METHODS FOR MODELING TEAM PERFORMANCE: A

CASE STUDY USING A DECISION SUPPORT SYSTEM

By

Erwin Alexis Baylot, Jr.

A Thesis
Submitted to the Faculty of
Mississippi State University
in Partial Fulfillment of the Requirements
for the Degree of Master of Science
in Engineering
in the Department of Industrial Engineering

Mississippi State, Mississippi

May 2000

UMI Number: 1399437

UMI

UMI Microform 1399437

Bell & Howell Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

METHODS FOR MODELING TEAM PERFORMANCE: A

CASE STUDY USING A DECISION SUPPORT SYSTEM

By

Erwin Alexis Baylot, Jr.

Approved:

Dr. Allen G. Greenwood
Associate Professor of Industrial Engineering
(Director of Thesis)

Dr. Stanley F. Bullington
Professor of Industrial Engineering
(Graduate Coordinator of the Department of Industrial Engineering and Committee Member)

Dr. Royce O. Bowden
Associate Professor of Industrial Engineering
(Committee Member)

Dr. A. Wayne Bennet
Dean of the College of Engineering

Name: Erwin Alexis Baylot, Jr.

Date of Degree: May 13, 2000

Institution: Mississippi State University

Major Field: Industrial Engineering

Major Professor: Dr. Allen G. Greenwood

Title of Study: METHODS FOR MODELING TEAM PERFORMANCE: A CASE STUDY USING A DECISION SUPPORT SYSTEM

Pages in Study: 130

Candidate for Degree of Master of Science

As a tool for process analysis, a procedure was developed to model and relate team in-process performance to the end-product while taking into account the human interactions of a team, and the variability of subject-matter-expert judgment. This was accomplished by combining a modified Analytical Hierarchy Process (AHP) with either a Multi-Variant Linear Regression (MVLR) or an Artificial Neural Network (ANN). Using the data obtained from a decision support system experiment with the U.S. Army Military Intelligence Officer Advanced Course students, a case study was used to develop a procedure for using the identified methods. Although the findings yielded little statistical difference in modeling performance of MVLR and ANN, the utility of such models was successfully used as a tool by helping to distinguish the characteristics and methods exhibited by the high performance versus the low performance teams.

TABLE OF CONTENTS

LIST OF TABLES

TABLE Page

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Background

During the last two decades, society has moved into what is becoming known as the information age. This information age was brought on by the technological revolution that has spawned the proliferation of inexpensive and powerful microchip-based computer systems. The commercial, private, governmental, and military sectors of society have all been touched by the information age, and thus changes have resulted.

Also occurring during this information age has been the downsizing of businesses and government due to global competition and the requirement to increase efficiency and improve the quality of service. Recent methods developed for process improvement have been employed to improve quality and at the same time to increase productivity. The basic premise is that if you improve the process, then there will be some corresponding increase in quality, thereby reducing the need for such non-value adding tasks as product testing, rework, and/or disposal. Resources used to perform processes range from automated machinery to strictly human. To improve productivity, engineers and scientists have developed numerous ways to gather information, but effective utilization of information to improve team performance is still an issue.

Problem Statement

A difficult process to improve is one involving a team of specialists developing a unique product. While the idea of using teams has been around for a long time, it has only been within the last ten years that this concept has gained popularity as a means to increase value to a process and/or increase efficiency. As with any method of process analysis such as value added analysis, the method must be measured and somehow evaluated. According to experts in this area (Meister 1986), there are insufficient methods to measure team performance due to the inherent dynamics. When modeling a team process, such as knowledge-based teams using decision support systems, analysts are sometimes unsure as to how to systematically evaluate process performance when the only certain element of the process is what the end-product must be. What is needed is a procedure to mathematically relate the in-process results to the end-product, taking in account the variability of team interactions. This thesis will address and develop such a procedure.

Purpose and Scope of Report

A comprehensive literature search was conducted in order to fully document the basis of the problem, the nature of the problem, and the approach to solve the problem. Two candidate methods, Multi-Variant Linear Regression and Artificial Neural Networks, were both found in the literature as being used for modeling human performance. The determination as to why these were selected methods, and why the Analytical Hierarchy Process was introduced into the procedure, are presented. The preferred method and procedure were discussed and data gathered from a live experiment demonstrated the procedure. Additionally, areas for further study will be identified.

CHAPTER II

LITERATURE REVIEW

This chapter presents an overview of the limited body of literature dealing with evaluating team performance. Emphasis is on recent literature since Meister's (1986) review indicated there is little information available prior to 1986. Meister noted that even though there was an extensive amount of available literature on performance measurement, there is virtually nothing on measuring team performance because of the uncertainty of what to measure. Team performance research before and during 1986 was focused on discovering what measures are important. The limited amount of published literature was also noted to be the case through 1999 as well, and will be presented in the forthcoming sections.

The following sections are intended to guide the reader through the development of the basis of the problem by providing background on improving effectiveness and value-added analysis to establish the premises for the basis of the problem and to familiarize the reader with topic areas that could benefit with the application of the procedure developed in this thesis. Candidate techniques for evaluating team performance were reviewed to provide a basis for relating team performance to end-product results. The literature review will define the basis of the problem and how the scope was determined.

Developing the Basis of the Problem

This section provides an introduction as to what events and changes have occurred in recent years that have brought about the need for advances in analyzing team performance. Emphasis is given to improving effectiveness within an organization and value-added analysis on team performance.

Improving Effectiveness

Corporations, as well as governments, are looking at re-engineering and other methods to cut costs while improving the effectiveness of the organization. Downsizing has produced reductions in staff across these organizations. While reducing staff certainly will reduce the payroll costs, effectiveness and increased production have not been automatic. In an effort to maintain or increase effectiveness of an organization, the team approach has gained interest in the corporate and government organizations (Escover 1994). The age-old notion that "two heads are better than one" comes into play. The old philosophy that it is impossible to accomplish low production cost while still maintaining high quality and flexibility of process has changed in recent years. Organizations that continue with the old philosophy either fail or make the needed adjustments. Companies are now interested in two questions; "Are the right things being done?" and "Are they being done well?" (Vokurka and Fliedner 1995).

It is well known that the role of the computer has provided a means to make improving efficiency and productivity in all aspects of life. Information is quite abundant and can sometimes be intimidating to the knowledge worker (Robinson 1991).

Requirements are being set for building decision support systems that will "smartly" take the abundance of information and "brilliantly" package the information for use by the decision maker. (U.S. Army 1996, Robinson 1991).

Acquisition of better technology to increase productivity should be accompanied by a commitment to base the acquisition on an integrated systems approach which takes into account the human and social aspects of automation as well as the computer systems at hand (Osborne and Rosenthal 1985). This total systems approach for acquisitions is based on long-term objectives and methods for measuring performance, (Vokurka and Fliedner 1995).

Value-Added Analysis on Team Performance

A value-added function has been defined simply as "a function which must be performed in order to produce a product or service" (Boza et al. 1990). Value-Added Analysis is an analytical approach to reviewing the merit of the elements of a process and hopefully eliminating the cost and complexity of unnecessary elements and improving the necessary elements. Emphasis is given to improving the elements that produce the product rather than those that do not actually produce the product (e.g., evaluate the quality or track the progress of the product).

A recent study has found that in the area of manufacturing, the average production worker in the United States only spends about 25 percent of his or her time adding value to a product while the remainder of the time is spent on nonproductive activities that do not actually add value to the product (Bradyhouse 1982). Bearing this in mind, it is not hard to see that someone trying to increase productivity and only spending their efforts on 25 percent of the problem will have diminishing returns. Much more emphasis has come to bear recently on reducing the amount of time spent on nonproductive activities. With this shift in focus, anything that does not provide value, risks elimination (Escover 1994).

An approach referred to by Boza et al. (1990) as "Process Engineering" requires a systematic analysis approach to determine the value added (VA) by each function. A team performs a process review with the product being either material, information, or service. According to Boza et al. (1990), Process Engineering has the two sub-fields, referred to as Process Re-engineering and Continuous Process Improvement. Process Re-engineering calls for drastic changes to processes and Continuous Process Improvement calls for gradual, but steady changes. This is somewhat over simplified because using Continuous Process Improvement may eventually require complete removal of a significant element or addition of a significant element within a process, but is a suitable explanation for the purposes of introducing these two sub-fields. To illustrate Process Re-Engineering and Continuous Process Improvement, an example is given in Figure 1. The process is illustrated using several steps for producing a car: design, testing the design by simulation, building and field testing a prototype, and then manufacturing the car after the successful completion of the latter elements of the process. The value-added elements are the design and manufacturing of the car. The non-value added elements are simulating the prototype car and building and field testing the car. They are non-value added elements because they are for testing and exposing design flaws. For this example, Continuous Process Improvement would probably work to reduce the number of prototype testing cycles for both simulation and field testing. Processing Re-Engineering in this example would call for eliminating the most resource intensive form of testing, the building of and field testing of a prototype.

