\mathbf{R}$). [8]

$$\mathbf{X}_r = \mathbf{X} \cdot \mathbf{L} \tag{3.5}$$

$$\mathbf{L} = \mathbf{A} \cdot \mathbf{\Lambda}^{1/2} \tag{3.6}$$

The sample correlation matrix ($\mathbf{R}$) is calculated (see Equation 3.7) using the sample covariance matrix ($\mathbf{C}$) and the diagonal matrix of the sample covariance matrix ($\mathbf{D}$). Finally, the sample covariance matrix ($\mathbf{C}$) is calculated (see Equation 3.8) using a matrix of data ($\mathbf{X}$), the mean vector of that data ($\overline{\mathbf{x}}$), and the number of exemplars in the data (n). [8]

$$\mathbf{R} = \mathbf{D}^{-1/2} \cdot \mathbf{C} \cdot \mathbf{D}^{-1/2} \tag{3.7}$$

$$\mathbf{C} = \frac{(\mathbf{X}' \cdot \mathbf{X}) - n \cdot (\overline{\mathbf{x}} \cdot \overline{\mathbf{x}}')}{n-1} \tag{3.8}$$

5. Statistical averages ($\overline{y}$) for measures of performance (y) are calculated by removing the highest and lowest values from the 10 runs and taking the average of the remaining 8 values (see Equation 3.9) [17]. This is done to eliminate the impact to the average for a result that is much greater or less (an extreme outlier on the error surface) than the other values, since there are only 10 runs of data at each design point.

$$\overline{y} = \frac{\sum_{i=1}^{10} y_i - \max_{i=1}^{10} y_i - \min_{i=1}^{10} y_i}{8} \tag{3.9}$$

6. Sample variances (s^2) for measures of performance (y) are calculated by taking the variance of the 8 values used in the average calculation (see Equation 3.10) [17].

$$s^2 = \frac{\sum_{i=1}^{8} (y_i - \overline{y})^2}{7} \tag{3.10}$$

7. The hypothesis test used to determine if an input feature is significant (a statistical difference in average between classes) is as follows [17]:

$H_0 : \mu_1 = \mu_2$

$H_a : \mu_1 > \mu_2$

Rejection Region: $Z > Z_\alpha$

The test statistic (Z) is calculated with Equation 3.11, the critical value (Z_α) has to be looked up in a Z distribution table, and μ_1 and μ_2 are approximated by $\overline{y}_1$ and $\overline{y}_2$ respectively. Using an α of .05, $Z_\alpha = 1.645$. The number of exemplars in class 1 is n_1 and the number of exemplars in class 2 is n_2. [17]

$$Z = \frac{\overline{y}_1 - \overline{y}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \tag{3.11}$$

8. The hypothesis test used to determine if a measure of performance is significant (a statistical difference in average between two results) is as follows [17]:

$H_0 : \mu_1 = \mu_2$

$H_a : \mu_1 > \mu_2$

Rejection Region: $t > t_\alpha$

The test statistic (t) is calculated with Equation 3.12, the critical value (t_α) has to be looked up in a t distribution table, and μ_1 and μ_2 are approximated by $\overline{y}_1$ and $\overline{y}_2$ respectively. Using an α of .05 with 14 degrees of freedom, based on 8 runs each for each average, $t_\alpha = 1.761$. This test assumes that $\overline{y}_1$ and $\overline{y}_2$ are independent samples from normal distributions with equal population variances ($\sigma_1^2 = \sigma_2^2$). [17]

$$t = \frac{\overline{y}_1 - \overline{y}_2}{s_p \cdot \sqrt{\frac{1}{4}}}, \text{ where } \; s_p = \sqrt{\frac{7 \cdot s_1^2 + 7 \cdot s_2^2}{14}} \tag{3.12}$$

3.2.2 Signal-to-Noise Ratio Saliency

The objective of SNR saliency analysis is input feature reduction. The goal is to remove those features which add no value to (or possibly reduce) supervised training classification. The SNR saliency method, modified from Bauer, Alsing, and Greene [2], is outlined below:

1. Add a noise [Uniform(0,1)] feature (x_n) to the input features set for each exemplar [2].

2. Randomize the training and validation sets.

3. Standardize the data.

4. Train the training set with a FFNN (see Figure 3.2 for FFNN architecture [10]). Refer to the FFNN section in this chapter for training and architecture details.

5. Calculate SNR saliency measures of performance using Equation 3.13, where SNR_i is the SNR saliency measure for input feature i, w_{ij} is the hidden layer weight

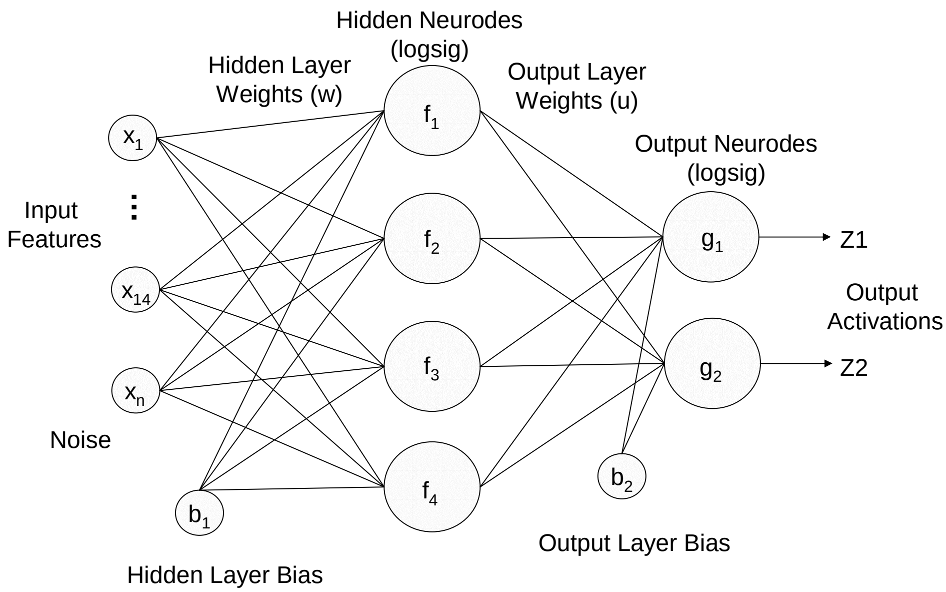

Figure 3.2 SNR Saliency FFNN Architecture

from input feature i to hidden neurode j, and w_{nj} is the weight from the noise input feature to hidden neurode j. An input feature is salient if its SNR is greater than 0. [2]

$$SNR_i = 10 \cdot \log_{base10} \left[\frac{\sum_{j=1}^{4} (w_{ij})^2}{\sum_{j=1}^{4} (w_{nj})^2} \right] \tag{3.13}$$

6. Simulate validation set output activations by using the hidden layer (w) weights, output layer (u) weights, and biases (b) from training. The output activations ($Z1$ and $Z2$) are posterior probabilities, approximated between {.1,.9} of class membership. The highest output activation is the predicted class. This is the result of using a .5 threshold for class membership (i.e. if $Z1 \geq .5$, then class 1, otherwise class 2).

7. All of the validation set exemplars (N) are evaluated to determine a classification accuracy (CA) measure of performance (see Equation 3.14). A true positive (TP)

occurs when target class 1 is predicted as class 1 and a true negative (TN) occurs when target class 2 is predicted as class 2.

$$CA = \frac{\sum TP + \sum TN}{N} \tag{3.14}$$

8. Identify the feature with the lowest SNR saliency measure of performance and remove it from further training [2].

9. Repeat steps 2 through 8 until the only input feature left is noise [2].

Finally, a reduced input feature set, with the noise feature removed, is selected for FFNN and GRNN supervised training. This decision is based on two factors:

a. The reduced input feature set must be salient (all SNRs > 0).

b. Reduce the number of input features as low as possible without losing significant classification accuracy.

3.2.3 Feed Forward Neural Network

The objective of FFNN supervised training is to train the reduced input feature set for each data case and provide measures of performance that can be used for comparison of the results of uncertainty and GRNN supervised training. The FFNN supervised training method is outlined below:

1. Randomize the training and validation sets.

2. Standardize, rotate and standardize, or standardize and rotate the data. All three methods will be investigated for each data case.

3. Train the training set with a FFNN (see Figure 3.3 for FFNN architecture [10]). The MATLAB® **newff** function is used to create a FFNN with 4 log-sigmoid hidden neurodes and 2 log-sigmoid output neurodes. This architecture was based on previous experience with the database and a few test runs. Max training is set at

1500 epochs and mean squared error (MSE) goal at .01 for the Alert case and .05 for the Effect cases. These setting were determined after a few test runs with epochs varied from 500 to 5000 and MSE goal varied from .1 to .01. Finally, each training run consists of 10 iterations.

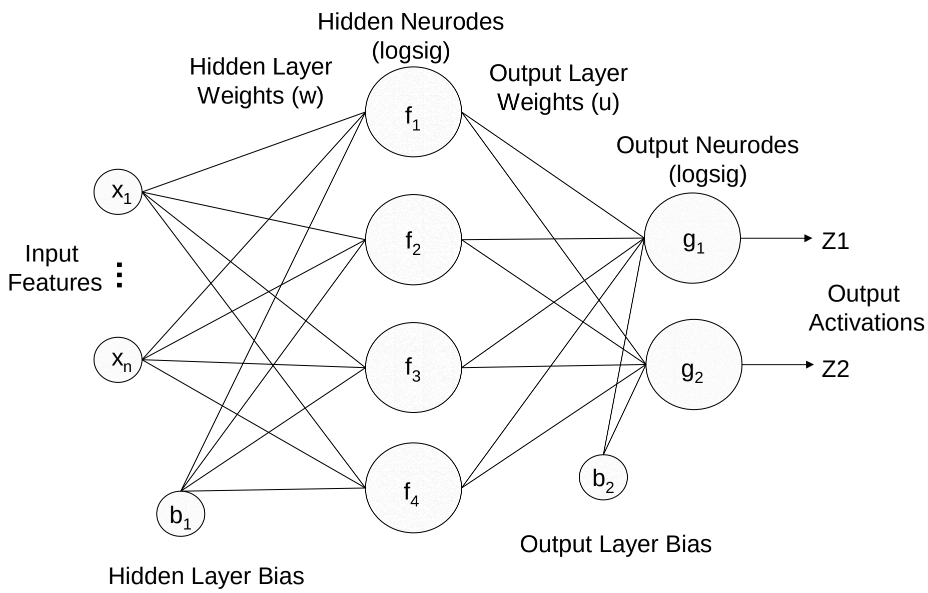

Figure 3.3 FFNN Architecture

The mathematical process for FFNN training, adapted from Looney [10], is as follows:

a. Hidden and output layer weights and biases are initialized at random (**newff** uses **initnw** to initialize weights and biases according to the Nguyen-Widrow initialization algorithm [4]).

b. Activations from hidden layer neurodes (f) are calculated with the log-sigmoid transfer function (see Equation 3.15), where s is a function of input features (x) and hidden layer weights (w), b_1 is hidden layer bias, and n is the number of input features [10].

$$f_i = \frac{1}{1+\exp^{-s+b_1}}, \text{ where } \; s = \sum_{j=1}^{n} w_{ji} \cdot x_j \tag{3.15}$$

c. Activations from output layer neurodes (g or Z) are calculated with the log-sigmoid transfer function (see Equation 3.16), where s is a function of hidden layer neurode activations (f) and output layer weights (u), b_2 is output layer bias, and m is the number of hidden layer neurodes [10].

$$g_i = \frac{1}{1+\exp^{-s+b_2}}, \text{ where } \; s = \sum_{j=1}^{m} u_{ji} \cdot f_j \tag{3.16}$$

d. Hidden and output layer weights are biases are updated (trained) each epoch using the MATLAB® **trainrp** batch mode gradient update function. This function uses a resilient backpropagation algorithm that is generally much faster than the standard steepest descent algorithm and eliminates harmful effects from having small slopes at the extreme ends of the log-sigmoid transfer function. Equation 3.17 is the gradient descent algorithm for backpropagation, where $\mathbf{x}_{k+1}$ is an updated vector of weights and biases, $\mathbf{x}_k$ is a vector of current weights and biases, $\mathbf{g}_k$ is the current gradient, and α_k is the learning rate. [4]

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \alpha_k \cdot \mathbf{g}_k \tag{3.17}$$

e. Training stops when the maximum number of epochs or the MSE goal is reached [10].

4. Simulate validation set output activations by using the hidden layer (w) weights, output layer (u) weights, and biases (b) from training. The output activations ($Z1$ and $Z2$) are posterior probabilities of class membership. The highest output activation is the predicted class. This is the result of using a .5 threshold for class membership (i.e. if $Z1 \geq .5$, then class 1, otherwise class 2).

5. All of the validation set exemplars (N) are evaluated to determine a classification accuracy (CA) measure of performance (same method as SNR saliency). Also, a measure of performance for the probability of a true positive [P(TP)] is calculated (see Equation 3.18), where n is the number of validation set exemplars with target class 1. Finally, a measure of performance for the probability of a true negative [P(TN)] is calculated (see Equation 3.19), where m is the number of validation set exemplars with target class 2.

$$P(TP) = \frac{\sum TP}{n} \tag{3.18}$$

$$P(TN) = \frac{\sum TN}{m} \tag{3.19}$$

6. A receiver operating characteristic (ROC) curve is plotted (see Figure 3.4) as adapted from Egan [9]. The curve plots the probability of false positive [P(FP)], target class 2 is predicted as class 1, versus the P(TP). Points on the ROC curve are generated for 101 thresholds between 0 and 1, incremented by .1. TP and FP determinations are made as follows:

If Z1>threshold, then classify as TP if actual class 1, else classify as FP.

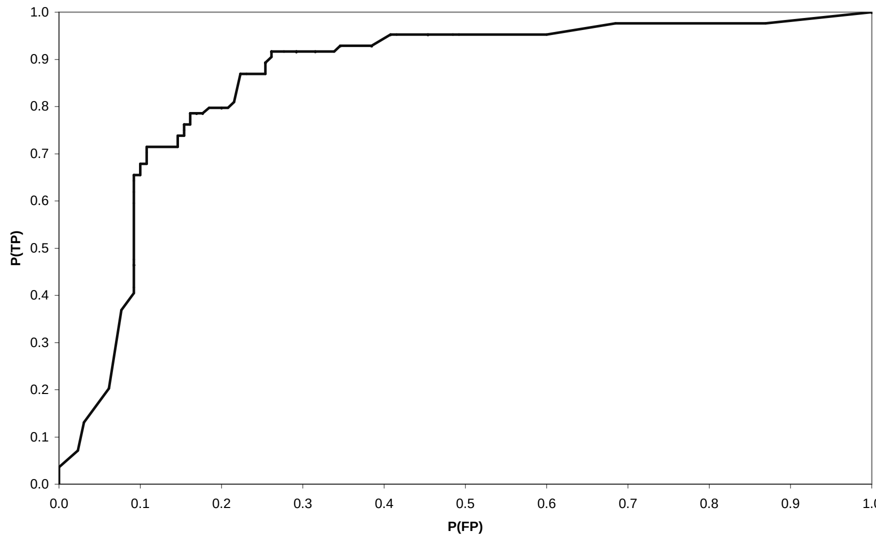

Figure 3.4 Sample ROC Curve

The P(TP) (same method as above) and the P(FP) (see Equation 3.20) are calculated , where m is the number of validation set exemplars with target class 2. It is important to note that since targets are coded with $\{.9,.1\}$, thresholds $>.9$ will always generate a ROC point at (0,0) and thresholds $<.1$ will always generate a ROC point at (1,1). This has no impact on the ROC curve.

$$P(FP) = \frac{\sum FP}{m} \tag{3.20}$$

7. Finally, an average metric distance (AMD) measure of performance is determined from the ROC curve as adapted from Alsing, Bauer, and Oxley [1]. AMD is a measure of the "robustness" of the classifier and is calculated (see Equation 3.21) using the average of the Euclidian distances between each ROC curve point and its corresponding point on the "chance" line, the diagonal line running between (0,0) and (1,1) on the ROC curve. $\mathbf{ROC}_i$ is the column vector of the ROC point for threshold i and $\mathbf{CH}_i$ is the column vector of the chance line point for threshold i.

$$AMD = \frac{\sum_{i=1}^{101}\sqrt{(\mathbf{ROC}_i - \mathbf{CH}_i)'(\mathbf{ROC}_i - \mathbf{CH}_i)}}{101} \tag{3.21}$$

3.2.4 General Regression Neural Network

The objective of GRNN supervised training is to train the reduced input feature set for each data case and provide measures of performance that can be used for comparison of the results of uncertainty and FFNN supervised training. The GRNN supervised training method is outlined below:

1. Randomize the training and validation sets.

2. Standardize, standardize and rotate, or rotate and standardize the data. All three methods will be investigated for each data case.