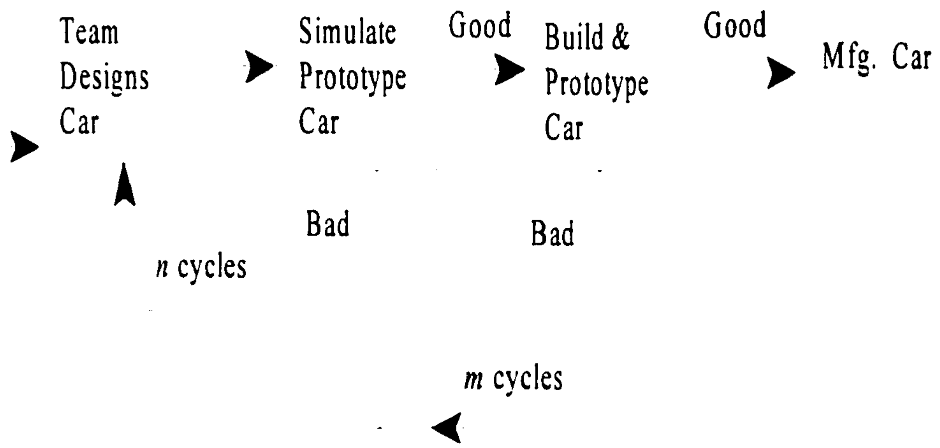


Figure 1 Simplified diagram of a process to design and manufacture a car illustrating value added and non-value added elements.

As stated in Boza et al. (1990), simply identifying and eliminating non-value added functions will not necessarily improve the process. Some non-valued added functions exist to accommodate poorly performing value-added functions. Thus, non-value added functions cannot necessarily be eliminated until performance is increased for value-added functions. Boza et al. (1990) stated that a poorly performing value-added function often is the cause for a non-value added function(s). Eliminating such non-value added functions will only make the process perform worse. Thus, we can conclude that improving the poorly performing value added function, which is responsible for the non-value added function, is the key to the eventual removal of the non value-added function.

But even with the two described methods, there is still great difficulty in measuring the productivity of the value-added activities, especially in automating a process that previously had been a manual system. Automated and/or computer-based resources available to a decision maker can enhance their ability to make complex decisions, but are often difficult to measure because there is no basis for comparison. It has proven difficult to quantify the professional's or knowledge worker's productivity (Osborne and Rosenthal 1985). Because of the knowledge worker's intelligence and flexibility to problem solve, he or she introduces a degree of complexity into measuring end-product results. Furthermore, the subject-matter-expert evaluating the final product is often at some level of management (e.g., commander of an army maneuver unit, as will be shown in the case study) and little research has been developed that effectively records the judgements of these subject-matter-experts and determines the validity of their judgements (Meister 1985).

According to Meister (1985), "None of our methods of evaluation is specifically team oriented . . . The special problem for behavioral evaluation is to find ways of relating personnel subsystem process and outward to the total system output. This is the great void in the evaluation of behavioral effectiveness." Thus, the lack of well developed measuring tools will make it especially difficult for the knowledge team (Baker and Salas 1992, Williges et al. 1992). Not only is there a need for understanding the dynamics of team work, but understanding the subject-matter-expert's judgement and its effectiveness to measure team performance is suspect (Osborne and Rosenthal 1985).

Developing an Approach

This section of the literature review is intended to guide the reader through the development of the approach to solve the problem discussed in the previous section. First, a background is provided to establish the general techniques for measuring performance. Next, a discussion is given on the importance of having the proper measures and then how best to obtain those measures. A discussion of screening studies is provided to introduce the reader to handling and reducing the volume of data prevalent in team performance studies. Finally, the applicability of Multi-Variant Linear Regression and Artificial Neural Networks as modeling paradigms and the Analytical Hierarchy Process for comparing the end product of team performance are discussed.

Background

One literature source indicated not to look within the team at the individuals or members, but to look at the effect on the process at the aggregate team level (Osborne and Rosenthal 1985). However, Meister (1986) suggested that team performance can be measured four possible ways: (1) the individual performances of team members; (2) the team output is measured without individual member performance considered; (3) the team output is measured as defined in "2," but in addition, the output measure is related to the total system output, the team being part of the larger system; and (4) all three levels (individual, team, system) with the goal to relate individual member performance to system output. The fourth approach is the ideal circumstance since it considers all of the previous methods. This thesis uses this approach.

Measures for the Knowledge Worker

Another area of concern when selecting measures is subjective versus objective and qualitative versus quantitative. According to Osborne and Rosenthal (1985), qualitative measures are more reliable at higher hierarchical levels since individual variances tend to balance out, whereas quantitative measures are easy to obtain at lower levels.

Vokuraka and Fliedner (1995) explained that many companies still rely on the "bottomline" measures of total cost and efficiency to indicate performance, but that time, quality, and service performed would be more appropriate. Moreover, the primary benefits to the knowledge worker are, "not necessarily speeding up the information flow, but in improving the depth of analysis and understanding of the available information." (Osborne and Rosenthal 1985). Osborne and Rosenthal (1985) further stated that the knowledge worker or professional staff productivity should be measured in terms of the information handled and processed." While this is ideally stated, it is not always obtainable due to resource limitations. This thesis research is based on readily obtainable quantitative measures of time, quantity, and the more subjective measures of quality. Generally speaking, measures that can be used across research and test settings lead to standardization and further aid in comparing results across such settings (Muckler and Seven 1992)

Measuring Team Performance

A study (Pasmore 1993) indicated that measuring teamwork activities is best performed by on-site observers. The team members themselves are very limited in judging their own or each other's performance. Morgan et al. (1986) developed the Critical Team Behavior Form (CTBF) to measure teamwork skills in tactical decision-making. It was found in a follow on study (Baker and Salas 1992) that the CTBF was successfully used to gauge the performance of 13 tactical-decision-making teams. The study found that high-performing teams demonstrated sixty-six percent more effective behaviors (appropriate actions) than low performing teams. Additional research studies (Muckler and Seven 1992) support the idea that on-site observers perform best at monitoring teamwork activities and further states that observers should be selected so that they will have minimal impact on the team activities. Furthermore, interaction must be minimized because too much interaction can unintentionally bias results. Although bias can be accounted for in an analysis, collecting data for team performance is resource intensive and sufficiently large samples needed to account for a bias may not be practical.

Screening Studies

As with any study where a large set of variables come into play, a screening study is recommended. The sheer amount of data collected for a study of team work activities, especially activities that include automation systems, can be overwhelming. The value of screening studies lies in the capability to assess a large number of variables with a relatively small number of observations (Beaudet and Williges 1988) rather than depending solely on the experiences of human factors practitioners (Kantowitz 1992). Measures must be well-defined and meet re-test criteria to insure reliability (Bittner 1992) and be suitable for cross comparison against other experiments.

Although screening studies have proven to be invaluable, literature on the use in human factor research is limited. Beaudet and Williges (1988) suggest using several screening techniques and comparing results to determine common correlations between variables. Moreover, when selecting variables to analyze using any technique, conduct a post-selection review to ensure that the given variables can be observed under desirable experimental conditions and was performed for the case study described within this text. The screening techniques used in this thesis are discussed in the next chapter, entitled Procedure.

Data Analysis and Modeling Paradigms

Multi-Variant Linear Regression (MVLR)

A MVLR model is one that uses MVLR to fit and subsequently forecast/predict the results of a process based on a set of variables and a linear relationship. The general form of the linear regression model, with normal error terms, and in terms of X independent variables is given in Equation 1. (Neter et al. 1989):

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_{p-1} X_{i,p-1} + \epsilon_i \qquad \textbf{(1)}$$

where: β are parameters, X are independent variables, and ϵ is an independent error term that is assumed to be normally distributed with a mean of zero and variance of σ^2. The values of i range from 1 to n observations.

The parameters β are sometimes referred to as the partial regression coefficients because they are the partial effect the corresponding independent variable will have on the value of Y when all other independent variables are held constant. MVLR and in general multiple regression, parcels out the least mean-squared error for each independent variable and then selects a regression function Y, that best fits the data based on the least amount of total squared error (Lykins and Chance 1992).

Although there is an abundance of research literature that documents the use of MVLR, no literature was found that showed the use of MLVR for modeling the performance of a team process. A paper by Lykins and Chance (1992) and another by Eksioglu et al (1996) document comparisons between a MVLR model and Artificial Neural Network models, and a paper by Charlton (1992) described a method to relate three forms of data; questionnaire data, in-process data, and end-process results. These approaches proved useful in developing the procedure described in Chapter IV of this thesis.

The Lykins and Chance study demonstrated that Artificial Neural Network models statistically performed better at forecasting/prediction than MVLR. However, in the study, only one data set was used in the analysis and thus no generalizations could be drawn as to whether Artificial Neural Networks would perform better than MVLR or vice-versa.

In Charlton's study (1992), the first step was to identify human factors as predictors of systems performance. The experiment assessed human information processing and memory capabilities required for computer systems in order to monitor space control systems. This experiment had the added challenge of not allowing interruption of the operator process. The experiment took a systems viewpoint in which human factor design aspects would lead to mission effectiveness. Analysis of data was performed with a small sample size that used multiple regression to determine the significant predictors of system performance. The findings indicated in his case study that the questionnaire data was a significant predictor of both system performance and operator performance.

Artificial Neural Networks (ANN)

ANN forecast/predict the results of a process based on multiple independent variables that are related in typically highly non-linear relationships. The varying functional relationships and equation descriptions for the different architectures of ANNs are quite extensive and are beyond the scope of this thesis. The reader is advised to refer to such texts as Haykin (1999) and Zurada (1992) for more detailed descriptions.

The explosion of interest in ANN has come about due to the many successful applications that have emerged over recent years. Statistical approaches require an analyst to determine how output data relate to input data, but an ANN does not require such (Kosko 1992). In one case, an example statistical approach using MVLR by Klimasauskas (1992), yielded a result that was correct only 85 percent of the time. Conversely, an ANN yielded a correct result 92 percent of the time. Although other studies could show the converse, the interesting point to note from the documentation was that most of the poor predictions yielded in the statistical approach were due to the analyst's lack of understanding of statistics, and thus a poor implementation of the methodology. This is a real problem! But, because the ANN required less understanding of statistics and was more of a "black-box," the analyst did a better job of modeling the system with ANN. This is not to say that an analyst shouldn't have a command of statistics, but like the human brain, ANN can recognize data patterns that remain undefinable. Kosko (1992) refers to this ability as "recognition without definition."

The non-linear behavior of people in teams lends itself to non-linear modeling. Furthermore, when subject-matter-experts are used to subjectively evaluate team performance, they often can't quantify why they give a particular score or rating, they just do. Thus, ANN is a logical approach to modeling the variability of a team process.

There are numerous types of ANN with many variations of each type. Some types of ANNs lend themselves to better solving the problem proposed than others. The literature cited (Kosko, 1992; Klimasauskas, 1992; Lykins and Chance 1992; Eksioglu et al, 1996), points toward Back-Propagation, or some variation as the one of choice. Back-Propagation offers the advantage of being a general purpose technique that attempts to minimize global error and can accommodate multi-dimensional functions. Disadvantages include possible need for a large data set or slow learning. Slow learning can be remedied with a high performance computer, but reducing the number of samples or factors measured in a data set may not provide enough data to properly train the ANN.

Back-propagation ANNs using the least-mean-square algorithm to minimize error assumes that all independent variables (processing units/elements) contribute to the error and therefore propagate the output error backward through the network. The ANN is trained by propagating the input forward through the network to the output layer, compares the predicted output to the actual output, and then determines the amount of error. The error is then propagated back again until a desirable root-mean-square error is obtained. The ANN "learns" how to behave by using a gradient descent rule which changes each weight (value of coefficient) based on the size and direction of the negative gradient on the error surface (Lykins and Chance 1992).

Analytic Hierarchy Process (AHP)

The AHP is a method that can be used to combine measures-of-effectiveness to compare the end-process results of team performance. It is a popular method for evaluation of decision alternatives in both government and industry (Buede and Maxwell 1995). It was first addressed in the late 1970's as a decision support methodology. It has been widely used in qualitative and quantitative analysis. The method is sometimes used to convert qualitative factors to quantitative scales (Saaty 1990). AHP requires subject-matter-experts to make pairwise comparisons of the various factors of a given problem. When the number of factors is too great (more than seven), a hierarchy formulation of the problem is recommended to make it more manageable.

Table 1 is an example scale that a subject-matter-expert (SME) or evaluator might use for determining the relative importance of performance factors using the AHP procedure.

Table 1. AHP Subjective Scale for Pairwise Comparisons

1 - Base factor roughly equivalent importance to other factor
3 - Base factor moderately more important than other factor
5 - Base factor essentially more important than other factor
7 - Base factor is much more important than other factor
9 - Base factor is overwhelmingly more important than other factor

The generalized form of AHP is represented as follows. C_{ij} is a value of a pairwise comparative relationship between base-factor *i* and other-factor *j* for *n* number of base-factors and n^2 number of comparisons for a scale of "1" to some positive endpoint *m*. A value of "1" would indicate equivalence and a value of *m* would indicate overwhelming importance of the base-factor to the other-factor. Comparisons are only necessary when *i* is greater than *j*. When *i* is equal to *j*, the base-factor and other-factor are one in the same, thus $C_{ii}=1$. Table 2 illustrates the expanded matrix form of these relationships.

Table 2. General matrix form of AHP pairwise relationships

Other Factor *j*						
		1	2	3	...	*n*
Base-Factor *i*	1	1	C_{12}	C_{13}	...	C_{1n}
	2	$1/C_{12}$	1	C_{23}	...	C_{2n}
	3	$1/C_{13}$	$1/C_{23}$	1	...	C_{3n}
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	1	$\vdots$
	n	$1/C_{1n}$	$1/C_{2n}$	$1/C_{3n}$	...	1

Although AHP is mathematically robust, it assumes linearity when there may be no evidence of linearity, and furthermore it becomes increasingly difficult to apply when more than one SME is involved. However, the procedure developed in the next section will discuss how this was handled with a modification.

Conclusions of Review

Based on the literature review, the following premises are apparent.

- With the advent of re-engineering, business, industry and government need methods to examine the way processes are designed, performed, and evaluated.
- The increased emphasis in team work has added value to processes and increased the capability of an organization to perform its task.
- Value-Added Analysis is a means to increase effectiveness of a process by eliminating non-value added tasks, but processes performed by teams are complex, dynamic, and difficult to analyze.
- Knowledge worker based teams require subject-matter-expert judgement to evaluate the end-product and performance is difficult to relate back to the elements of the process.
- MVLR and ANN, with either coupled with AHP, are good candidate modeling paradigms for predicting team performance.

What is needed to perform process analysis for knowledge worker based teams is a procedure to model and relate the in-process team performance to the end-product while taking in account team interactions and the variability of subject-matter-expert judgement.

CHAPTER III

PROCEDURE

The need and applicability to business, industry, and government for a procedure to evaluate team performance has been stated in the literature review. Furthermore, the tools available to construct a procedure have been discussed in the Approach section of the literature review. However, the direct application of such a contribution may not be clearly apparent. This thesis will first solve a real problem and expand the applicability to other potential problems.

This section describes the procedure undertaken to accomplish the objective of this thesis. The literature review and characteristics of the selected case study, suggested that the application of MVLR, ANN and AHP held considerable promise to accomplishing the goals of this thesis stated in the Problem Statement. Like the studies found in Lykins and Chance (1992) and Eksioglu et al (1996), MVLR and ANN were compared to determine which modeling paradigm was most suitable for analysis of human response, but will be extended to a team process. The end-product evaluations performed by the subject-matter-experts (SME) were computed with the AHP to derive and combine the evaluation scores into a single factor to train with the MVLR and ANN. Figure 2 graphically portrays the procedure employed to develop the two models and order of discussion within this chapter following the selection of the case problem. Each box label in Figure 2 is synonymous with the section headings following the description of the case study.

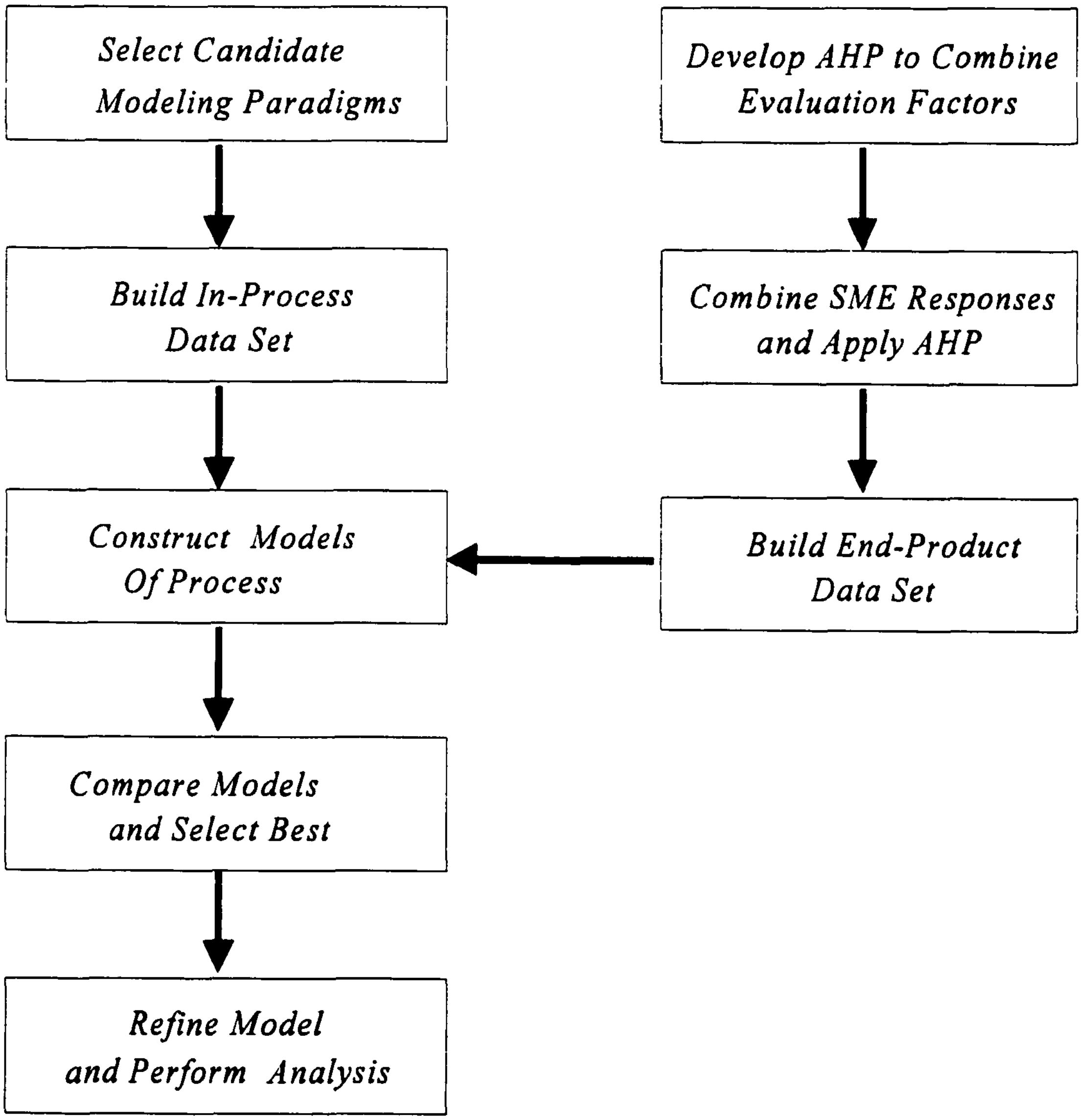


Figure 2 Diagram of model development methodology

Description of Case Study

To compare MVLR and ANN as modeling paradigms for a team process, a data set was needed that would be representative of various team processes evaluated by SME judgement to demonstrate the value of the thesis findings. Such a data set was found in the study by the U.S. Army Engineer Research and Development Center that reported the value added for introducing computer-assisted procedures into the Intelligence Preparation of the Battlefield (IPB) process (Deliman et al. 1997).[1]

Like business and industry, the U.S. Army is using decision support systems and will continue to introduce such systems in the future. Downsizing of U.S. military forces has necessitated the need to win the "information war" (U.S. Army 1996). The study was sponsored by the Office of the Deputy Chief of Staff for Intelligence, Department of the Army, and was conducted to assess time savings, improved quality, and prioritization of future automation efforts at the brigade organizational level and below. It will hereafter be referred to as the IPB Study.