3. Train the training set with a GRNN (see Figure 3.5 for GRNN architecture [15]) using a spread parameter (σ) between .1 and 1, incremented by .1. This decision was based on previous experience with the database and a few test runs. The spread parameter with the best CA will be used for further analysis. The MATLAB® **newgrnn** function is used to create a GRNN with n (the number of training set exemplars) exponential hidden neurodes, 3 linear summation neurodes, and 2 conditional probability of class output neurodes.

The mathematical process for GRNN training (simulation of validation set exemplars), adapted from Specht [15], is as follows:

a. The input vector (v) weights are initialized using one training set exemplar for each input vector [15].

b. Activations from hidden layer neurodes (hidden neurodes a have input vectors (v) from training set target class 1 and hidden neurodes b have input vectors (v) from training set target class 2) are calculated for each validation set exemplar with an exponential (radial basis) function (see Equation 3.22), [15]. D_i^2 is the

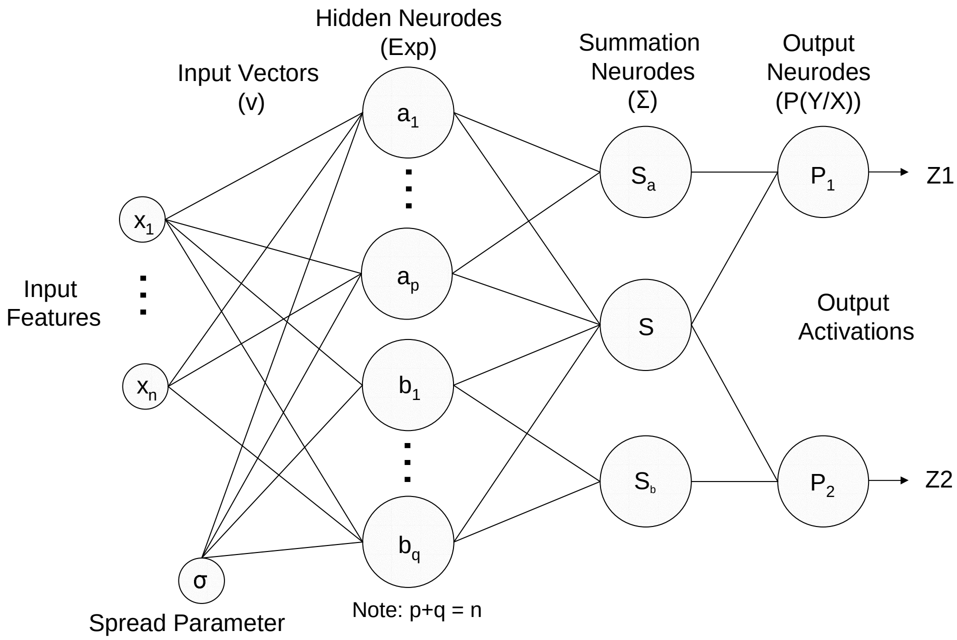

Figure 3.5 GRNN Architecture

squared Euclidian distance between a column vector of input features ($\mathbf{x}$) and a column vector of input vector weights ($\mathbf{v}_i$) (see Equation 3.23). [15]

$$a_i \text{ or } b_i = \exp\left[-D_i^2/(2 \cdot \sigma^2\right] \tag{3.22}$$

$$D_i^2 = (\mathbf{x} - \mathbf{v}_i)'(\mathbf{x} - \mathbf{v}_i) \tag{3.23}$$

c. Activations from summation layer neurodes (S) are calculated using linear functions (see Equation 3.24), where p is the number of training set exemplars with target class 1 and q is the number of training set exemplars with target class 2 [15].

$$S_a = \sum_{i=1}^{p} a_i, \quad S_b = \sum_{i=1}^{q} b_i \quad \text{and} \quad S = S_a + S_b \tag{3.24}$$

d. Activations from output layer neurodes (P or Z) are calculated using Bayes Theorem (see Equation 3.25), where $P(X|Y)P(Y) = S_a$ for target class 1, $P(X|Y)P(Y) = S_b$ for target class 2, and $P(X) = S$ (see Equation 3.26). P_1 is the posterior probability of being in class 1 and P_2 is the posterior probability of being in class 2. The highest output activation is the predicted class. This is the result of using a .5 threshold for class membership (i.e. if $Z1 \geq .5$, then class 1, otherwise class 2). [15]

$$P(Y|X) = \frac{P(X|Y)P(Y)}{P(X)} \tag{3.25}$$

$$P_1 = S_a/S \quad \text{and} \quad P_2 = S_b/S \tag{3.26}$$

e. Training stops when all of the validation set exemplars have been processed through the GRNN.

4. All of the validation set exemplars are evaluated to determine measures of performance for CA, P(TP), and P(TN) (same method as FFNN).

5. A ROC curve is plotted (same method as FFNN).

6. Finally, an AMD measure of performance is determined from the ROC curve (same method as FFNN).

3.3 Summary

The methods used in this thesis fall into two categories: data preparation and multivariate analysis. Data preparation involves taking the "raw" ATO and MISREP databases and reducing them into one final usable database with 15 features of interest. Subsequently, one data case with a low level of uncertainty (Alert) and three data cases with a high level of uncertainty (Effect) are selected for multivariate analysis. Multivariate analysis includes common statistical practices and three techniques: SNR saliency, FFNN supervised training, and GRNN supervised training. Measures of performance from multivariate analysis are used to determine the impact of uncertainty on the data.

4. Numerical Results

This chapter contains all of the numerical results from data preparation and multivariate statistical analysis. First, the mean vector section compares means between classes for input features. Next, the SNR saliency, FFNN, and GRNN sections present measures of performance from output feature prediction. Finally, the FFNN versus GRNN section seeks to find the superior neural network classifier.

4.1 Mean Vectors

The mean vectors for both of the classes are calculated for each of the four data cases: Alert, Effect .85, Effect .75, and Effect .65. The purpose of this is to determine if any of the 14 input feature have significant statistical difference between the two classes. These significant features provide a first look at the data and give some indication of those features that may be significant in SNR saliency.

The mean vectors for the Alert case (see Table 4.1) show that MSN Type, # AC, and STGS are significantly greater for the Alert class and MSN Time, TOS, and AC Type are significantly greater for the Pre-Planned class. Significant features for the Alert case are rank ordered in Table 4.2. It makes since that mission type has the greatest difference in means because certain mission types lend themselves to alert missions.

Table 4.1 Alert Case Mean Vectors

Class	MSN Type	# AC	MSN Time	TOS	AC Type	# ESM	# SAF
Alert	2.2740	2.1731	3.9877	1.4146	10.48	0.2404	0.2356
Pre-Plan	1.5466	1.8866	4.5397	1.6419	11.59	0.3360	0.2348
Significant	Yes	Yes	Yes	Yes	Yes	No	No

Class	SAF Type	STGS	# ORD	ORD Type	ALT	TGT Type	Effect
Alert	0.2837	0.2356	34.63	16.93	19304	4.0576	0.7959
Pre-Plan	0.3482	0.1336	31.61	16.19	20380	4.2660	0.8097
Significant	No	Yes	No	No	No	No	No

Table 4.2 Alert Case Significant Feature Rank Order

Rank	Feature	Z value
1	MSN Type	12.4037
2	# AC	3.6915
3	AC Type	2.2838
4	STGS	2.1163
5	MSN Time	1.8376
6	TOS	1.7668

The mean vectors for the Effect .85 case (see Table 4.3) show that MSN Time, TOS, # SAF, SAF Type, and # ORD are significantly greater for the Success class and MSN Type and # AC are significantly greater for the Failure class. Significant features for the Effect .85 case are rank ordered in Table 4.4.

Table 4.3 Effect .85 Case Mean Vectors

Class	MSN Type	# AC	MSN Time	TOS	AC Type	# ESM	# SAF
Success	2.6646	1.9238	4.8744	1.9585	10.60	0.2530	0.3232
Failure	3.0161	2.0677	4.3027	1.5808	11.04	0.2968	0.2290
Significant	Yes	Yes	Yes	Yes	No	No	Yes

Class	SAF Type	STGS	# ORD	ORD Type	ALT	TGT Type	Alert
Success	0.4207	0.1860	42.03	15.44	20139	4.1586	0.3445
Failure	0.2935	0.1387	23.18	15.89	20210	3.9720	0.3161
Significant	Yes	No	Yes	No	No	No	No

Table 4.4 Effect .85 Case Significant Feature Rank Order

Rank	Feature	Z value
1	TOS	3.3500
2	MSN Type	2.5317
3	# AC	2.4439
4	MSN Time	2.4138
5	SAF Type	2.004
6	# SAF	1.9783
7	# ORD	1.7252

The mean vectors for the Effect .75 case (see Table 4.5) show that TOS, # SAF, SAF Type, # ORD, and TGT Type are significantly greater for the Success class and MSN Type, ORD Type, and ALT are significantly greater for the Failure class. Significant features for the Effect .75 case are rank ordered in Table 4.6.

Table 4.5 Effect .75 Case Mean Vectors

Class	MSN Type	# AC	MSN Time	TOS	AC Type	# ESM	# SAF
Success	2.6848	2.0000	4.7044	1.8435	10.87	0.2488	0.3104
Failure	3.1296	1.9815	4.3860	1.6411	10.71	0.3241	0.2130
Significant	Yes	No	No	Yes	No	No	Yes

Class	SAF Type	STGS	# ORD	ORD Type	ALT	TGT Type	Alert
Success	0.3981	0.1825	42.77	15.36	19803	4.2712	0.3507
Failure	0.2824	0.1250	13.54	16.23	20896	3.6707	0.2917
Significant	Yes	No	Yes	Yes	Yes	Yes	No

Table 4.6 Effect .75 Case Significant Feature Rank Order

Rank	Feature	Z value
1	# ORD	3.1368
2	MSN Type	2.5631
3	TGT Type	2.1546
4	# SAF	2.0766
5	SAF Type	1.7858
6	ORD Type	1.7723
7	TOS	1.7410
8	ALT	1.7241

The mean vectors for the Effect .65 case (see Table 4.7) show that MSN Time, TOS, # SAF, SAF Type, # ORD, and TGT Type are significantly greater for the Success class and MSN Type and # ESM are significantly greater for the Failure class. Significant features for the Effect .65 case are rank ordered in Table 4.8.

Table 4.7 Effect .65 Case Mean Vectors

Class	MSN Type	# AC	MSN Time	TOS	AC Type	# ESM	# SAF
Success	2.7205	1.9934	4.7289	1.8578	10.85	0.2489	0.3035
Failure	3.1278	1.9944	4.2601	1.5644	10.73	0.3389	0.2111
Significant	Yes	No	Yes	Yes	No	Yes	Yes

Class	SAF Type	STGS	# ORD	ORD Type	ALT	TGT Type	Alert
Success	0.3952	0.1703	39.99	15.46	19895	4.2944	0.3450
Failure	0.2667	0.1444	14.76	16.17	20881	3.4915	0.2944
Significant	Yes	No	Yes	No	No	Yes	No

Table 4.8 Effect .65 Case Significant Feature Rank Order

Rank	Feature	Z value
1	TGT Type	2.8116
2	# ORD	2.6925
3	MSN Type	2.5631
4	TOS	2.4540
5	SAF Type	1.9335
6	# SAF	1.8762
7	MSN Time	1.7816
8	# ESM	1.6757

Overall, MSN Type and TOS are common significant features for all four cases. This indicates that certain mission types and the amount of time aircraft spend on station are generally invariant to changes in uncertainty. Additionally, TOS, # SAF, SAF Type, and # ORD are common significant features for the Success class of all three Effect cases. This makes sense since the more time an aircraft spends on station and the more ordnance that is dropped, one would expect that the probability of success would increase. Also, an aircraft that spends more time on station would be expected to be subjected to more surface-to-air firings. Finally, MSN Type is a common significant feature for the Failure class of all three Effect cases. This makes sense since some mission types are historically more successful than others.

4.2 Signal-to-Noise Ratio Saliency

SNR saliency analysis is performed for each of the four data cases: Alert, Effect .85, Effect .75, and Effect .65. The purpose of this is to reduce the initial 14 input features as much as possible without losing significant classification accuracy.

The SNR saliency results for the Alert case are summarized in Table D.1 in Appendix D, where the lowest SNR for each design point is italicized. A plot of the number of input features versus validation set CA is shown in Figure 4.1. The best combination of CA and saliency is found at 5 input features, MSN Type, # AC, MSN Time, AC Type, and # ORD. All of these except # ORD are significant features in the mean vector.

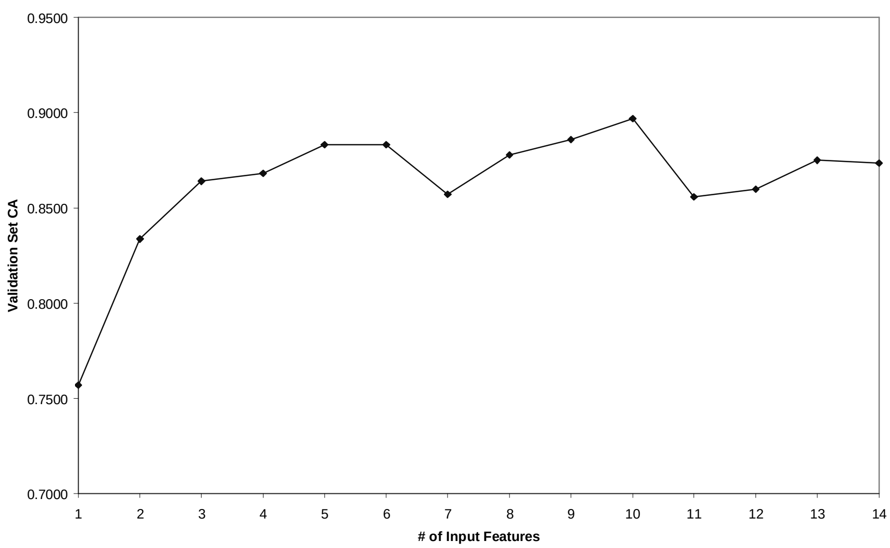

Figure 4.1 Alert Case SNR Classification Accuracy

The SNR saliency results for the Effect .85 case are summarized in Table D.2 in Appendix D, where the lowest SNR for each design point is italicized. A plot of the number of input features versus validation set CA is shown in Figure 4.2. The best combination of CA and saliency is found at 3 input features, MSN Type, # ESM, and # ORD. All of these except # ESM are significant features in the mean vector.