To assess the utility and benefits of automating mobility-related functions (vehicle routing, etc.) and products, quantitative measures were devised to compare performance

---

[1]The IPB integrates enemy doctrine with the weather and terrain to determine and evaluate enemy capabilities, vulnerabilities, and feasibility of enemy courses of action. IPB products support commanders and their staffs in the decision making process. IPB provides a graphic intelligence estimate that portrays probable enemy courses of action and a graphic operations order (U.S. Army 1989).

of automated and manual brigade staffs.[1] Brigade staffs qualify as a knowledge team, an important consideration for this study (Robinson 1991). Measurements were gathered using controlled experiments and questionnaires. The Military Intelligence Officer Advanced Course (MIOAC) at Fort Huachuca, Arizona was selected as the test environment. Two experiments were conducted in conjunction with the MIOAC. A team of typically nine officers simulated a brigade staff performing the IPB process in support of the commander's (student evaluator) decisions. Six teams were involved in each experiment. Three performed the process by manual means and three performed the process using a government developed decision support system. These teams using the decision support system were known as Computer-Assisted Squads (CASs). The interactions of more than 100 officers were studied.

Any further discussion of this study is viewed as inconsequential to the development of the objectives of the thesis. Therefore, the specifics of the study and how the data were gathered is documented in Appendix A and further in a Technical Report (Deliman et al. 1997). Appendix B provides a list of terms and definitions to help with understanding the terminology used in the case study description and data discussion.

---

[1]A brigade staff is the planning and operations element for the Army brigade commander. The commander is a full colonel and his staff officers range in rank from captain to major. The staff includes, but is not limited to an Executive Officer, Adjutant Officer, Operations Officer, Intelligence Officer, Logistics Officer, Engineer Officer, Fire Support Officer, Assistant Operations Officer, Assistant Intelligence Officer, Non-Commissioned Officers, etc.

Select Candidate Modeling Paradigms

The first step in this procedure as indicated by Figure 2, was selecting the candidate modeling paradigms. The literature review as indicated by Lykins and Chance (1992) and Eksioglu et al (1996) showed that MVLR and ANN held promise for developing candidate models for a knowledge team process. Initial sample training sets were analyzed with MVLR in the software package JMP 3.1 and with the default settings (e.g. convergence criteria) for Backpropagation ANN in the software package NeuralWare Professional II/Plus (SAS Institute, Inc. 1995, NeuralWare 1995b). It was found from the provided measures of performance from the case study that these two paradigms were indeed good candidates.

Build In-Process Data Set

From the literature review, it was anticipated that three categories of data are required. They are (1) characteristics of team members, (2) duration of in-process elements and number of team members involved, and (3) evaluation scores of end-products. The initial training data set gave emphasis on the latter two, however a better predictive model was developed after the introduction of category one. Thus the in-process data set was composed of categories one and two.

As with any analysis, a sufficiently large data set with adequate replications is desirable to insure robustness. The experimental data from the case problem for Mission Analysis was sufficiently large, with ample replications to provide a demonstration.

The characteristics of team members (category 1) were obtained from the MIOAC questionnaires. The questionnaire form is found in Appendix C and the Mission Analysis Evaluation Form is found in Appendix D. The findings for the individual team members were grouped by their squad identifier and averaged. Average team characteristics were used to correspond with the team's in-process performance data and end product data. The complete data set for the initially selected characteristics are provided in Appendix E, Table E1 (Mission Analysis Squad Average Characteristics). The column headings for Table E1 are synonymous with the questions found in Appendix C. A corresponding set of data for in-process performance (category 2) is provided in Appendix E, Table E2 (Mission Analysis In-Process Performance Parameters and Values).

Practical Exercise (PE) 1 was conducted in the months of March and May 1996, while PE 7 was conducted in the months of April and June 1996. The monthly indicators shown in the previously named tables identify which PE was being conducted. The membership of the teams remained constant from PE 1 to PE 7, but the role of the Computer-Assisted Squads and Manual Means Squads were exchanged. In May, a new set of students replaced the others and the experiment was replicated. Although by design, PE 1 and PE 7 used the Mission Analysis process to analyze different, but similar situations, they yielded the same product. The IPB study revealed no statistical difference in the evaluation scores due to the PE (Deliman et al. 1997). Therefore, all data from both PEs where treated as being from the same population. The evaluation scores (category 3) given by the evaluators for each squad's Mission Analysis "product" are discussed later, in the *Build End-Product Data Set* section of this thesis on page 35.

As with many experiments, data collection errors are a possible source of apparent error and the method or collection means should be reviewed when outlier data is evident. Sometimes such outlier data is apparent to the trained observer while at other times it must be parsed by an analysis. The case study showed such an error and it was discovered that the observer had forgotten to stop the clock when recording the time for an in-process element. This sample from the data set (Squad 2, April) was eliminated from the analysis. Thus 23 samples of the 24 were used for this thesis research, and according to the JMP software requirements, at least 20 samples were needed to provide a robust set of statistics for a MVLR (SAS Institute, Inc. 1995).

In order to test and validate the goodness and generalization of each model, a test data set is required for cross-validation (Haykin 1999, Stone 1974). A rule-of-thumb for ANN suggested that the test and validation data set be approximately 25 percent in size of the complete data set and be a balanced representation of the complete data set (NeuralWare 1995b). This computes to approximately a value of six. Haykin (1999) provided two equations (Equation 2 and Equation 3) for determining the size of a testing or validation

$$f = \frac{\sqrt{2P - 1} - 1}{2(P - 1)} \tag{2}$$

$$f = \frac{1}{\sqrt{2P}} \tag{3}$$

data set where f is the fractional value of the testing data set that determines its portion of the entire data set. The variable P is the number of independent parameters used as input. For the case when P is greater than or near the quantity of independent data sets, Equation 3 approximates Equation 2. For this case study the number of independent parameters initially used was 18 and the number of independent data samples was 23. Thus, from Equation 3, the value of f for the described case study was 0.1667 or 17 percent of the complete data set which computes approximately to a value of 4. However, as required for the MVLR Model, 20 samples were needed for developing the model. Therefore, only 3 samples remained for testing and validation instead of 4.

Statistically speaking, measuring the performance of the two candidate models with only 3 data sets for testing will not likely provide conclusive results because of the loss in degrees of freedom. Twenty or greater samples is a more statistically reasonable number when using such evaluators as the "paired t-test." To obtain 20 or greater samples, seven trials of distinct training and test data sets were uniformly randomly selected from the 23 available samples. Each trial had 20 samples for training and 3 samples for testing. See Figure 3 for an illustration of these trials. Since the training and testing data sets were randomly generated, there were some instances where a sample within a test data set was common to one of the other 6 test data sets. Discussion of how this effect was minimized will be discussed in the upcoming section *Compare Models and Choose Best.*

The statistical significance of each attribute of the team and the performance of the in-process elements are also an important consideration when building the sample data

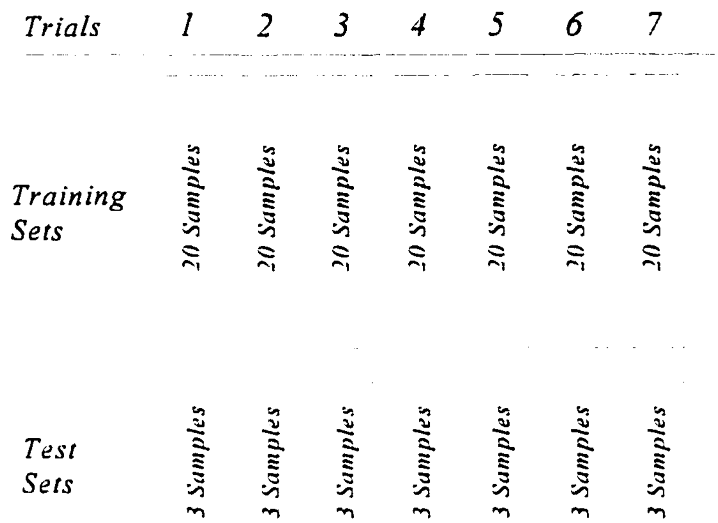


Figure 3 Diagram illustrating parceling of samples into training and test data sets for seven trials.

sets. However, tests for significance are different for the MVLR and the ANN, and will be discussed in subsequent sections that describe the procedures for both of the candidate methods.

Develop AHP for Combining Evaluation Factors

Evaluation of end-product performance will often require multiple factors in rating performance. In many cases, the relative weight or importance each of these factors has in regard to one another is unknown. The case study required that one evaluation score be given to determine the quality of the end product and thus, this thesis addressed the case for multi-factor performance analysis in order to generalize the process and provide a means to determine the relative weight of each factor. The widely used Analytical Hierarchy Process (AHP) was selected as a means to combine the multiple evaluation factors. This step in the methodology is not needed for a single evaluation factor. This section has an expanded discussion of AHP in order to fully describe the modifications made to make AHP less restrictive and awkward.

A total of 21 questions were used by the evaluators to score (determine quality) each squad's Mission Analysis end product. These 21 questions are found on the Mission Analysis Form found in Appendix D. The Mission Analysis Evaluation Form for PE 1 actually had a rating scale of 1 (no go), 2 (marginal go), and 3 (go) while PE 7 had a rating scale of 1 (no go), 3 (marginal go), and 5 (go) with intermediate scores of 2 and 4 permitted. This rating scale of 1 through 3 was already in use at the study site and was expanded to 1 through 5 to obtain a better representation of performance. To maintain the fidelity of the five point scale, the values of the evaluations using the three point scale where converted to the five point scale by setting a 1 to equal 1, a 2 to equal 3, and a 3 to equal 5. The original evaluation factors were not weighted and the relative importance between evaluation factors was unknown prior to evaluation. The evaluation factors were simply added linearly without regard to whether factors of low importance gave undue influence in the final evaluation. This procedure described herein used AHP to determine the relative weights for the evaluation factors.

AHP requires SMEs to make pairwise comparisons between factors. As previously discussed on page 17, when the number of factors are too great, a hierarchy formulation of the problem is recommended to make it more manageable. This procedure was accomplished and the 21 questions were organized into six groups after careful consultation with subject-matter experts. These groupings are found in Appendix F.

Each grouping has a hierarchical label {A-F} inserted before each group to help with cross-referencing the questions found in Appendix D that are also labeled as such. These labels used as identifiers of the groups are further used in explanation of the process and the data produced.

As a supplemental initiative to the IPB study, seven SMEs (MIOAC Instructors) where asked to conduct a pairwise comparison of the level of importance to Mission Analysis between these question groupings. Page 1 of Appendix F is the full description of the question groupings and a subject scale for pairwise comparison presented to the SMEs and page 2 of Appendix F is the data sheet used to record the comparisons. The order of the groupings where randomly (uniform) jogged (A,B,F,C,D,E) to help prevent the SMEs from mentally following the Mission Analysis process and possibly discrediting the importance of one factor over another due to its occurrence in the process. As called for by the AHP, the center diagonal of the data sheet was filled with one's and the lower left side of the data sheet was omitted from review (grayed blocks).

Like Table 1 found on page 17, 18, a subjective scale for pairwise comparisons was provided to SMEs to compare the evaluation factors. However, the scale used in this study is less restrictive. The scale on page 17, 18 assumes the order of importance is known between factors. Thus, the factors are arranged from left to right and top to bottom in an increasing order of importance. This is intended to yield values of the comparisons that are between the values of 1 to 9. Even if this assumption is relaxed, a SME likely would have to use reciprocal values (1/3, 1/5, 1/7, 1/9) to show that the base factor is less important than the other factor. Not only would this be awkward to the

SMEs, but it would be impossible to directly determine the average of the SMEs comparisons without first performing a scaling translation.

To alleviate these restrictions and awkwardness a modification of the AHP was developed. The application of the new scale is shown in Appendix F page 1. The new scale has values ranging from -9 to 9. The value -9 is the extreme case of where the base factor is overwhelmingly less important than the other factor and directly corresponds to a value of 1/9 in the AHP. The value of 9 is the extreme case of where the base factor is overwhelmingly more important than the other factor and directly corresponds to a value of 9 in the AHP. The negative numbers proved to be less awkward to the SMEs than the reciprocal values. Once the data were taken, the negative values were translated back to the corresponding reciprocal values.

Combine SME Responses and Apply AHP

Because the performance evaluation of the teams in the case study required that all scores to the questions of the evaluation sheet be combined, the procedure given in this thesis must also be combined. The relative rank of the pairwise comparisons performed by the seven SMEs as discussed in the previous section are shown in Appendix F, Tables F1, F3, F5, F7, F9, F11, F13. To determine the average response of the SMEs for each pairwise comparison, a transformation was required because the AHP scale is linearly unbalanced (zero is the midpoint between -3 and 3, not 1). Figure 4 illustrates this transformation. A value less than 1 indicates that when one factor is compared pairwise with a factor other than the one being compared (base factor) it is less important,

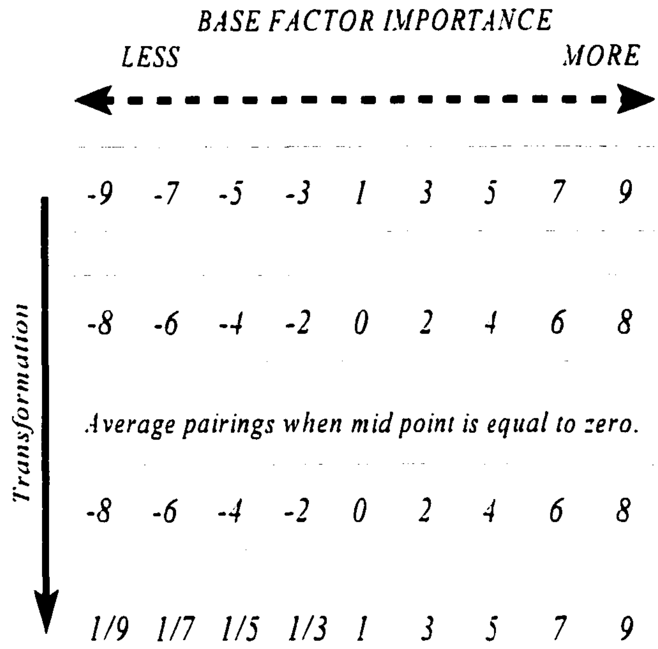


Figure 4 Diagram of transformation procedure used for modifications to AHP when combining SME pairings of evaluation factors.

when equal to 1 the same importance, and when greater than 1 it is of greater importance. Therefore, the original AHP scale was shifted by one towards zero and compressed from -8 to 8 shown in going from the first to second block in Figure 4. This translation is given by Equation 4 and the case study adjusted values are given in Appendix F, on Tables F2, F4, F6, F8, F10, F12, F14.

$$C_{ij} = \begin{cases} C_{ij} - 1 & C_{ij} \geq 0 \\ C_{ji} + 1 & C_{ij} < 0 \end{cases} \qquad (4)$$

Translate to Zero as Midpoint

C_{ij} is a value of a pairwise comparative relationship between base-factor *i* and other-factor *j* for *n* number of base-factors and n^2 number of comparisons. Referring to page 17, where C_{ij} is defined in more detail, Equation 5 provides the formula for computing the average of each pairwise comparison given by the SMEs. The variable "e" is the total number of SME values and the variable "k" is used to increment C_{ij}. The average values for the case study are found in Appendix F, Table F15.

$$\overline{C_{ij}} = \frac{\sum_{k=1}^{e} C^{k}_{ij}}{e} \qquad (5)$$

Average Pairwise Comparisons to aggregate

Before the AHP can be used with the average pairwise comparisons, the translation has to be reversed (third and fourth blocks in Figure 4). Once this is accomplished, Equation 6 is applied to complete the matrix (Appendix F, Table F16). Subsequently, Equations 7 and 8 were applied (Appendix F, Table F17) to produce weight factors for the question groupings.

$$C_{ij} = \begin{matrix} C_{ij} & i < j \\ \frac{1}{C_{ji}} & i > j \\ 1 & i = j \end{matrix} \qquad i,j = 1,\ldots,n \qquad (6)$$

Invert values for where $i > j$

Once the comparison values are computed, the relative weights are calculated either using an averaging or eigenvector method. Equations 7 through 9 are the equations needed to normalize and develop the AHP function for the averaging method. Variables are defined as: "η" is the normalized value of "C"; "w" is the factor weight computed from the normalized value; "x" is some parameter with a factor value; and "f" defines the resulting equation. For the case study the factor (Base Question) weights are given under the heading "AVG" in Appendix F, Table F17.

$$\eta_{ij} = \mathop{\Big|}\limits_{i=1}^{n} \mathop{\Big|}\limits_{j=1}^{n} \frac{C_{ij}}{\sum_{k=1}^{n} C_{ik}} \tag{7}$$

(7) Normalize pairwise comparisons

$$w_j = \mathop{\Big|}\limits_{j=1}^{n} \frac{\sum_{i=1}^{n} \eta_{ij}}{n} \tag{8}$$

(8) Compute factor weights

$$f = \sum_{i=1}^{n} x_i w_i \tag{9}$$

(9) Develop weighted average composite function

At this point Equation 9 could have been completed for the question groupings, but the question groupings were not what the evaluators responded to. Therefore, the weights had to be dis-aggregated across the 21 questions. This was accomplished by multiplying each question grouping weight factor by the number of questions that made up the question grouping and dividing by the total number of questions used in the evaluation. Equation 10 mathematically describes the resulting transformation of Equation 9.

$$f = \sum_{i=1}^{n} \sum_{j=1}^{m_i} x_{ij} \left(w_i \frac{m_i}{k} \right) \qquad (10)$$

Dis-Aggregate weighted average

Variables are defined as: "f" represents the derived evaluation score; "n" is the number of question groupings; "m" is the number of questions within a grouping; "x" is the individual question score; "w" is the question grouping weight;"k" is the total number of questions; "i" increments the question groupings; "j" increments the questions within a question grouping. For the case study the dis-aggregated weights per question are given under the column heading "Question Weight" in Appendix F, Table F17.