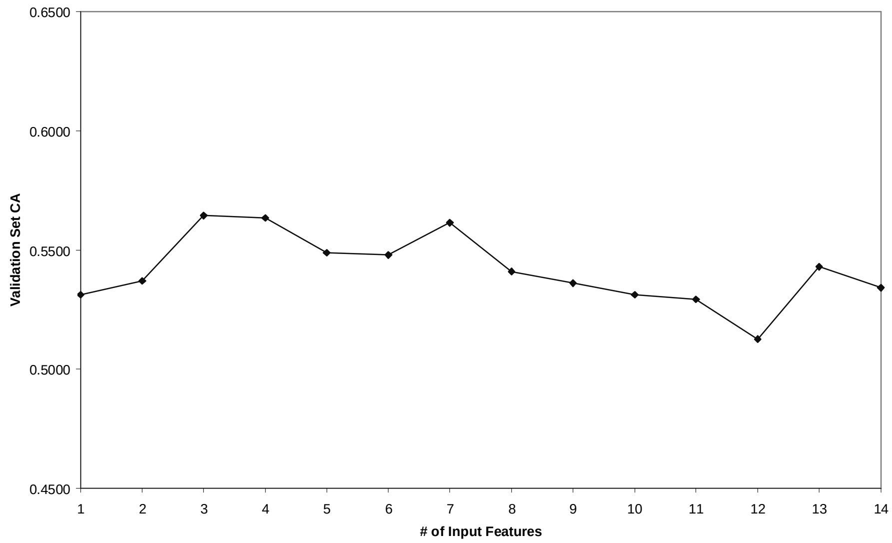

Figure 4.2 Effect .85 Case SNR Classification Accuracy

The SNR saliency results for the Effect .75 case are summarized in Table D.3 in Appendix D, where the lowest SNR for each design point is italicized. A plot of the number of input features versus validation set CA is shown in Figure 4.3. The best combination of CA and saliency is found at 5 input features, MSN Type, STGS, # ORD, ORD Type, and Alert. All of these except STGS and Alert are significant features in the mean vector.

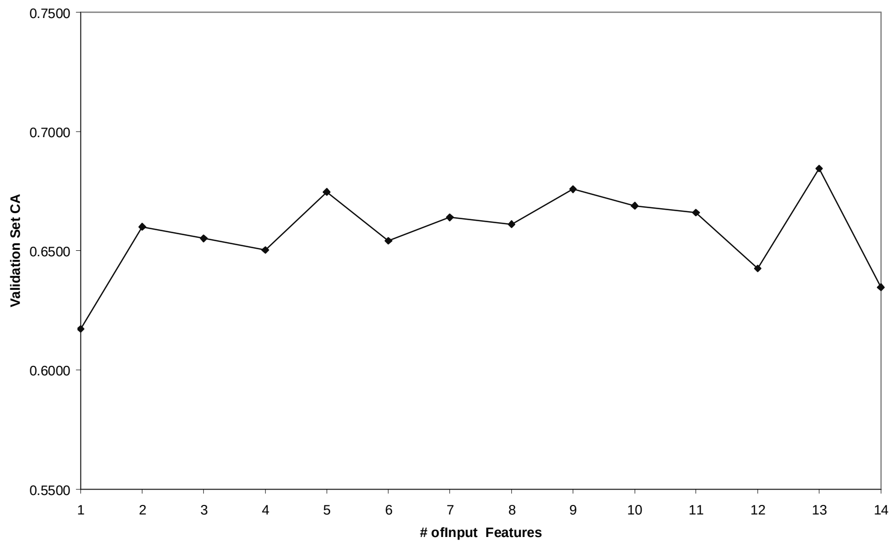

Figure 4.3 Effect .75 Case SNR Classification Accuracy

The SNR saliency results for the Effect .65 case are summarized in Table D.4 in Appendix D, where the lowest SNR for each design point is italicized. A plot of the number of input features versus validation set CA is shown in Figure 4.4. The best combination of CA and saliency is found at 4 input features, MSN Type, AC Type, # ORD, and ORD Type. All of these except AC Type and ORD Type are significant features in the mean vector.

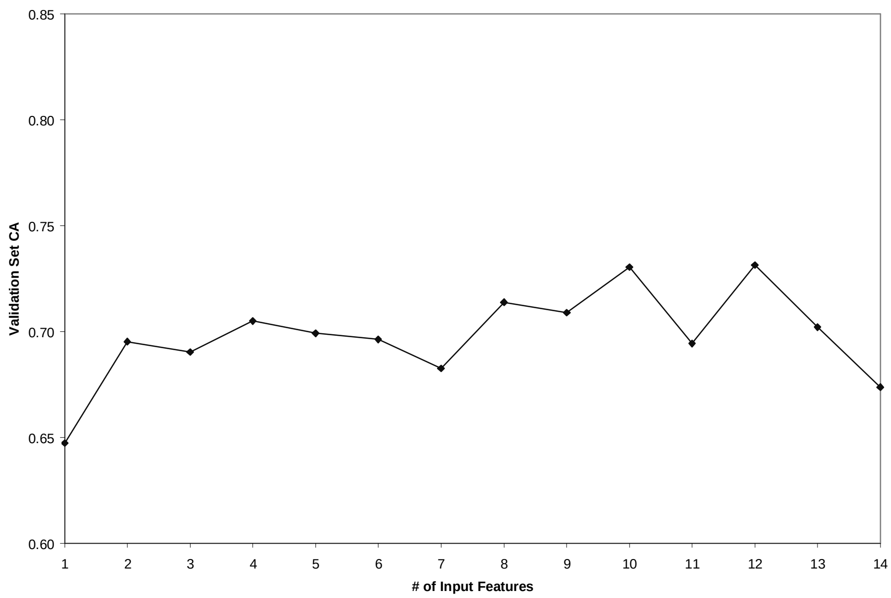

Figure 4.4 Effect .65 Case SNR Classification Accuracy

Overall, MSN Type and # ORD are common reduced input features for all four cases. Both of these features are significant in the mean vector for the three Effect cases and only MSN Type is significant in the mean vector for the Alert case. This shows that there are some strong ties between significant features in the mean vector and reduces features in SNR saliency. Additionally, # ORD is the most important and MSN Type is the second most important feature for all three Effect cases. Again, this is expected because the more ordnance that is dropped and certain mission types lead to a higher expected probability of success. As expected, the CA for the low uncertainty Alert case (0.8832) is significantly higher than the best CA (0.7051) for the high uncertainty Effect cases (Effect .65). Finally, Figure 4.5 is a plot of the effect threshold versus CA for the reduced input feature data (noise still present). It clearly shows that CA increases as the effect threshold decreases.

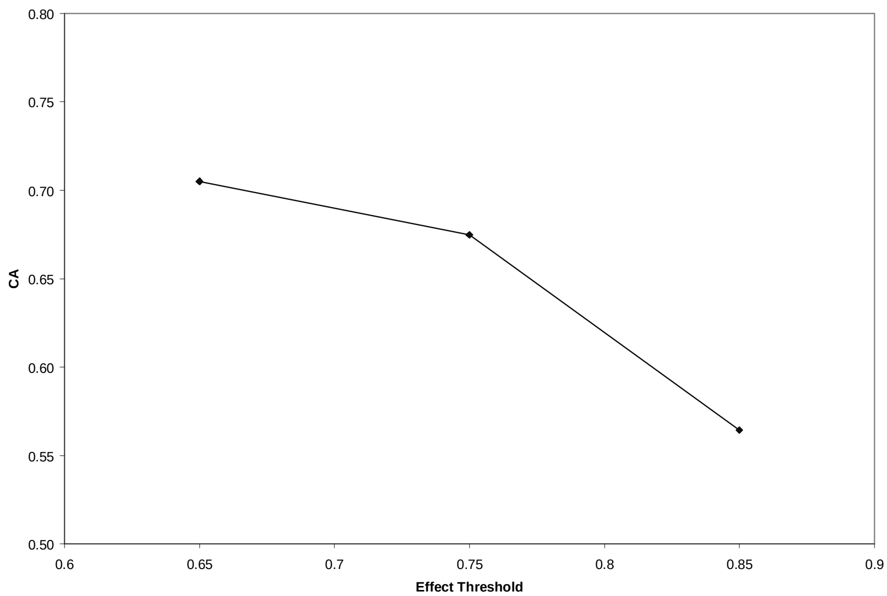

Figure 4.5 SNR Effect Threshold Versus Classification Accuracy

4.3 Feed Forward Neural Network

FFNN supervised training is performed for each of the four reduced input feature data cases: Alert (5 input features), Effect .85 (3 input features), Effect .75 (5 input features), and Effect .65 (4 input features). The purpose of this is output feature prediction. Three data methods are investigated: standardization only (coded as # of input features-S), standardization then rotation (coded as # of input features-SR), and rotation then standardization (coded as # of input features-RS). The "best" overall result is determined by the method with the highest CA.

FFNN ROC curve results for the Alert case are plotted in Figure 4.6. It is difficult to determine if the 5-S or 5-SR method gives the "best" results from the ROC curves alone.

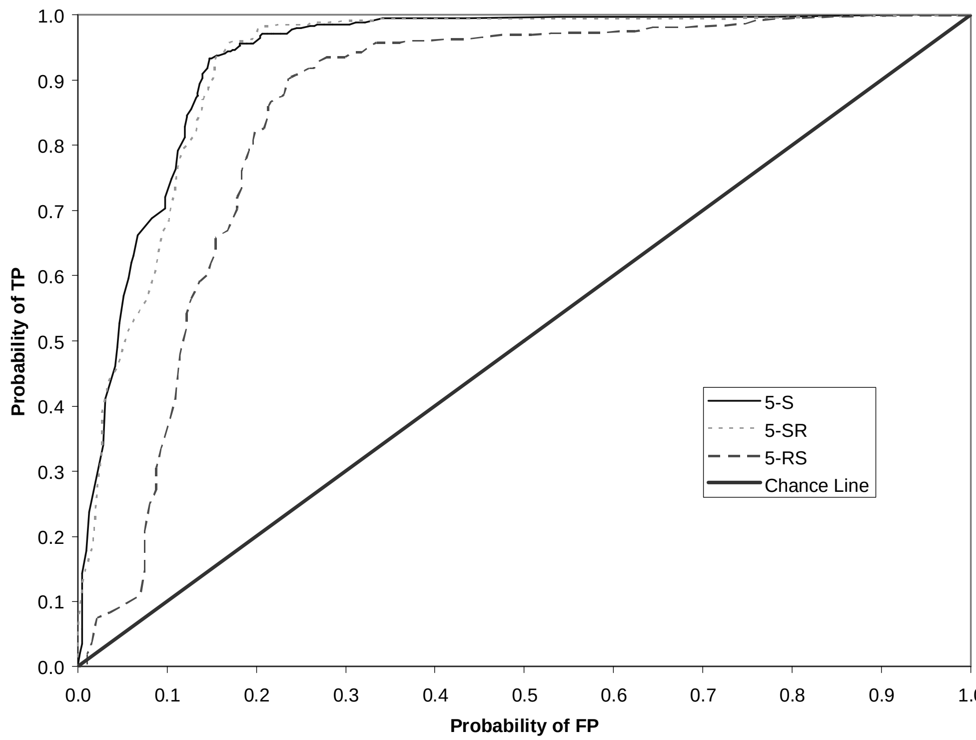

Figure 4.6 FFNN Alert Case Validation Set ROC Curves

Measures of performance results for the Alert case are summarized in Table 4.9. The 5-S method has the highest CA, P(TN), and is the most "robust" classifier, while the 5-SR method has the highest P(TP). The significance of these results is summarized in Table 4.10. This shows that there is not a significant difference between the two methods with the "best" ROC curve results, 5-S and 5-SR.

Table 4.9 FFNN Alert Case Validation Set Measures of Performance

Method	CA	P(TP)	P(TN)	AMD
5-S	0.8852	0.9150	0.8536	0.3932
5-SR	0.8805	0.9258	0.8470	0.3802
5-RS	0.8091	0.8722	0.7789	0.3376

Table 4.10 FFNN Alert Case Validation Set Significance Matrix

Measure	5-S vs. 5-SR	5-S vs. 5-RS	5-SR vs. 5-RS
CA	No	Yes	Yes
P(TP)	No	No	Yes
P(TN)	No	Yes	Yes
AMD	No	Yes	No

FFNN ROC curve results for the Effect .85 case are plotted in Figure 4.7. It is difficult to determine which of the three methods give the "best" results from the ROC curves alone.

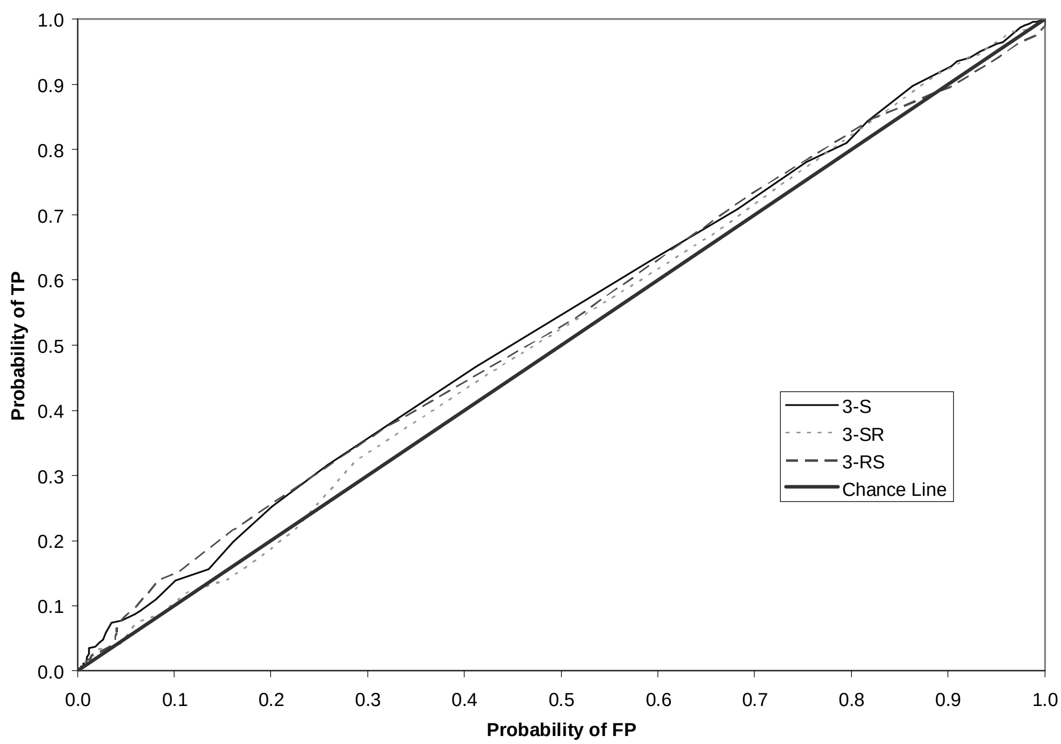

Figure 4.7 FFNN Effect .85 Case Validation Set ROC Curves

Measures of performance results for the Effect .85 case are summarized in Table 4.11. The 3-S method has the highest CA and P(TP), while the 3-RS method has the highest P(TN) and is the most "robust" classifier. The significance of these results is summarized in Table 4.12. This shows that there is not a significant difference between the three methods except for AMD using the 3-RS method.

Table 4.11 FFNN Effect .85 Case Validation Set Measures of Performance

Method	CA	P(TP)	P(TN)	AMD
3-S	0.5195	0.6280	0.4089	0.2982
3-SR	0.5049	0.5909	0.4285	0.2977
3-RS	0.5039	0.5441	0.4831	0.3157

Table 4.12 FFNN Effect .85 Case Validation Set Significance Matrix

Measure	3-S vs. 3-SR	3-S vs. 3-RS	3-SR vs. 3-RS
CA	No	No	No
P(TP)	No	No	No
P(TN)	No	No	No
AMD	No	Yes	Yes

FFNN ROC curve results for the Effect .75 case are plotted in Figure 4.8. It is difficult to determine which of the three methods give the "best" results from the ROC curves alone.

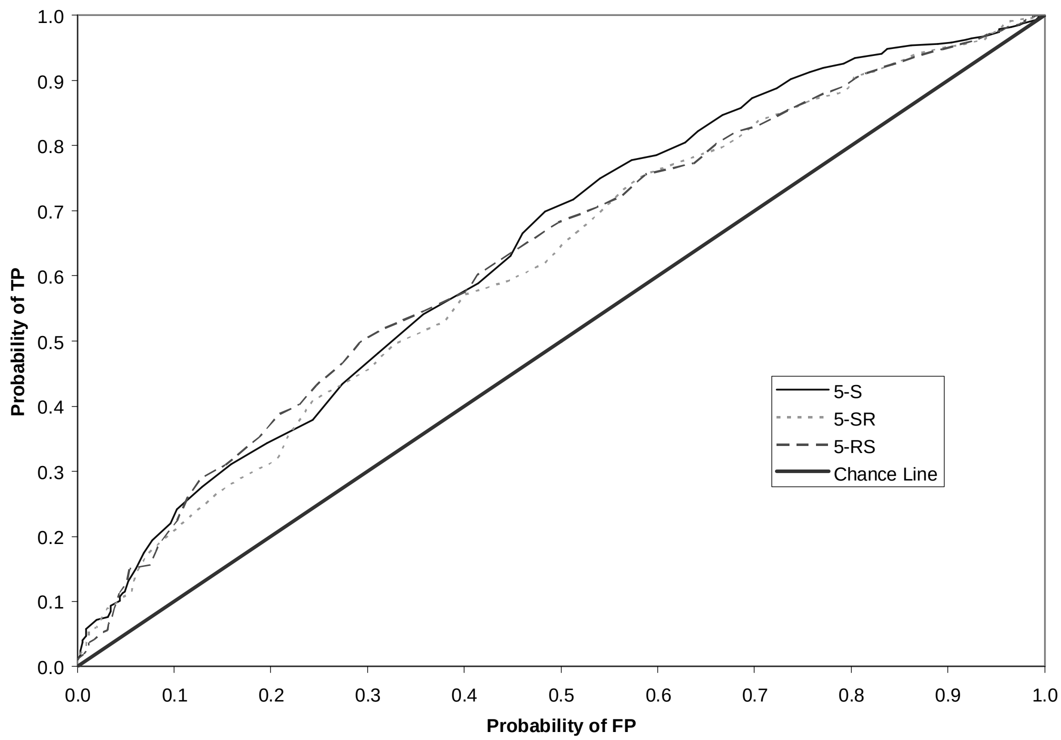

Figure 4.8 FFNN Effect .75 Case Validation Set ROC Curves

Measures of performance results for the Effect .75 case are summarized in Table 4.13. The 5-S method has the highest CA and P(TN), the 5-SR method has the highest P(TP), and the 5-RS method is the most "robust" classifier. The significance of these results is summarized in Table 4.14. This shows that there is not a significant difference between the three methods except for CA using the 5-S method.

Table 4.13 FFNN Effect .75 Case Validation Set Measures of Performance

Method	CA	P(TP)	P(TN)	AMD
5-S	0.6738	0.8725	0.3028	0.2507
5-SR	0.6396	0.8806	0.2127	0.2606
5-RS	0.6406	0.8305	0.2957	0.2680

Table 4.14 FFNN Effect .75 Case Validation Set Significance Matrix

Measure	5-S vs. 5-SR	5-S vs. 5-RS	5-SR vs. 5-RS
CA	Yes	Yes	No
P(TP)	No	No	No
P(TN)	Yes	No	No
AMD	No	No	No

FFNN ROC curve results for the Effect .65 case are plotted in Figure 4.9. It is difficult to determine if the 4-S or 4-SR method gives the "best" results from the ROC curves alone.

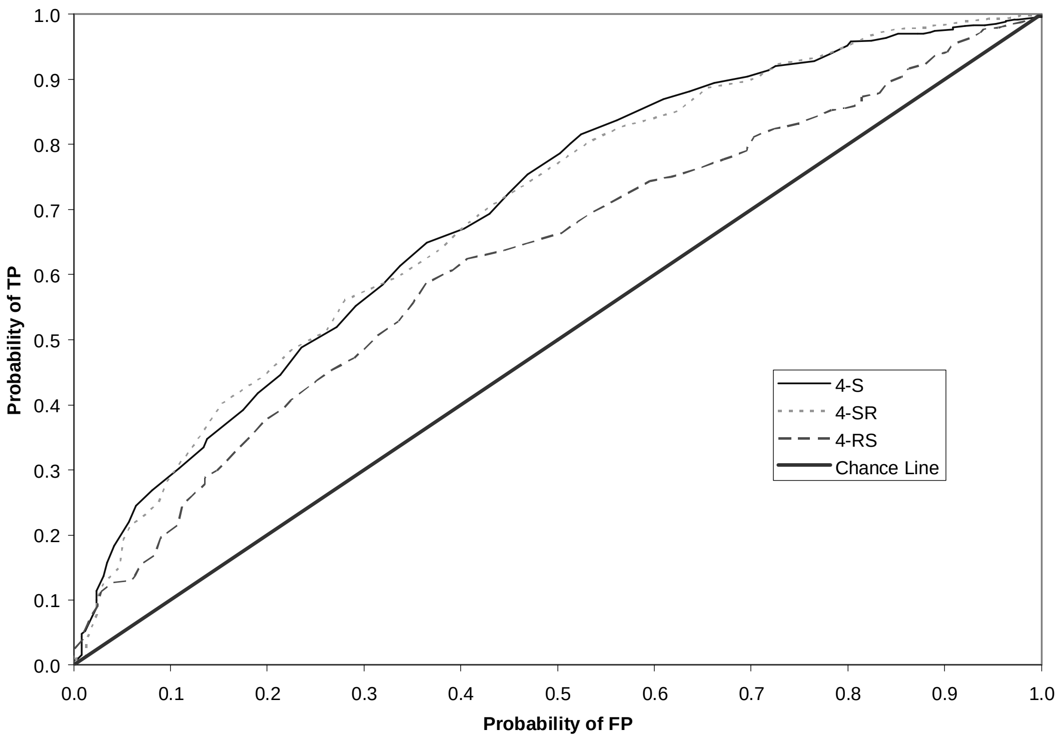

Figure 4.9 FFNN Effect .65 Case Validation Set ROC Curves

Measures of performance results for the Effect .65 case are summarized in Table 4.15. The 4-S method has the highest CA, the 4-SR method has the highest P(TP) and is the most "robust" classifier, and the 4-RS method has the highest P(TN). The significance of these results is summarized in Table 4.16. This shows that there is not a significant difference between the two methods with the "best" ROC curve results, 5-S and 5-SR, except for P(TP) and AMD.

Table 4.15 FFNN Effect .65 Case Validation Set Measures of Performance

Method	CA	P(TP)	P(TN)	AMD
4-S	0.7461	0.9211	0.2753	0.2729
4-SR	0.7383	0.9663	0.1781	0.3048
4-RS	0.6338	0.7221	0.4306	0.2569

Table 4.16 FFNN Effect .65 Case Validation Set Significance Matrix

Measure	4-S vs. 4-SR	4-S vs. 4-RS	4-SR vs. 4-RS
CA	No	Yes	Yes
P(TP)	Yes	Yes	Yes
P(TN)	No	No	Yes
AMD	Yes	No	Yes

Overall, the standardization only method provides the "best" results in terms of CA for all four cases. This result is significantly higher for only the Effect .75 case. Also, the other three measures of performance do not have a common "best" result method. Therefore, the standardization only method will be used for comparison with GRNN results for all four cases. As expected, the CA for the low uncertainty Alert case (0.8852) is still significantly higher than the best CA (0.7461) for the high uncertainty Effect cases (Effect .65). Figure 4.10 is a plot of the effect threshold versus the four measures of performance for the reduced input feature data. It clearly shows that CA and P(TP) increase as the effect threshold decreases, P(TN) decreases as the effect threshold decreases, and AMD is invariant to changes in threshold. This results in ROC curves moving up and to the right as the threshold decreases. This leads to almost all of the missions being predicted as successful at a .65 threshold, which hinders the purpose of classification.

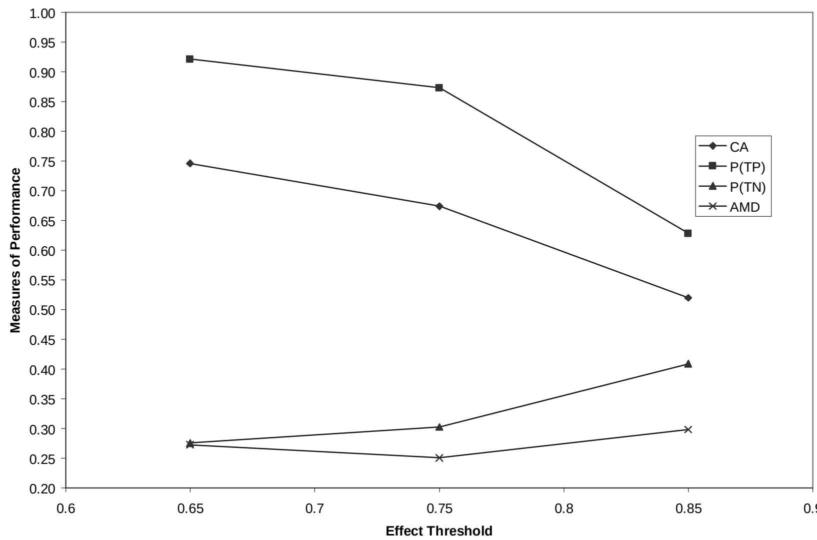

Figure 4.10 FFNN Effect Threshold Versus Measures of Performance

4.4 General Regression Neural Network

GRNN supervised training is performed for each of the four reduced input feature data cases: Alert (5 input features), Effect .85 (3 input features), Effect .75 (5 input features), and Effect .65 (4 input features). The purpose of this is output feature prediction. Three data methods are investigated: standardization only (coded as # of input features-S), standardization then rotation (coded as # of input features-SR), and rotation then standardization (coded as # of input features-RS). The "best" overall result is determined by the method with the highest CA.

The effects of spread for the Alert case are summarized in Table 4.17. The optimal spread was selected at 0.1 for all three methods.

Table 4.17 GRNN Alert Case Spreads

Spread	5-S CA	5-SR CA	5-RS CA
1.0	0.8324	0.8558	0.5426
0.9	0.8407	0.8585	0.5316
0.8	0.8819	0.8695	0.5343
0.7	0.8956	0.8832	0.5522
0.6	0.9011	0.8654	0.5261
0.5	0.8846	0.8791	0.6154
0.4	0.8887	0.8736	0.5852
0.3	0.8915	0.8819	0.5632
0.2	0.8984	0.8929	0.6827
0.1	0.9038	0.8956	0.6992

GRNN ROC curve results for the Alert case are plotted in Figure 4.11. It is difficult to determine if the 5-S or 5-SR method gives the “best” results from the ROC curves alone.

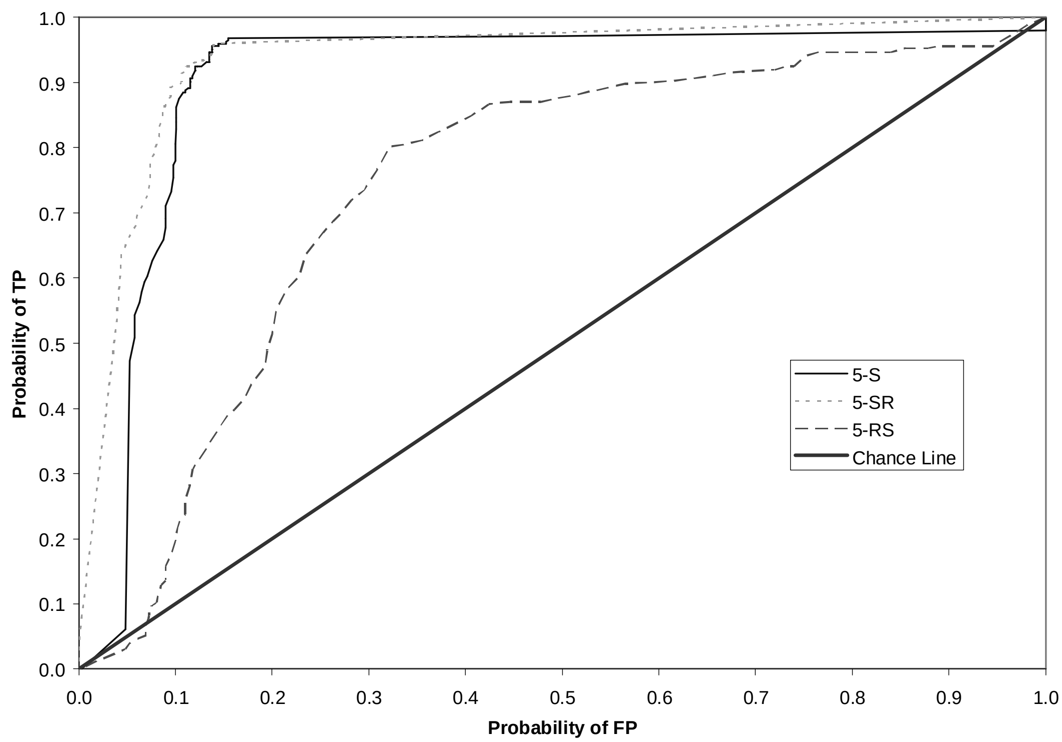

Figure 4.11 GRNN Alert Case Validation Set ROC Curves

Measures of performance results for the Alert case are summarized in Table 4.18. The 5-S method has the highest CA, P(TP), and P(TN), while the 5-SR method is the most "robust" classifier. The significance of these results is summarized in Table 4.19. This shows that there is not a significant difference between the two methods with the "best" ROC curve results, 5-S and 5-SR.

Table 4.18 GRNN Alert Case Validation Set Measures of Performance

Method	CA	P(TP)	P(TN)	AMD
5-S	0.9052	0.9213	0.8925	0.4733
5-SR	0.8984	0.9170	0.8886	0.4834
5-RS	0.6992	0.6820	0.7361	0.2157

Table 4.19 GRNN Alert Case Validation Set Significance Matrix

Measure	5-S vs. 5-SR	5-S vs. 5-RS	5-SR vs. 5-RS
CA	No	Yes	Yes
P(TP)	No	Yes	Yes
P(TN)	No	Yes	Yes
AMD	No	Yes	Yes

The effects of spread for the Effect .85 case are summarized in Table 4.20. The optimal spread was selected at 0.7 for the 3-S method, 0.5 for the 3-SR method, and 0.8 for the 3-RS method.