Build End-Product Data Set

As discussed in the *Build In-Process Data Set* section of this thesis, a third category of data was required, the evaluation scores. The actual IPB Study evaluation scores by question and the linear average evaluation score of these questions are given in Appendix F, Table F18 under the column heading "AVG." The dis-aggregated weights per question in Appendix F, Table F17 were multiplied to the corresponding question score value given in Table F18 to produce the values given in Table F19 as the final results of applying Equation 10 to the evaluation scores. The column under the heading

"SUM" contains the final derived composite score for each squad evaluation shown by row. The computation of these derived scores was the goal of the modified AHP introduced in this thesis and is what the team characteristics and in-process performance data were modeled against in the case study. The derived scores served as the dependent variable of each model.

Construct Models

This section discusses the approaches that were used to construct the two comparative models, one using MVLR and the other using ANN. The objective of this thesis was to develop a procedure for modeling team performance from acknowledged and well documented paradigms for modeling human performance, and not to explore all the possible methods for developing a MVLR or ANN model. The case study is real data, and thus some steps in the procedure was demonstrated with the case study data for illustrative purposes although tests for significance indicated that a procedural step was unnecessary for the data set. Such instances are so noted in the description of the procedure. Detailed discussion of the mathematical basis for these models and approaches are beyond the scope of this thesis; the interested reader to explore should review the provided references.

The mean squared error (MSE) was chosen as the common measure of performance to compare the test data set derived score sample values against the corresponding predicted score sample values of the test data set for each of the seven trials. MSE is a commonly used measure of performance and is good for exposing shortcomings in the accuracy of predictions. Equation 11 is the formula for MSE. The error or difference in the observed value Y_i and the predicted value Y_i^* is squared for all samples n before the mean is computed.

$$MSE = \frac{\sum_{i=1}^{n} (Y_i - Y_i^*)^2}{n} \tag{11}$$

(11) Mean Squared

All training and testing performed with each model paradigm was accomplished with identical training and testing data set pairs within a trial. As previously described on page 27, seven trials were needed. Therefore, all predictions and MSEs were directly compared between the MVLR and the ANN models.

Multi-Variate Linear Regression

A readily apparent assumption of MVLR is that the relationship between the input and output factors is linear for the solution space of interest. Moreover, most of the techniques available for Multi-Variate analysis assume the data were generated from a multi-variate normal distribution (Johnson and Wichern 1988). While these assumptions are limiting, MVLR has proven to be a robust method nonetheless. An off-the-shelf commercial software package called JMP was used to perform the MVLR (SAS Institute, Inc. 1995).

Iterate for each trial:

1. *Construct the MVLR models.*

a. *Screen factors that exhibit colinearity or duplication of information* in order to reduce modeling complexity, reduce future data collection by eliminating uninfluential factors, and increase sensitivity. Using the complete data set of 23 samples, a correlation matrix generated with JMP was used to determine colinearity between independent factors. Tables G1-G4 in Appendix G provide the goodness-of-fit value between the given factors for the case study (Johnson and Wichern 1988). Linear correlation is evident between two factors when the value of r is close to a value of 1 or -1. No factors showed a strong correlation between one another, thus no factors were eliminated due to duplication of information. An example of two factors that were suspected to be near duplicates and were not shown to be were "Age" and "Years in service." The paired correlation coefficient value was 0.118 for these two factors. A value greater than 0.85 or less than -0.85 indicates strong correlation (Anastusi 1976).

2. *Fit data with MVLR techniques*:

   a. *Import the complete data set into JMP*, exclude the test samples from the training samples from the regression algorithm, and conduct a MVLR model fit. Record the goodness-of-fit statistics for the training samples and the predicted values of the test samples.

   b. *Use backward stepwise regression as a screening method* to eliminate the uninfluential factors on the model prediction, exclude those factors from the regression algorithm, and conduct a MVLR model fit[1]. Again, record the goodness-of-fit statistics for the training samples and the predicted values of the test samples. The number of occurrences that each input factor was eliminated from the seven MVLR models are given in Table 3. This does not indicate that these factors were nonessential and should not be gathered for other case studies. It generally indicates that with the given data set, these particular factors had only a slight variation of value.

c. *Perform Regression and compute mean squared error (MSE) for each test data set.* The fit of the regression model, as measured by the adjusted r^2 of the training data for each trial, is provided in Table 4 for the before and after factor elimination cases. The adjusted r^2 for the training set increased favorably after factor screening was performed. However, the predictions with the test set showed no improvement as shown in Table 5. The MSE value of all test sample predictions before screening was 2.101 and the MSE value of all test sample predictions after screening was 2.325. A paired t test was performed on all test sample predictions before and after screening to determine whether there is a significant difference in the predictions. No significant difference was found with α=0.05. Had there been a significant difference, and the MSE of the screened test sample predictions was larger

[1]Stepwise regression is an approach for selecting a subset of factors for a regression model. In an iterative fashion the procedure removes input factors from the regression equation that do not show a greater than 0.10 probably of receiving a greater F statistic. Thus, by design, the procedure will increase the adjusted r^2 and reduce the MSE until no other factors meet the screening criteria (SAS Institute, Inc. 1995).

than the MSE of the non-screened test sample predictions, then the regression model based on the non-screened data would have be used regardless of the practical aspects and economics of reducing factor data collection and analysis.

Table 3. Elimination of the MVLR uninfluential input factors for seven trials

Input Factor Description	Occurrences
Age of Members	1
Gender of Members	0
Rank of Members	2
Years in Service of Members	2
OTC of Members	2
Computer Skill Level of Members	3
IPB Experience of Members	5
NTC of Members	2
Practical Exercise Performed	6
Computer Aided Squad or Manual	0
Number of Members developing MCOO	3
MCOO Development Duration	3
Number of Members working TPL what-ifs	1
Number of TPL what-ifs developed	0
Time Phase Line (TPL) Duration	1
Number of Members developing SitMap	3
Situational Map (SitMap) Duration	4
Evaluator scoring squad	3

Table 4. Comparison of adjusted r^2 values before and after screening for each trial

Trial	Adjusted r^2 Before Screening*	Adjusted r^2 After Screening
1	0.946	0.970
2	0.790	0.937
3	0.662	0.817
4	0.575	0.868
5	0.116	0.756
6	0.709	0.931
7	0.980	0.991

* As an adjusted r^2 value approaches 1.0, so does the goodness of the model fit.

Table 5. Comparison of predicted MVLR values before and after screening with derived scores by sample

Trial	Sample	Derived Score	Before Screening	After Screening
1	1	3.520	3.581	3.672
1	2	2.830	5.440	5.380
1	3	4.290	3.162	3.144
2	1	3.290	5.999	6.289
2	2	3.400	1.359	1.176
2	3	2.970	0.545	0.093
3	1	2.780	3.518	3.596
3	2	4.290	4.682	4.382
3	3	2.830	4.505	4.954
4	1	4.360	3.869	3.814
4	2	2.830	4.390	4.554
4	3	4.140	4.450	4.410
5	1	3.090	1.282	2.511

Table 5. (Continued)

5	2	3.860	3.039	3.280
5	3	3.290	3.956	4.241
6	1	3.950	4.464	3.950
6	2	3.400	2.268	2.555
6	3	4.370	3.083	3.046
7	1	4.360	3.782	3.949
7	2	4.070	2.214	2.039
7	3	4.370	5.376	5.754
		MSE	2.101	2.325

Further discussion of the results is given in the section *entitled Compare Models and Select Best*. Serving as an example of the mathematical form of a MVLR, the model before any input factors were eliminated for trial 1 is given in Appendix H.

Artificial Neural Network

As discussed on page 13, statistical approaches require an analyst to determine how output data mathematically relates to input data. This is also true for traditional mathematical programming, but ANN does not require knowing whether the relationship is linear or what form of non-linearity. However, there is a need to follow a procedure that will construct the best possible model with the provided data. Simply "throwing" data at an ANN will not provide a good model. The procedure developed for modeling with an ANN is as follows.

1. *Construct the Backpropagation (Backprop) ANN models.*

   a. *Select ANN software package or develop custom ANN software program.* An off-the-shelf commercial software package called NeuralWare Professional II/Plus was used (NeuralWare 1995a).

   b. *Determine number of training cycles.* Three additional trials of training and test/validation data sets were created in order to determine the number of training cycles needed to avoid overfitting or memorization. There was no rationale for choosing three more trials other than the fact that some estimation was needed and by just using one trial an analyst runs the risk of using a value for training cycles that only makes sense for that one trial. In order for the ANN model being developed to make consistently good predictions, generalization must be achieved. Thus, overfitting must be minimized. Overfitting is evident when the Root Mean Square (RMS is square root of the MSE) computed from the original training set output layer (derived score) and the predicted training set output layer (predicted score) is small in comparison to the RMS computed from the test set output layer (derived score) and the predicted test set output layer (predicted score) (Amari et al. 1996). Generalization will occur when these two measures are nearly equal. The Early Stopping Method of Training was employed to determine a suitable number of training cycles (Amari et al. 1996)[1]. As seen in Figure 5, the early-stopping point for the three test/validation sets would be

---

[1]Simply stated, the Early Stopping Method calls for checking the MSE or RMS of a test/validation samples at a pre-determined training cycle (i.e. 1,000) for a minimum value as the ANN is being trained with the training samples.