Table 4.20 GRNN Effect .85 Case Spreads

Spread	3-S CA	3-SR CA	3-RS CA
1.0	0.4893	0.5107	0.4766
0.9	0.5137	0.5254	0.4775
0.8	0.5156	0.5205	0.4824
0.7	0.5371	0.5078	0.4648
0.6	0.4990	0.5215	0.4600
0.5	0.5244	0.5361	0.4805
0.4	0.5068	0.5293	0.4707
0.3	0.5010	0.4971	0.4746
0.2	0.5088	0.5088	0.4795
0.1	0.4961	0.5029	0.4639

GRNN ROC curve results for the Effect .85 case are plotted in Figure 4.12. It is difficult to determine if the 3-S or 3-SR method gives the "best" results from the ROC curves alone.

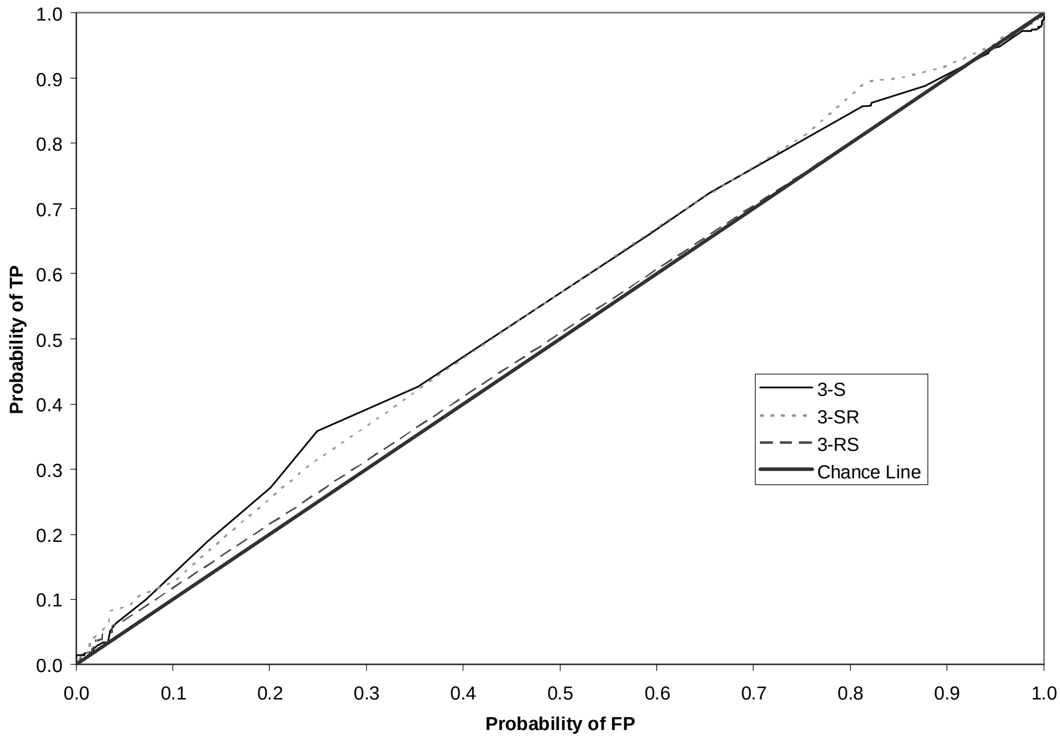

Figure 4.12 GRNN Effect .85 Case Validation Set ROC Curves

Measures of performance results for the Effect .85 case are summarized in Table 4.21. The 3-S method has the highest CA and P(TN), while the 3-RS method has the highest P(TP) and is the most "robust" classifier. The significance of these results is summarized in Table 4.22. This shows that there is not a significant difference between the two methods with the "best" ROC curve results, 3-S and 3-SR.

Table 4.21 GRNN Effect .85 Case Validation Set Measures of Performance

Method	CA	P(TP)	P(TN)	AMD
3-S	0.5332	0.6600	0.4081	0.3020
3-SR	0.5283	0.6865	0.3846	0.2981
3-RS	0.4814	0.7130	0.2911	0.3347

Table 4.22 **GRNN Effect .85 Case Validation Set Significance Matrix**

Measure	3-S vs. 3-SR	3-S vs. 3-RS	3-SR vs. 3-RS
CA	No	Yes	Yes
P(TP)	No	No	No
P(TN)	No	No	No
AMD	No	Yes	Yes

The effects of spread for the Effect .75 case are summarized in Table 4.23. The optimal spread was selected at 0.7 for the 5-S method, 0.5 for the 5-SR method, and 0.6 for the 5-RS method.

Table 4.23 **GRNN Effect .75 Case Spreads**

Spread	5-S CA	5-SR CA	5-RS CA
1.0	0.6406	0.6494	0.6484
0.9	0.6680	0.6689	0.6660
0.8	0.6602	0.6533	0.6514
0.7	0.6846	0.6553	0.6602
0.6	0.6787	0.6797	0.6777
0.5	0.6797	0.6865	0.6553
0.4	0.6582	0.6836	0.6504
0.3	0.6592	0.6709	0.6621
0.2	0.6494	0.6592	0.6641
0.1	0.6387	0.6563	0.6533

GRNN ROC curve results for the Effect .75 case are plotted in Figure 4.13. It is difficult to determine if the 5-S or 5-SR method gives the “best” results from the ROC curves alone.

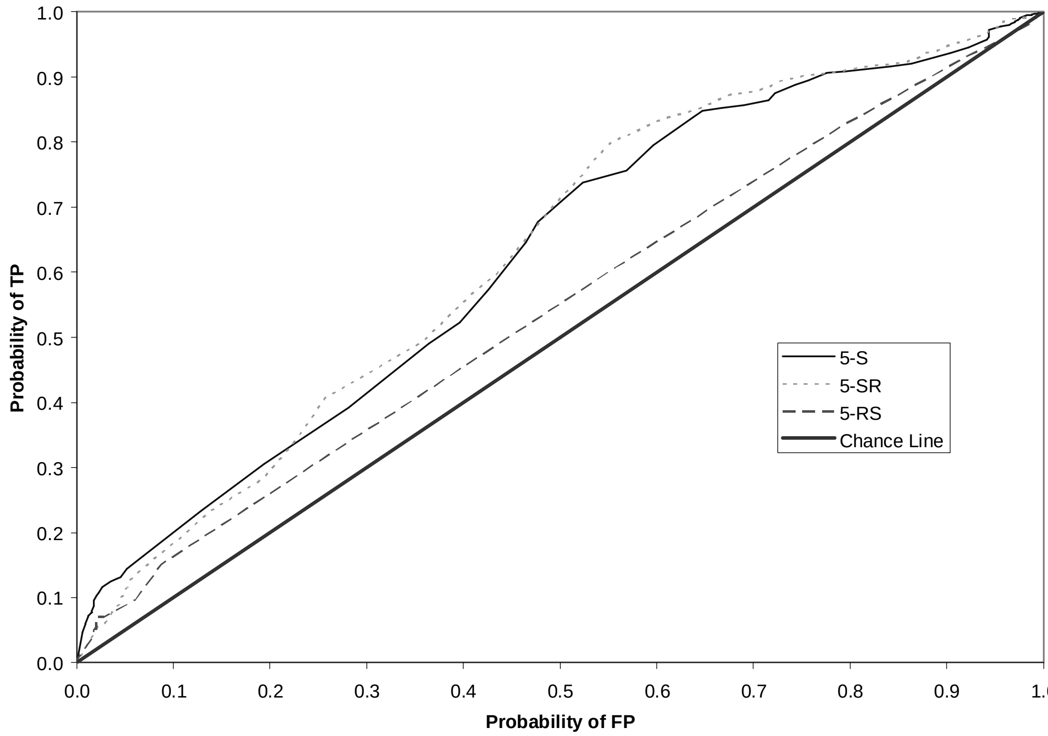

Figure 4.13 GRNN Effect .75 Case Validation Set ROC Curves

Measures of performance results for the Effect .75 case are summarized in Table 4.24. The 5-S method has the highest CA, the 5-SR method has the highest P(TN), and the 5-RS method has the highest P(TP) and is the most "robust" classifier. The significance of these results is summarized in Table 4.25. This shows that there is a significant difference between the two methods with the "best" ROC curve results, 5-S and 5-SR, except for P(TP).

Table 4.24 GRNN Effect .75 Case Validation Set Measures of Performance

Method	CA	P(TP)	P(TN)	AMD
5-S	0.6914	0.9059	0.2244	0.2781
5-SR	0.6777	0.8862	0.2821	0.2594
5-RS	0.6719	0.9985	0.0000	0.3465

Table 4.25 GRNN Effect .75 Case Validation Set Significance Matrix

Measure	5-S vs. 5-SR	5-S vs. 5-RS	5-SR vs. 5-RS
CA	Yes	No	No
P(TP)	No	Yes	Yes
P(TN)	Yes	Yes	Yes
AMD	Yes	Yes	Yes

The effects of spread for the Effect .65 case are summarized in Table 4.26. The optimal spread was selected at 0.6 for the 4-S method, 0.7 for the 4-SR method, and 0.2 for the 4-RS method.

Table 4.26 GRNN Effect .65 Case Spreads

Spread	4-S CA	4-SR CA	4-RS CA
1.0	0.7197	0.6963	0.7061
0.9	0.7295	0.7041	0.7168
0.8	0.7314	0.7100	0.7178
0.7	0.7295	0.7422	0.7139
0.6	0.7490	0.7246	0.7188
0.5	0.7480	0.7158	0.7080
0.4	0.7295	0.7188	0.7051
0.3	0.7041	0.7070	0.7002
0.2	0.6914	0.7100	0.7217
0.1	0.7119	0.6904	0.7168

GRNN ROC curve results for the Effect .65 case are plotted in Figure 4.14. It is difficult to determine if the 4-S or 4-SR method gives the "best" results from the ROC curves alone.

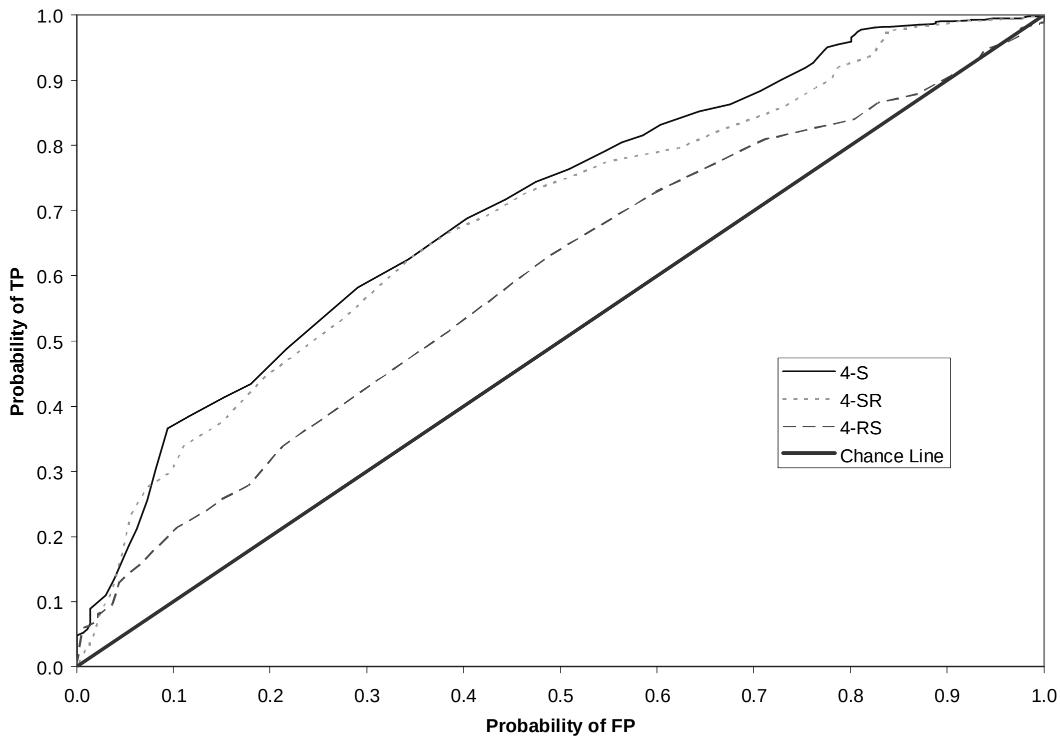

Figure 4.14 GRNN Effect .65 Case Validation Set ROC Curves

Measures of performance results for the Effect .65 case are summarized in Table 4.27. The 4-S method has the highest CA and P(TN), while the 4-RS method has the highest P(TP) and is the most "robust" classifier. The significance of these results is summarized in Table 4.28. This shows that there is no significant difference between the two methods with the "best" ROC curve results, 4-S and 4-SR, except for P(TN) and AMD.

Table 4.27 GRNN Effect .65 Case Validation Set Measures of Performance

Method	CA	P(TP)	P(TN)	AMD
4-S	0.7480	0.9632	0.1991	0.2983
4-SR	0.7383	0.9686	0.1646	0.3095
4-RS	0.7256	0.9880	0.0040	0.3428

Table 4.28 GRNN Effect .65 Case Validation Set Significance Matrix

Measure	4-S vs. 4-SR	4-S vs. 4-RS	4-SR vs. 4-RS
CA	No	Yes	No
P(TP)	No	Yes	Yes
P(TN)	Yes	Yes	Yes
AMD	Yes	Yes	Yes

Overall, the standardization only method provides the "best" results in terms of CA for all four cases. This result is significantly higher for only the Effect .75 case. Also, P(TP) and AMD have a common "best" result in the rotation then standardization method for all three Effect cases. Therefore, the standardization only method will be used for comparison with FFNN results for all four cases. As expected, the CA for the low uncertainty Alert case (0.9052) is still significantly higher than the best CA (0.7480) for the high uncertainty Effect cases (Effect .65). Figure 4.15 is a plot of the effect threshold versus the four measures of performance for the reduced input feature data. It clearly shows that CA and P(TP) increase as the effect threshold decreases, P(TN) decreases as the effect threshold decreases, and AMD is invariant to changes in threshold. These are the same results as FFNN supervised training.

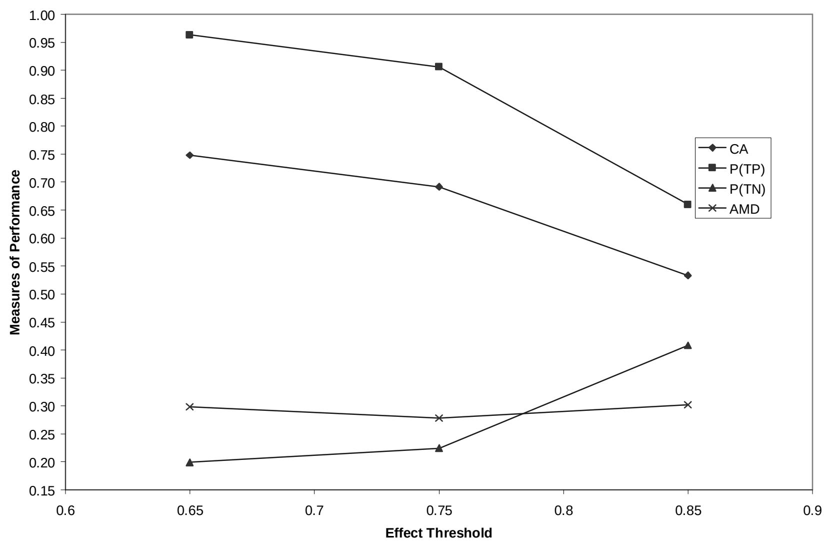

Figure 4.15 GRNN Effect Threshold Versus Measures of Performance

4.5 FFNN Versus GRNN

A comparison of the "best" results from FFNN supervised training and GRNN supervised training is made for each of the four reduced input feature data cases: Alert (5 input features), Effect .85 (3 input features), Effect .75 (5 input features), and Effect .65 (4 input features). The purpose of this is to determine which neural network classifier is superior. Only the "best" data method from both FFNN and GRNN supervised training is investigated: standardization only (coded as # of input features-S). The "best" overall classifier is determined by the neural network with the highest CA.

ROC curve results for the Alert case are plotted in Figure 4.16. It is difficult to determine if FFNN 5-S or GRNN 5-S gives the "best" results from the ROC curves alone.

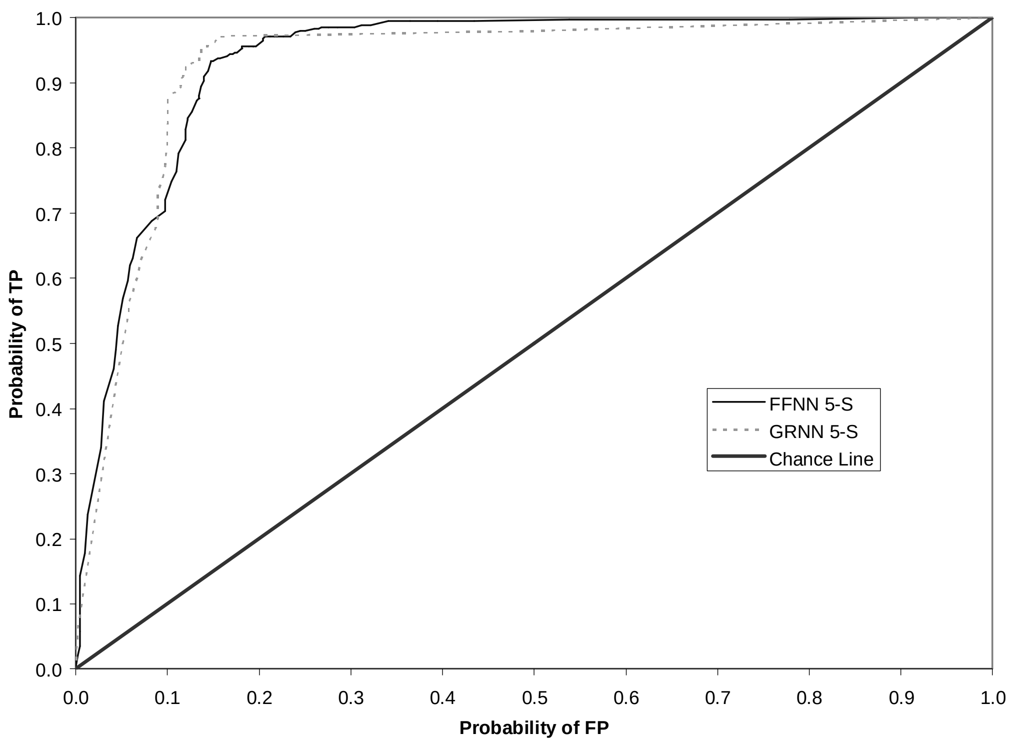

Figure 4.16 FFNN vs. GRNN Alert Case Validation Set ROC Curves

Measures of performance results and significance for the Alert case are summarized in Table 4.29. The GRNN 5-S has the highest CA, P(TP), P(TN), and is the most "robust" classifier. There is not a significant difference between the two classifiers except for P(TN) and AMD.

Table 4.29 FFNN vs. GRNN Alert Case Validation Set Measures of Performance

Classifier	CA	P(TP)	P(TN)	AMD
FFNN 5-S	0.8852	0.9150	0.8536	0.3932
GRNN 5-S	0.9052	0.9213	0.8925	0.4733
Significant	No	No	Yes	Yes

ROC curve results for the Effect .85 case are plotted in Figure 4.17. It appears that GRNN 3-S gives the "best" results from the ROC curves.

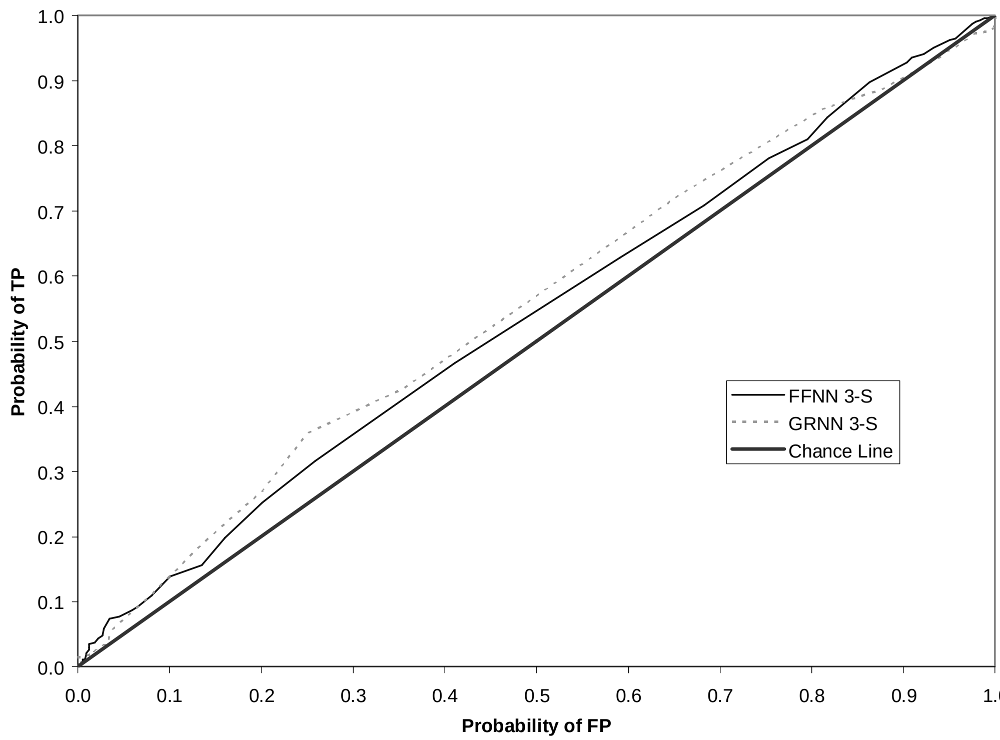

Figure 4.17 FFNN vs. GRNN Effect .85 Case Validation Set ROC Curves

Measures of performance results and significance for the Effect .85 case are summarized in Table 4.30. The GRNN 3-S has the highest CA, P(TP), and is the most "robust" classifier, while FFNN 3-S has the highest P(TN). There is not a significant difference between the two classifiers.

Table 4.30 FFNN vs. GRNN Effect .85 Case Validation Set Measures of Performance

Classifier	CA	P(TP)	P(TN)	AMD
FFNN 3-S	0.5195	0.6280	0.4089	0.2982
GRNN 3-S	0.5332	0.6600	0.4081	0.3020
Significant	No	No	No	No

ROC curve results for the Effect .75 case are plotted in Figure 4.18. It is difficult to determine if FFNN 5-S or GRNN 5-S gives the "best" results from the ROC curves alone.

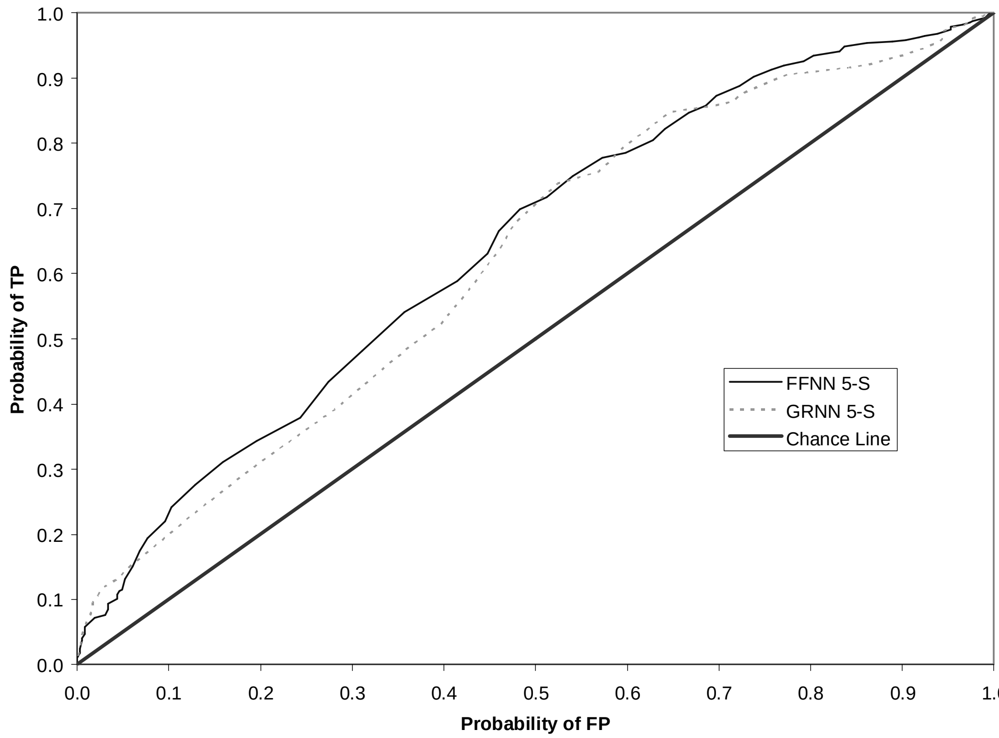

Figure 4.18 FFNN vs. GRNN Effect .75 Case Validation Set ROC Curves

Measures of performance results and significance for the Effect .75 case are summarized in Table 4.31. The GRNN 5-S has the highest CA, P(TP), and is the most "robust" classifier, while FFNN 5-S has the highest P(TN). There is a significant difference between the two classifiers in all measures except CA.

Table 4.31 FFNN vs. GRNN Effect .75 Case Validation Set Measures of Performance

Classifier	CA	P(TP)	P(TN)	AMD
FFNN 5-S	0.6738	0.8725	0.3028	0.2507
GRNN 5-S	0.6914	0.9059	0.2244	0.2781
Significant	No	Yes	Yes	Yes

ROC curve results for the Effect .65 case are plotted in Figure 4.19. It is difficult to determine if FFNN 4-S or GRNN 4-S gives the "best" results from the ROC curves alone.

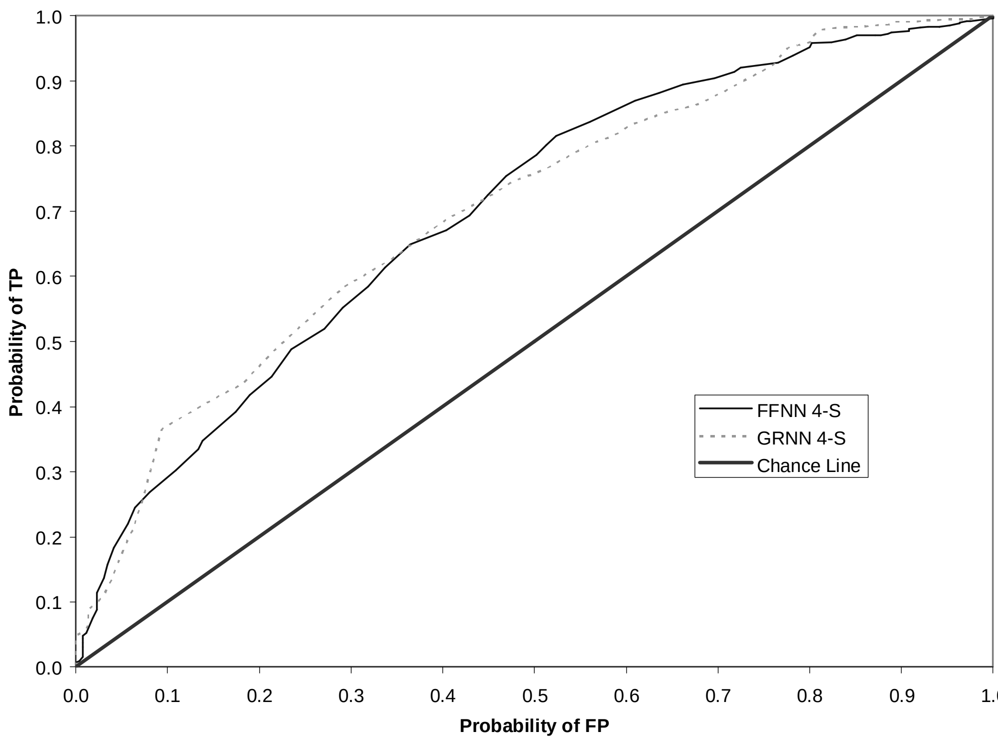

Figure 4.19 FFNN vs. GRNN Effect .65 Case Validation Set ROC Curves

Measures of performance results and significance for the Effect .65 case are summarized in Table 4.32. The GRNN 4-S has the highest CA, P(TP), and is the most "robust" classifier, while FFNN 4-S has the highest P(TN). There is no significant difference between the two classifiers except for P(TP) and AMD.

Table 4.32 FFNN vs. GRNN Effect .65 Case Validation Set Measures of Performance

Classifier	CA	P(TP)	P(TN)	AMD
FFNN 4-S	0.7461	0.9211	0.2753	0.2729
GRNN 4-S	0.7480	0.9632	0.1991	0.2983
Significant	No	Yes	No	Yes

Overall, the GRNN is the superior neural network classifier for this data. This was expected based on previous work with this data. GRNN supervised training provides the "best" results in terms of CA for all four cases. This result is not significantly higher than the FFNN for any of the cases. Also, P(TP) and AMD have a common "best" result in the GRNN and P(TN) has a common "best" result in the FFNN for all three Effect cases. As expected, the CA for the low uncertainty Alert case (0.9052) is still significantly higher than the best CA (0.7480) for the high uncertainty Effect cases (Effect .65).

4.6 Summary

Numerical results are obtained from analyzing the mean vector and performing SNR saliency, FFNN supervised training, and GRNN supervised training on all four data cases. Mean vector analysis shows that there are common significant features for all four cases and among the success and failure classes of the three Effect cases. SNR saliency analysis reduced the number of input features from 14 down to 3,4, or 5, depending on the case. Most of these reduced features are common to the significant features in the mean vector. The same two input features, MSN Type and # ORD, are the most important for all four cases. This shows the invariance of these features to uncertainty. FFNN supervised training measures of performance show that the standardize only method provides the best CA. The same holds true for GRNN supervised training. The GRNN is superior to the FFNN for all four data cases. Uncertainty had a big impact in all of these techniques. CA for the low uncertainty case (Alert) is much higher than any of the high uncertainty cases (Effect). As the threshold of Effect decreases, CA and P(TP) increase, P(TN) increases, and AMD is invariant.

5. Conclusions and Future Research

This chapter wraps up the overview, literature review, methodology, and numerical results presented in the previous four chapters. First, special problems encountered in the course of this research are described. Next, conclusions based on numerical results are presented. Finally, recommendations for further research are provided.

5.1 Special Problems Encountered

Special problems were encountered in two areas: MISREP data and EBO objectives. The MISREP data provided for this thesis was very inconsistent in both format and content. This was the cause of a lot of uncertainty in the features, in particular SAF Type, TGT Type, and Effect. Text-based inputs for all three of these features made it difficult to identify the intended data field entry. It seems clear that this problem, and a lot of related uncertainty, can be dealt with by implementing a standardized format for MISREPs that limits each data field to pre-determined options.