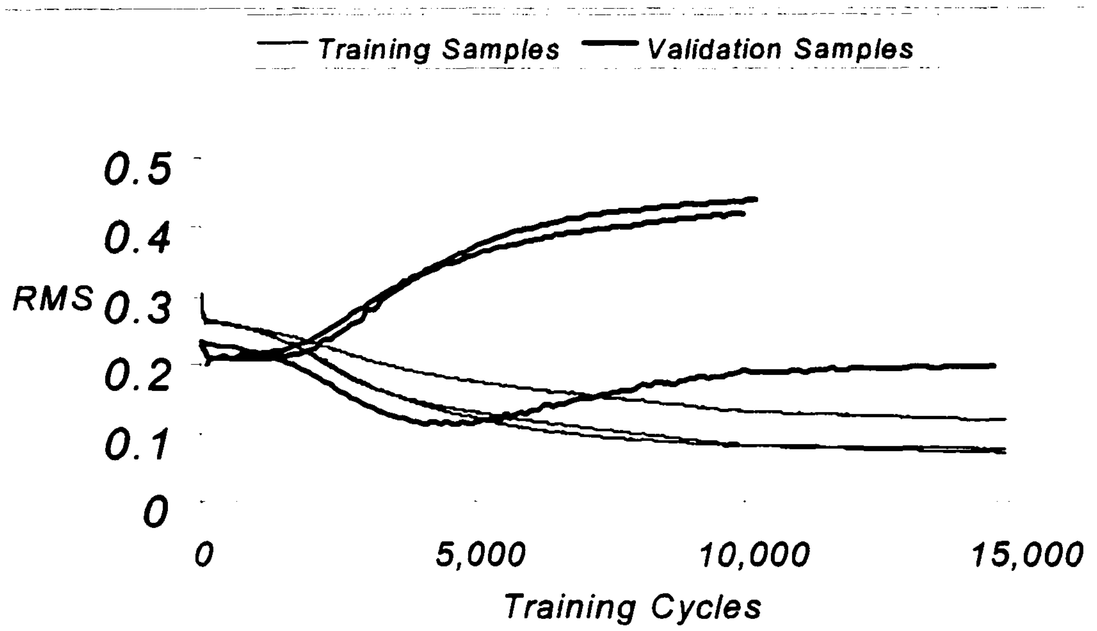


Figure 5 Plot of RMS versus training cycles for the training and test/validation samples chosen to determine early-stopping point.

somewhere between 2,000 and 5,000 training cycles. Initially 2,000 training cycles were used as the early-stopping point. However, for the given network structure and number of data samples available for this case study, only using 2,000 training cycles did not sufficiently allow the ANN training procedure to properly learn and adjust the influence of each input factor on the output factor from the initial state conditions of the ANN. Therefore, the other end of the range, (5,000 training cycles) was used and did permit the ANN to apply emphasis to the appropriate input factors and pruning would be possible. Pruning is synonymous with screening as described in the previous section.

c. *Import the training and test data set into NeuralWare Professional II/Plus.*

d. *Select Backprop candidate learning rules, transfer functions, number of hidden layers, number of hidden layer elements, and convergence criteria.* The defaults for the NeuralWare Professional II/Plus software used the Normalized-cumulative delta-rule as the learning rule, the sigmoid function as the transfer function, and the minimization/stabilization of the RMS as the convergence criterion. NeuralWare (1995b) suggested that the number of hidden layers for behavioral models be determined by dividing the number of training cases by the product of the number five and the sum of the number of input data elements and the number of output data elements. If the value is less than one, then set it equal to one. Therefore, one hidden layer was used for the case study. The initial number of hidden layer elements was chosen to be roughly 50 percent of the number of in-process. The references Haykin (1999) and Kosko (1992) confirmed that the defaults as described by NeuralWare are suitable. Figure 6 is an image of the menu and values used for establishing and training the ANN.

Iterate for each trial:

2. *Train Backprop ANN Models to preselected convergence criteria.*

   a. Train the Backprop ANN models to preselected convergence criteria or early-stopping point and analyze MSE and RMS for goodness. To improve performance as measured by the MSE, re-train the ANN as needed by eliminating or pruning input data factors that do not influence the objective function by more than 5 percent by weight as suggested by NeuralWare. An ANN of minimum size

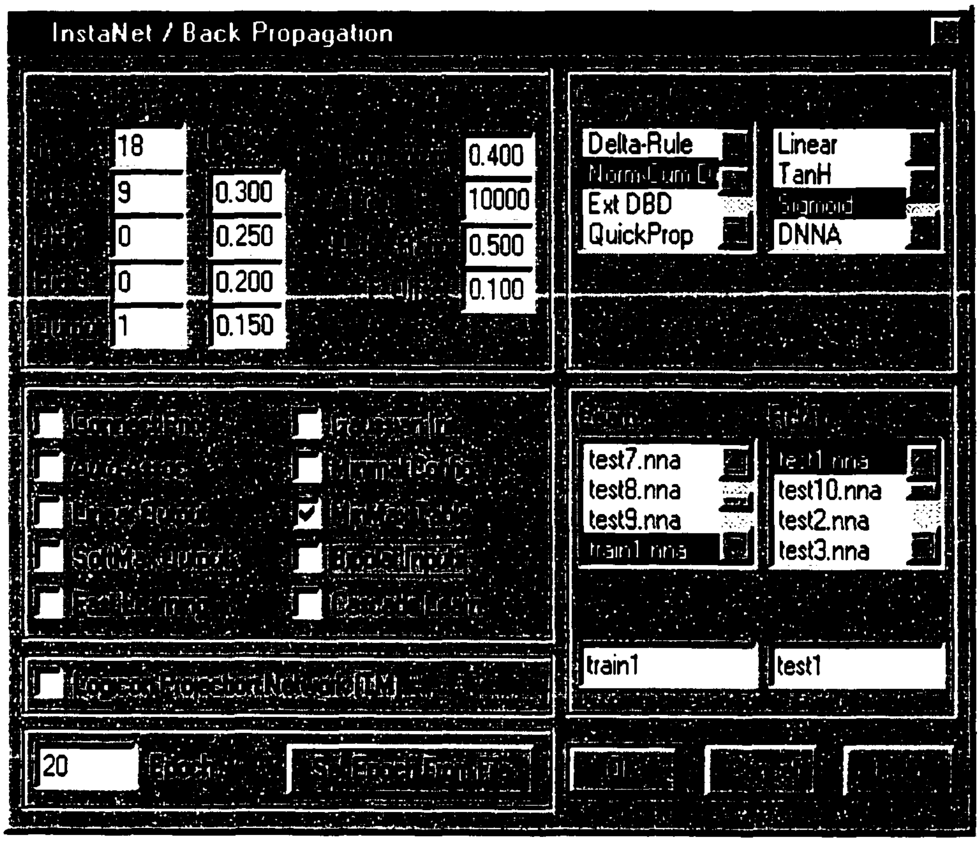


Figure 6 Menu of Backpropagation ANN setup for first sample set (Neural Ware II/Plus)

is less likely to learn noise in the training data and may generalize better (Haykin 1999). As mentioned in the previous section, there are practical aspects and economical advantages of reducing the number of input factors, and pruning is recommended. For the purpose of developing a modeling procedure and not over complicating this thesis with the variety of pruning methods available, only one iteration of the simplistic pruning technique as described above was performed. Consult Haykin (1999) for additional methods.

b. *Compute MSE for each Backprop ANN test set.*

The occurrence of pruned input data factors throughout the seven trials for the total system process (Mission Analysis) are given in Table 6. These factors did not achieve significant influence in ANN model prediction. This does not indicate that these factors were nonessential and should not be measured for other studies or analysis purposes. It generally indicates that with the given data set, these particular factors had only slight variation of value.

As shown in Table 7, the MSE value of all predictions before pruning was 0.433 and the MSE value of all predictions after pruning was 0.287. A paired *t* test was performed on all test sample predictions before and after pruning to determine whether there was a significant difference in the predictions. A *t* value of 0.989, $\alpha=0.05$ level of significance for 20 degrees of freedom indicated that there was no significant difference. Although it was apparent from the paired *t* test results that there was no significant difference or improvement after pruning, the ANN model developed after pruning is recommended due to the practical aspects of pruning and the reduction of the MSE.