The ATO and MISREP data provided for this thesis did not connect tactical (mission) level results to EBO objectives. This causes a lot of uncertainty in determining whether or not a mission is a success at the operational level. For this thesis, success was determined by interpreting the results of each individual weapon and setting a success threshold at the mission level. This naturally does not give a complete picture of mission success. Ideally, mission results would be measured by comparing tactical level results to the impact on EBO objectives and setting a success threshold.

5.2 Conclusions

The results of this research show that uncertainty has a significant impact on real-world aerospace mission data. Results are obtained from four data cases: low level of uncertainty (Alert) and high level of uncertainty (Effect .85, Effect .75, and Effect .65). These cases are compared using five analytical techniques: mean vectors, SNR saliency, FFNN supervised training, GRNN supervised training, and FFNN versus GRNN.

Mean vector analysis shows that there are significant statistical differences between the two classes for each case. MSN Type and TOS are common significant features for all four cases. TOS, # SAF, SAF Type, and # ORD are common significant features for the Success class of all three Effect cases. MSN Type is a common significant feature for the Failure class of all three Effect cases. Most of these significant features appear again as reduced input features in SNR saliency analysis.

SNR saliency analysis shows that the original 14 input features can be reduced to 5 for the Alert and Effect .75 cases, 4 for the Effect .65 case, and 3 for the Effect .85 case. The two most important input features for all four cases are MSN Type and # ORD. Uncertainty has a direct impact on CA as the low uncertainty Alert case has 88% CA compared to 70% CA for the best Effect case (Effect .65). In fact, CA decreases for the Effect case as the threshold increases. This is expected since an overwhelming majority of US aerospace missions are historically successful, which would indicate a lower threshold for success.

FFNN supervised training shows that the standardization only method provides the best results, in terms of CA, for all four cases. There is no best method for P(TP), P(TN), and AMD measures of performance. This indicates that data rotation by PCA loadings has no impact on FFNN supervised training. As with SNR saliency, uncertainty has a direct impact on CA as the low uncertainty Alert case has 88% CA compared to 74% CA for the best Effect case (Effect .65). For the Effect cases, CA and P(TP) decrease, P(TN) increases, and AMD is invariant as the threshold increases. This causes the ROC curve to move up and to the right (instead of the desired up and to the left) as the threshold decreases.

GRNN supervised training shows that the standardization only method provides the best results, in terms of CA, for all four cases. The best method for P(TP) and AMD is rotation then standardization. As with SNR saliency and FFNN, uncertainty has a direct impact on CA as the low uncertainty Alert case has 90% CA compared to 74% CA for the best Effect case (Effect .65). For the Effect cases, CA and P(TP) decrease, P(TN) increases, and AMD is invariant as the threshold increases. This causes the ROC curve to move up and to the right (instead of the desired up and to the left) as the threshold decreases.

A FFNN versus GRNN comparison shows that the GRNN using the standardization method provides the best results, in terms of CA, for all four cases. Clearly, the GRNN is the superior classifier.

Overall, analysis of the low uncertainty case (Alert) and high certainty cases (Effect .85, Effect .75, Effect .65) shows that it is very difficult to predict mission success when there is a high level of uncertainty in quantifying the measure of success and failure in the first place. The results from the Alert case clearly show that a very high CA can be obtained from this database using multivariate analysis techniques. Therefore, the uncertainty in the Effect feature must be the main cause of lower (not much higher than chance) CA for the Effect cases. Higher CA can be obtained by lowering the Effect threshold, but this has the effect of predicting almost every mission a success (high P(TP) and low P(TN)). As a result, the uncertainty in the Effect feature must be lowered in order to obtain reasonable (greater than 70%) CA for mission success.

5.3 Recommendations for Future Research

There are several different areas for future research with this thesis. First, the source data can be manipulated and changed to create different final useable databases for multivariate analysis. For example, a different three ATO day period can be chosen from the OIF data or source data from a different conflict, such as Kosovo or Afghanistan, can be used. Next, there might be other significant input features in the ATO and MISREP databases used in the thesis, or data not available for this thesis, that improve CA or reduce the number of input features. Also, different neural network architectures of the FFNN and GRNN, or a completely different classifier, can be investigated to determine if similar, or even better, CA results can be obtained.

Additionally, it is possible to measure mission success with methods different than the Effect method presented in this thesis. Possibilities include measuring success on the individual weapon level or in multiple stages, which might include successful weapon release, successful hit, successful high order explosion, and successful strategic level effect.

More importantly, collecting and/or generating data connecting tactical level results to EBO objectives is critical to reducing the level of uncertainty and improving CA for predicting mission success or failure. The combination of that data and the multivariate analysis techniques used in this thesis could produce a tool for decision makers in the AOC to measure one mission setup versus another to determine the best choice for mission success.

5.4 Contributions

The primary contribution of this thesis is that real-world aerospace mission data can be used to classify mission success. The low uncertainty Alert case validated the multivariate analysis techniques used in this thesis and the high uncertainty

Effect cases followed with a mathematical foundation for classifying mission success. A secondary contribution is that sources of uncertainty in ATO and MISREP data were identified and hopefully improvements to data collection and analysis techniques can be made in the future.

Appendix A. List of Acronyms

AAA	Anti-Aircraft Artillery
ACO	Airspace Control Order
ADOCS	Automated Deep Operations Coordination System
AFDD	Air Force Doctrine Document
AMD	Average Metric Distance
AO	Air Operations
AOC	Aerospace Operations Center
AOP	Air Operations Plan
ATO	Air Tasking Order
BDA	Battle Damage Assessment
CA	Classification Accuracy
CAS	Close Air Support
CI	Confidence Interval
COD	Combat Operations Division
CPD	Combat Plans Division
EBO	Effects Based Operations
FFNN	Feed Forward Neural Network
FP	False Positive
GRNN	General Regression Neural Network
ISDS	Intelligence Shared Data Server
ITS	Interim Targeting Solution
JFC	Joint Force Commander
JFACC	Joint Force Air Component Commander
JGAT	Joint Guidance Apportionment and Targeting

JIPTL	Joint Integrated Prioritized Target List
LB	Lower Bound
MAAP	Master Air Attack Plan
MSE	Mean Squared Error
OIF	Operation Iraqi Freedom
PCA	Principal Components Analysis
ROC	Receiver Operating Characteristic
ROE	Rules of Engagement
SAM	Surface-to-Air Missile
SD	Strategy Division
SEAD	Suppression of Enemy Air Defense
SNR	Signal-to-Noise Ratio
SPINS	Special Instructions
TAP	Theater Air Planner
TBMCS	Theater Battle Management Core System
TN	True Negative
TP	True Positive
TPW	Target Planning Worksheet
UMD	Unit Manning Document
UB	Upper Bound

Appendix B. Final Database Sample

Table B.1 is a 25 exemplar sample from the final database. Please contact Dr. J. O. Miller via email (john.miller@afit.edu) for the complete database.

Table B.1 Final Database Sample

MSN Type	# AC	MSN Time	TOS	AC Type	# ESM	# SAF	SAF Type	STGS	# ORD	ORD Type	ALT	TGT Type	Alert	Effect
2	2	2.28	1.00	0	0	0	0	0	2	28.00	9000	5.00	1	0.00
1	2	3.22	0.25	14	1	0	0	0	3	13.00	20000	2.00	0	0.00
3	2	1.83	0.50	2	0	0	0	0	1	14.00	17200	0.00	0	0.00
6	1	3.65	2.25	15	0	0	0	0	2	14.00	18000	1.00	0	0.00
2	1	4.50	3.75	15	0	0	0	1	3	19.33	15000	0.00	0	0.00
7	2	5.35	1.58	14	0	0	0	0	2	13.00	12500	2.00	0	0.00
2	2	1.58	0.83	15	0	0	0	0	2	12.00	20000	10.00	1	0.00
2	2	1.92	1.50	14	0	0	0	0	2	28.00	8000	0.00	1	0.00
2	1	6.20	6.08	15	0	0	0	1	2	22.00	15000	0.00	0	0.00
3	2	3.75	0.50	7	0	0	0	0	1	13.00	22000	9.00	0	0.00
2	2	1.58	0.50	15	1	0	0	0	9	19.80	18000	10.00	1	0.00
3	4	1.77	0.17	14	0	0	0	1	3	30.00	13000	2.00	1	0.00
1	2	1.75	0.67	19	1	1	4	0	2	31.00	16000	3.00	0	0.00
3	2	4.58	3.00	19	0	2	1	1	1	28.00	9000	10.00	1	0.00
4	2	4.42	0.50	12	1	0	0	0	1	7.00	20800	4.00	0	0.00
1	2	1.75	0.67	19	1	1	4	0	2	31.00	16000	3.00	0	0.00
6	2	4.13	1.17	8	0	1	1	0	2	12.00	21500	6.00	0	0.00
2	2	2.67	0.75	19	2	2	3	0	1	31.00	17000	2.00	1	0.00
6	2	2.08	0.50	14	0	0	0	0	5	17.20	21400	10.00	0	0.20
1	2	3.52	0.50	14	0	0	0	0	4	13.00	20713	2.00	0	0.25
1	2	3.27	0.08	14	0	0	0	0	4	13.00	13000	3.75	0	0.25
3	2	2.67	0.50	15	0	0	0	0	10	12.80	19000	2.00	1	0.30
7	2	5.95	1.75	14	1	0	0	0	3	9.00	25000	4.00	0	0.33
4	2	4.67	2.00	12	0	0	0	0	3	13.00	18567	4.00	0	0.33
1	2	2.67	0.25	14	0	0	0	0	3	16.67	18000	0.00	0	0.33

Appendix C. MATLAB® Code

```
% Title: Effects of Uncertainty on Real World Aerospace Mission Data
%
% Author: Capt Steven J. Barosko
%
% Abstract:  This Matlab code takes data exemplars and randomizes, standardizes, and rotates
% the data as needed before performing SNR saliency, FFNN, and GRNN multivariate techniques.
%
% Data exemplars:
%   Column 1: Noise Input Feature
%   Columns 2-15: 14 Input Features
%   Columns 16-17: Output Feature Targets
%   Note: Only 1 sample exemplar is shown.
%
D = [0.0029  1  3  1  14.15  3.72  5  0  1  1  0  12  18.00  38000  6.11  0.1  0.9];
%
% Randomize the data.
%
R=[rand(size(D,1),1),D];
R=sortrows(R,1);
%
% Calculate cut-off for training set (80%) and validation set (20%) data exemplar split.
%
C=round(.8*size(D,1));
%
% Separate the data into a training set input feature matrix (X), validation set input
% feature matrix (Y), training set target matrix (Tx), and validation set target matrix (Ty).
%
X=[R(1:C,2:16)];
Y=[R((C+1):size(D,1),2:16)];
Tx=[R(1:C,17:18)];
Ty=[R((C+1):size(D,1),17:18)];
%
% Input feature matrix adjustment used to drop off unused input features.
%
X=[X(:,2:3),X(:,11:13)];
Y=[Y(:,2:3),Y(:,11:13)];
%
% Calculate the correlation matrix of the training set input feature matrix.
%
Cor = corrcoef(X);
%
% Calculate the correlation matrix eigenvalues and eigenvectors.
%
[V,E]=eig(Cor);
%
% Calculate the correlation matrix PCA loadings.
%
Lambdar = diag(diag(sqrt(E)));
Loadr = V*Lambdar;
%
% Rotate the training and validation set input feature matrices with PCA loadings (data may
% be roatated before or after standardization, or not at all).
%
X=X*Loadr;
Y=Y*Loadr;
```

```
%
% Standardize the training and validation set input feature matrices.
%
X=(X-ones(size(X,1),1)*mean(X))*inv(sqrt(diag(diag(cov(X)))));
Y=(Y-ones(size(Y,1),1)*mean(Y))*inv(sqrt(diag(diag(cov(Y)))));
%
% Start of FFNN supervised training (n iterations).
%
n=10;
Zx=[zeros(2,C)];
Zy=[zeros(2,(size(D,1)-C))];
S=[zeros(1,size(X,2))];
W=[zeros(4,size(X,2))];
SNR=[];
for i=1:n
%
% Create a two-layer FFNN: the first layer has 4 LOGSIG nuerodes and the second layer has
% 2 LOGSIG nuerodes.  Uses a resilient backpropogation batch mode gradient update function.
%
    net=newff(minmax(X'),[4 2],{'logsig' 'logsig'},'trainrp');
%
% Train the network for 1500 epochs with MSE goal of 0.05 (Effect cases) or 0.01 (Alert case).
%
    net.trainParam.epochs=1500;
    net.trainParam.goal=0.05;
    net=train(net,X',Tx');
%
% Calculate output activations for the training and (Zx) validation sets (Zy).
%
    Zx=Zx+sim(net,X');
    Zy=Zy+sim(net,Y');
%
% Save the hidden layer weights.
%
    W=W+net.IW{1};
%
% End of FFNN supervised training.
%
end
%
% Average the output activations for the training (Zx) and validation sets (Zy).
%
for j=1:C
    Zx(:,j)=Zx(:,j)/n;
end
for k=1:(size(D,1)-C)
    Zy(:,k)=Zy(:,k)/n;
end
%
% Average the hidden layer weights.
%
for m=1:size(X,2)
   W(:,m)=W(:,m)/n;
end
```

```
%
% Calculate SNRs for each of the input features.
%
for q=1:size(X,2)
    for p=1:4
        S(1,q)=S(1,q)+W(p,q)^2;
    end
end
for r=1:size(X,2)
    SNR(1,r)=10*log10(S(1,r)/S(1,1));
end
%
% Create a two-layer GRNN (this section must be commented out when FFNN results are desired
% and vice versa).  The spread parameter is adjusted from 1.0 to 0.1 in increments of 0.1.
%
SPREAD=.6;
net=newgrnn(X',Tx',SPREAD);
%
% Calculate output activations for the training (Zx) and validation sets (Zy).
%
Zx=sim(net,X');
Zy=sim(net,Y');
%
% Calculate data points for ROC curve. 101 points are calculated for thresholds varying from
% 0 to 1, incremented by 0.1.  The corresponding chance line data points are calculated as well.
%
    n=100;
    ROC=[zeros(2,n+1)];
    DLINE=[zeros(2,n+1)];
    AMD=[];
    for i=0:n
        count1=0;
        count2=0;
        for j=1:(size(D,1)-C)
            if Ty(j,1)==.9
                count1=count1+1;
                if Zy(1,j)>(i*(1/n))
                    ROC(1,i+1)=ROC(1,i+1)+1;
                end
            end
            if Ty(j,1)==.1
                count2=count2+1;
                if Zy(1,j)>(i*(1/n))
                    ROC(2,i+1)=ROC(2,i+1)+1;
                end
            end
        end
        ROC(1,i+1)=ROC(1,i+1)/count1;
        DLINE(1,i+1)=1-(i*(1/n));
        ROC(2,i+1)=ROC(2,i+1)/count2;
        DLINE(2,i+1)=1-(i*(1/n));
    end
```

```
%
% Calculate average metric distance (euclidian) of ROC curve data points from corresponding
% chance line data points.
%
ED=0;
    for i=0:n
        ED=ED+sqrt((ROC(1,i+1)-DLINE(1,i+1))^2+(ROC(2,i+1)-DLINE(2,i+1))^2);
    end
%
% Display AMD.
%
AMD=[AMD,ED/(n+1)]
%
% Save desired output data for Excel spreadsheet analysis.
%
outpt1 = Ty;
outpt2 = Zy';
outpt3 = SNR';
outpt4 = ROC';
save('Insert Path Here\Tgt.xls','outpt1','-ascii','-tabs');
save('Insert Path Here\Val.xls','outpt2','-ascii','-tabs');
save('Insert Path Here\SNR.xls','outpt3','-ascii','-tabs');
save('Insert Path Here\ROC.xls','outpt4','-ascii','-tabs');
%
% Code is finished.
%
disp('Done!')
```

Appendix D. SNR Saliency Summaries

Table D.1 Alert Case SNR Saliency Summary

Features	CA	MSN Type	# AC	MSN Time	TOS	AC Type	# ESM	# SAF
14	0.8736	23.57	9.15	10.45	8.82	9.15	2.97	5.02
13	0.8750	18.76	6.75	9.60	5.50	4.76	3.40	3.14
12	0.8599	21.41	8.98	8.25	5.45	8.18	*0.82*	2.95
11	0.8558	21.72	5.64	9.51	7.39	6.17	N/A	4.95
10	0.8970	21.31	7.07	9.60	8.87	8.32	N/A	*1.80*
9	0.8860	20.05	7.37	7.16	7.89	8.32	N/A	N/A
8	0.8777	19.43	7.16	11.00	6.68	6.52	N/A	N/A
7	0.8571	18.39	6.79	7.79	4.88	6.67	N/A	N/A
6	0.8832	23.59	8.90	12.85	*6.64*	9.13	N/A	N/A
5	0.8832	27.00	13.99	16.74	N/A	13.57	N/A	N/A
4	0.8681	24.04	12.78	18.20	N/A	*9.60*	N/A	N/A
3	0.8640	23.68	*8.85*	15.83	N/A	N/A	N/A	N/A
2	0.8338	19.37	N/A	*18.44*	N/A	N/A	N/A	N/A
1	0.7569	16.74	N/A	N/A	N/A	N/A	N/A	N/A

Features	SAF Type	STGS	# ORD	ORD Type	ALT	TGT Type	Effect
14	7.64	5.45	6.43	4.37	7.27	4.36	*2.85*
13	2.66	*1.79*	6.95	1.96	4.41	2.74	N/A
12	4.78	N/A	8.56	2.05	5.31	4.37	N/A
11	*1.50*	N/A	7.45	2.93	6.38	3.21	N/A
10	N/A	N/A	8.04	4.24	8.52	3.98	N/A
9	N/A	N/A	6.34	2.10	5.94	*0.16*	N/A
8	N/A	N/A	9.05	*3.94*	5.20	N/A	N/A
7	N/A	N/A	3.32	N/A	*3.30*	N/A	N/A
6	N/A	N/A	9.90	N/A	N/A	N/A	N/A
5	N/A	N/A	*13.23*	N/A	N/A	N/A	N/A
4	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Table D.2 Effect .85 Case SNR Saliency Summary

Features	CA	MSN Type	# AC	MSN Time	TOS	AC Type	# ESM	# SAF
14	0.5342	3.54	*-0.45*	2.83	0.29	1.25	1.10	2.53
13	0.5430	4.61	N/A	3.02	2.85	2.34	0.14	-0.36
12	0.5127	6.05	N/A	2.96	3.94	0.76	1.05	0.31
11	0.5293	9.69	N/A	5.95	4.08	*2.11*	5.69	4.34
10	0.5313	4.11	N/A	2.00	2.60	N/A	2.91	0.83
9	0.5361	7.81	N/A	6.19	4.37	N/A	5.67	3.13
8	0.5410	4.84	N/A	4.20	3.49	N/A	3.83	*-1.90*
7	0.5615	4.31	N/A	3.14	3.30	N/A	2.82	N/A
6	0.5479	6.27	N/A	1.07	0.58	N/A	3.14	N/A
5	0.5488	8.43	N/A	*2.57*	2.64	N/A	2.66	N/A
4	0.5635	5.42	N/A	N/A	*-0.29*	N/A	-0.28	N/A
3	0.5645	2.38	N/A	N/A	N/A	N/A	*2.30*	N/A
2	0.5371	*1.85*	N/A	N/A	N/A	N/A	N/A	N/A
1	0.5313	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Features	SAF Type	STGS	# ORD	ORD Type	ALT	TGT Type	Alert
14	-0.30	0.97	8.10	0.22	0.34	0.75	0.26
13	2.84	0.58	9.79	*-1.88*	3.66	1.74	3.10
12	1.77	*-1.19*	9.55	N/A	3.29	3.24	1.50
11	4.16	N/A	11.21	N/A	3.60	4.72	2.84
10	*0.17*	N/A	6.28	N/A	3.20	3.20	2.54
9	N/A	N/A	10.65	N/A	6.00	3.62	*2.37*
8	N/A	N/A	8.27	N/A	0.60	0.14	N/A
7	N/A	N/A	8.48	N/A	*1.21*	2.67	N/A
6	N/A	N/A	6.85	N/A	N/A	*-0.48*	N/A
5	N/A	N/A	7.18	N/A	N/A	N/A	N/A
4	N/A	N/A	8.80	N/A	N/A	N/A	N/A
3	N/A	N/A	8.83	N/A	N/A	N/A	N/A
2	N/A	N/A	6.79	N/A	N/A	N/A	N/A
1	N/A	N/A	-0.74	N/A	N/A	N/A	N/A

Table D.3 Effect .75 Case SNR Saliency Summary

Features	CA	MSN Type	# AC	MSN Time	TOS	AC Type	# ESM	# SAF
14	0.6348	7.57	1.67	4.55	1.72	3.30	2.24	2.35
13	0.6846	7.21	4.06	3.57	2.97	4.29	0.24	2.62
12	0.6426	9.32	4.56	4.36	5.88	6.18	*2.22*	2.67
11	0.6660	7.01	4.01	5.33	*0.54*	5.75	N/A	3.44
10	0.6689	5.10	*-0.24*	4.72	N/A	4.58	N/A	0.69
9	0.6758	6.29	N/A	5.36	N/A	4.23	N/A	*1.81*
8	0.6611	10.71	N/A	3.37	N/A	*0.79*	N/A	N/A
7	0.6641	9.07	N/A	*4.32*	N/A	N/A	N/A	N/A
6	0.6543	10.08	N/A	N/A	N/A	N/A	N/A	N/A
5	0.6748	6.99	N/A	N/A	N/A	N/A	N/A	N/A
4	0.6504	1.69	N/A	N/A	N/A	N/A	N/A	N/A
3	0.6553	2.67	N/A	N/A	N/A	N/A	N/A	N/A
2	0.6602	*1.58*	N/A	N/A	N/A	N/A	N/A	N/A
1	0.6172	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Features	SAF Type	STGS	# ORD	ORD Type	ALT	TGT Type	Alert
14	2.75	5.10	15.53	3.31	3.80	*0.59*	0.77
13	*0.19*	3.26	17.21	3.67	2.18	N/A	2.46
12	N/A	6.45	17.33	6.04	3.40	N/A	3.09
11	N/A	4.99	16.90	2.94	2.24	N/A	3.32
10	N/A	6.51	14.49	0.73	1.45	N/A	0.20
9	N/A	3.23	18.06	4.31	4.33	N/A	4.08
8	N/A	4.87	15.73	3.45	5.09	N/A	4.04
7	N/A	7.59	20.00	4.93	5.80	N/A	6.05
6	N/A	5.71	19.68	5.70	*3.35*	N/A	6.61
5	N/A	*0.59*	12.71	1.52	N/A	N/A	2.23
4	N/A	N/A	10.01	*-0.36*	N/A	N/A	0.97
3	N/A	N/A	16.85	N/A	N/A	N/A	*2.19*
2	N/A	N/A	14.65	N/A	N/A	N/A	N/A
1	N/A	N/A	14.20	N/A	N/A	N/A	N/A

Table D.4 Effect .65 Case SNR Saliency Summary

Features	CA	MSN Type	# AC	MSN Time	TOS	AC Type	# ESM	# SAF
14	0.6738	3.28	-1.58	1.82	1.86	1.03	2.43	1.56
13	0.7021	1.31	*-4.79*	-0.39	0.91	2.19	-0.45	-0.34
12	0.7314	2.46	N/A	2.15	1.71	1.59	*-0.02*	0.05
11	0.6943	2.78	N/A	4.21	0.05	1.90	N/A	-0.63
10	0.7305	4.43	N/A	*-0.73*	-0.11	1.73	N/A	0.87
9	0.7090	6.94	N/A	N/A	*0.81*	3.97	N/A	1.86
8	0.7139	6.04	N/A	N/A	N/A	4.80	N/A	0.53
7	0.6826	9.87	N/A	N/A	N/A	7.95	N/A	*2.68*
6	0.6963	2.65	N/A	N/A	N/A	2.97	N/A	N/A
5	0.6992	8.12	N/A	N/A	N/A	8.04	N/A	N/A
4	0.7051	5.66	N/A	N/A	N/A	4.78	N/A	N/A
3	0.6904	3.92	N/A	N/A	N/A	*3.86*	N/A	N/A
2	0.6953	*-2.78*	N/A	N/A	N/A	N/A	N/A	N/A
1	0.6475	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Features	SAF Type	STGS	# ORD	ORD Type	ALT	TGT Type	Alert
14	1.11	*-2.14*	15.27	2.50	-0.40	3.02	1.49
13	-0.99	N/A	13.07	3.88	-3.14	0.81	-0.71
12	0.75	N/A	11.64	0.36	1.93	0.81	0.53
11	4.49	N/A	12.51	0.66	*-1.39*	1.77	2.54
10	0.07	N/A	16.54	1.67	N/A	0.35	2.17
9	1.34	N/A	15.23	2.69	N/A	2.98	3.70
8	2.22	N/A	14.81	4.44	N/A	*0.25*	4.63
7	4.95	N/A	17.71	4.69	N/A	N/A	6.80
6	*-1.61*	N/A	12.51	1.33	N/A	N/A	1.44
5	N/A	N/A	17.08	6.76	N/A	N/A	*6.05*
4	N/A	N/A	14.76	*2.73*	N/A	N/A	N/A
3	N/A	N/A	10.47	N/A	N/A	N/A	N/A
2	N/A	N/A	9.62	N/A	N/A	N/A	N/A
1	N/A	N/A	3.55	N/A	N/A	N/A	N/A

Bibliography

1. Alsing, Stephen G., Kenneth W. Bauer, and Mark E. Oxley. "Convergence for Receiver Operating Charateristic Curves and the Performance of Neural Networks," *Smart Engineering System Design*, 4: 133-145 (2002).

2. Bauer Jr., Kenneth W., Stephen G. Alsing, and Kelley A. Greene. "Feature screening using signal-to-noise ratios," *Neurocomputing*, 31: 29-44 (2000).

3. Burns, Robert. "To a Mouse." *Poems and Songs.* Vol. VI. The Harvard Classics. New York: P.F. Collier and Son, 1909-14.

4. Demuth, Howard and Mark Beale. *Neural Network Toolbox User's Guide, Version 3.0.* The MathWorks Inc., 1998.

5. Department of the Air Force. *Organization and Employment of Aerospace Power.* AFDD 2. Washington: HQ USAF, 17 February 2000.

6. Department of Defense. *Doctrine for Joint Operations.* Joint Pub 3-0. Washington: Joint Chiefs of Staff, 10 September 2001.

7. Department of Defense. *Command and Control for Joint Air Operations.* Joint Pub 3-30. Washington: Joint Chiefs of Staff, 5 June 2003.

8. Dillon, William R. and Matthew Goldstein. *Multivariate Analysis: Methods and Applications.* New York: John Wiley and Sons, 1984.

9. Egan, J. P. *Signal Detection Theory and ROC Analysis.* New York: Academic Press, 1975.

10. Looney, Carl G. *Pattern Recognition Using Neural Networks: Theory and Algorithms for Engineers and Scientists.* New York: Oxford University Press, 1997.

11. McAree, Paul W. "RE: Analysts in the AOC Information." Electronic Message. 11:08AM, 12 November 2003.

12. McLean III, Charles F. "Operations Analysis Software and Information Services (OASIS): An Unsolicted Proposal Based on the Lessons Learned from Operation Iraqi Freedom." Unpublished White Paper. Intelligent Software Solutions Branch Office, Hampton VA, 30 May 2003.

13. McLean III, Charles F. "Operation Iraqi Freedom: Effects Based Targeting Wartime Systems Architecture Observations and Lessons Laerned." Microsoft Powerpoint Briefing. Intelligent Software Solutions Branch Office, Hampton VA, 3 Nov 2003.

14. Myers, Gene. "Effects-Based Operations: Everything old is new again, as concept reveals," *Armed Force Journal International*, 140:47-49 (June 2003).

15. Specht, D. F. “A general regression neural network,” *IEEE Transactions on Neural Networks* 2(6): 568-576 (1991).

16. Von Clausewitz, Carl. *On War*, Michael E. Howard and Peter Parent, eds. Princeton, NJ: Princeton University Press, 1976.

17. Wackerly, Dennis D. and others. *Mathematical Statistics with Applications* (6th Edition). Australia: Duxbury, 2002.

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

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From – To)*
05-03-2004	**Master's Thesis**	Jun 2003 – Mar 2004

4. TITLE AND SUBTITLE	
EFFECTS OF UNCERTAINTY ON REAL WORLD AEROSPACE MISSION DATA	**5a. CONTRACT NUMBER**
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	**5d. PROJECT NUMBER** ENR2003-078
Barosko, Steven, J., Captain, USAF	**5e. TASK NUMBER**
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(S)	**8. PERFORMING ORGANIZATION REPORT NUMBER**
Air Force Institute of Technology Graduate School of Engineering and Management (AFIT/EN) 2950 Hobson Way, Building 642 WPAFB OH 45433-7765	AFIT/GOR/ENS/04-01
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	**10. SPONSOR/MONITOR'S ACRONYM(S)** AFSAA
Air Force Studies and Analyses Agency 1777 Kent Street Rosslyn VA 22209	**11. SPONSOR/MONITOR'S REPORT NUMBER(S)** N/A

12. DISTRIBUTION/AVAILABILITY STATEMENT
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

13. SUPPLEMENTARY NOTES
N/A

14. ABSTRACT

Uncertainty is an age old element of warfare. How does one measure success or failure on the battlefield? This thesis explores the effect of uncertainty on real world aerospace mission data from Operation Iraqi Freedom (OIF). Sources of uncertainty are identified during data preparation and the effects of uncertainty are investigated using multivariate analysis techniques. Two distinct cases are analyzed: one case with a low level of uncertainty in predicting whether or not a mission is pre-planned or alert generated, and another case with a high level of uncertainty in predicting whether or not a mission is a success or a failure. Three multivariate analysis techniques are used on both cases: Signal-to-Noise Ratio (SNR) saliency, Feed Forward Neural Network (FFNN) supervised training, and General Regression Neural Network (GRNN) supervised training. Measures of performance are gathered from each of these techniques and analyzed to determine the effect of uncertainty on the data.

15. SUBJECT TERMS
Operation Iraqi Freedom, Signal-to-Noise Ratio, Feed Forward Neural Network, General Regression Neural Network, Multivariate Analysis, Uncertainty, Aerospace Operations Center, ROC Curves

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON** Dr. John O. Miller, AFIT/ENS
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			**19b. TELEPHONE NUMBER** *(Include area code)*
U	**U**	**U**	**UU**	103	(937) 255-6565, ext 4326; e-mail: john.miller@afit.edu

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39-18