Table 6. Elimination of the ANN uninfluential input factors for seven trials

Input Factor Description	Occurrences
Age of Members	6
Gender of Members	2
Rank of Members	0
Years in Service of Members	3
OTC of Members	2
Computer Skill Level of Members	5

Table 6. (Continued)

IPB Experience of Members	2
NTC of Members	3
Practical Exercise Performed	3
Computer Aided Squad or Manual	0
Number of Members developing MCOO	0
MCOO Development Duration	5
Number of Members working TPL what-ifs	1
Number of TPL what-ifs developed	5
Time Phase Line (TPL) Duration	2
Number of Members developing SitMap	6
Situational Map (SitMap) Duration	3
Evaluator scoring squad	1

Table 7. Comparison of predicted ANN values before and after pruning with derived scores

Trial	Sample	Derived Score	Before Pruning	After Pruning
1	1	3.520	3.620	3.792
1	2	2.830	3.279	3.032
1	3	4.290	3.669	3.851
2	1	3.290	4.419	4.263
2	2	3.400	3.260	3.312
2	3	2.970	3.154	3.154
3	1	2.780	3.083	3.383
3	2	4.290	4.031	3.535
3	3	2.830	4.088	3.457

Table 7. (Continued)

4	1	4.360	4.095	4.100
4	2	2.830	2.988	3.046
4	3	4.140	4.152	4.121
5	1	3.090	3.555	3.619
5	2	3.860	3.356	3.595
5	3	3.290	4.479	3.618
6	1	3.950	3.588	3.478
6	2	3.400	3.183	3.200
6	3	4.370	3.515	3.540
7	1	4.360	4.039	3.997
7	2	4.070	2.520	2.841
7	3	4.370	4.083	3.874
		MSE	0.433	0.287

Further results will be discussed in the section entitled *Compare Models and Select Best*. Serving as an example of the mathematical form of an ANN, the model (written as a *C* program function) before any input factors were eliminated for trial 1 is given in Appendix H.

Compare Models and Select Best

The development of the procedure and discussion of input and output data thus far has been to support the comparative analysis found in this section. To compare the two models, (1) hypotheses will be stated and proved or disproved regarding the predictive ability of modeling a knowledge team process, (2) hypotheses will be stated and proved or disproved regarding whether there was any significant difference in the two models, (3) the sensitivity of the two models will be explored, and (4) the findings discussed.

Hypotheses Development

The first step in model comparison was to show whether one model was significantly better than or no different than the other for predicting the derived scores. Thus, the following hypotheses are postulated and subsequently tested.

Null Hypothesis 1:

$H1_0$: With regard to the test samples by trial, the derived scores and the MVLR predicted scores are equal.

Alternative Hypothesis 1:

$H1_1$: With regard to the test samples by trial. the derived scores and the MVLR predicted scores are not equal.

Null Hypothesis 2:

$H2_0$: With regard to the test samples by trial, the derived scores and the ANN predicted scores are equal.

Alternative Hypothesis 2:

$H2_1$: With regard to the test samples by trial, the derived scores and the ANN predicted scores are not equal.

The second step in model comparison was to show whether there was a statistical difference between the two model predictions. Thus, the following hypotheses are postulated and subsequently tested.Null Hypothesis 3:

$H3_0$: With regard to the test samples by trial, the MVLR and the ANN predicted scores are equal.

Alternative Hypothesis 3:

$H3_1$: With regard to the test samples by trial, the MVLR and the ANN predicted scores are not equal.

As stated in the procedure for generation of the seven trials of sample data on page 27, there is a slight bias introduced with using repetitive data. Additionally, the population variance is unknown, and by a Shapiro-Wilk Test, the distribution of the 21 derived score samples is not normally distributed (SAS Institute, Inc. 1995). To get around this limitation and use such tests as a paired *t* test, the samples have been averaged within trials and are subsequently normally distributed according to the Central Limit Theorem. Thus, comparisons with the *t* statistic will be more robust (Miller and Freund 1977). Had the 21 samples of derived scores been normally distributed, it would be recommended to use the values of the 21 samples rather than the 7 trials.

Therefore, the paired *t* tests were performed on the mean of the samples within each trial. The level of significance selected for this analysis is α=0.05 with 6 degrees of freedom (Miller and Freund 1977). Table 8 provides the statistics of averaging the seven trials for the derived scores, the MVLR predicted scores, and the ANN predicted scores.

When testing the null hypothesis $H1_0$, the paired *t* test yielded a *t* value of 0.094, α=0.05 with 6 degrees of freedom for having larger differences between the derived scores and the MVLR predicted scores. Because the *t* value did not exceed 1.943, the null hypothesis $H1_0$ cannot be rejected. When testing the null hypothesis $H2_0$, the paired *t* test yielded a t value of 0.484, α=0.05 with 6 degrees of freedom for having larger differences between the derived scores and the ANN predicted scores. Because the *t*

Table 8. Statistics of seven trials for the derived scores, the MVLR predicted, and the ANN predicted scores

	Trial	Derived Score	MVLR Predicted	ANN Predicted	Derived Score minus MVLR	Derived Score minus ANN	MVLR minus ANN
	1	3.547	4.065	3.558	-0.519	-0.012	0.507
	2	3.220	2.519	3.577	0.701	-0.357	-1.058
	3	3.300	4.311	3.458	-1.011	-0.158	0.853
	4	3.777	4.259	3.756	-0.482	0.021	0.503
	5	3.413	3.344	3.611	0.069	-0.197	-0.267
	6	3.907	3.184	3.406	0.723	0.501	-0.222
	7	4.267	3.914	3.571	0.353	0.696	0.343
Mean		3.633	3.657	3.562	-0.024	0.071	0.094
Variance		0.139	0.440	0.013	0.444	0.149	0.422
STD		0.373	0.663	0.112	0.666	0.386	0.650
MSE					0.381	0.132	

value did not exceed 1.943, the null hypothesis $H2_0$, cannot be rejected. When testing the null hypothesis $H3_0$, the paired *t* test yielded a *t* value of 0.383, $\alpha=0.05$ with 6 degrees of freedom for having larger differences between the MVLR and the ANN predicted scores. Because the *t* value did not exceed 1.943, the null hypothesis $H3_0$, cannot be rejected.

From the analysis using the paired *t* test, both models' prediction was not significantly different from the derived score, nor was there a statistical difference between the predicted scores of the two models even though the MVLR model yielded a greater MSE than the ANN model.

Sensitivity Analysis

Regression models such as MVLR or even the ANN are built from an existing data set and thus model well within or just outside of the relevant range of the data. However, in practice the model might be required to make broader extrapolations than the relevant range defined by the original data set it was created from, although partitioning the data into training and test sets helped minimize the risk of venturing outside the relevant range. Thus, model predictive sensitivity from sample data not yet seen, was explored.

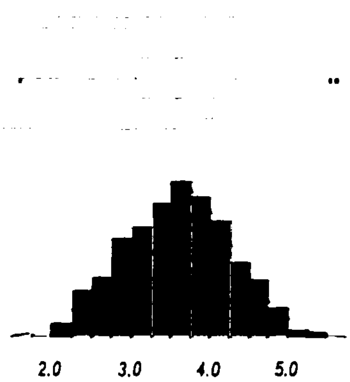
To test the sensitivity of the two models, two new data sets were produced from the complete data set for a sensitivity analysis. The first data set contained one thousand samples with the input factors not screened for less than five of the seven trials in the MVLR development (see Table 3). The second data set contained one thousand data records with the input factors not screened (pruned) for less than five of the seven trials in the ANN development (see Table 6). In these two data sets, the values of the input factors were randomly generated from either a normal distribution with a mean and standard deviation of the complete data set if the distribution was found to be normal from a Shapiro-Wilks W Test. If the distribution was not found to be normally distributed, a uniform distribution with the minimum and maximum values from the complete data set appropriate for the given factor was used. Normally distributed input factors are noted in Table 9. Factors that require discrete values were rounded to the nearest integer. Infrequently, values of factors were generated to be near zero, but negative in sign. These values were rounded to zero.

Table 9. Complete list of input factors with or without normally distributed data.

Factor	Normal* (✔)
Age	✔
Gender	
Rank	
Years of Service	
OTC Graduate	
Computer Skill Level	✔
IPB Experience	
NTC Rotations	
Practical Exercise	
CAS or MMS	
Members of MCOO	
MCOO Time	✔
Members on TPL	
Whatifs on TPL	
TPL Time	
Members on SITMAP	
SITMAP Time	
Evaluator	✔
Derived Score	✔
*Shapiro-Wilk W Test used to test normality, α=0.05	

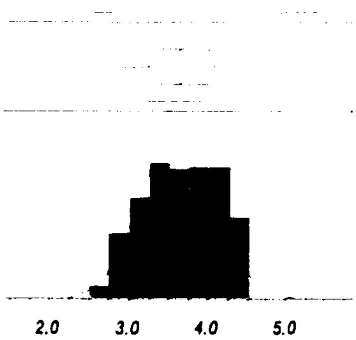
Figure 7 provides a summary of the sensitivity results for the 1,000 sample MVLR data set predictions. Figure 8 provides a summary of the sensitivity results of the 1,000 sample ANN data set predictions. The results showed that the mean predicted response for both models were within 3% of the original 23 sample mean given in Appendix F,

Table F20. Neither model predicted below the minimum possible score of 1.0, but the MVLR model predicted above the maximum possible score of 5.0 roughly 3% of the time and thus went outside of the intended relevant range. The ANN model did not predict above this maximum value.



Quantiles			Moments	
maximum	100.0%	5.6239	Mean	3.579
	99.5%	5.4206	Std Dev	0.704
	97.5%	4.9257	Std Err Mean	0.022
	90.0%	4.5096	upper 95% Mean	3.623
quartile	75.0%	4.0642	lower 95% Mean	3.536
median	50.0%	3.5871	N	1000.000
quartile	25.0%	3.0824	Sum Wgts	1000.000
	10.0%	2.6218		
	2.5%	2.2474		
	0.5%	1.7483		
minimum	0.0%	1.5907		

Figure 7. MVLR Predictions with 1,000 randomly generated samples



Quantiles			Moments	
maximum	100.0%	4.5392	Mean	3.640
	99.5%	4.4929	Std Dev	0.460
	97.5%	4.3925	Std Err Mean	0.015
	90.0%	4.2545	upper 95% Mean	3.669
quartile	75.0%	4.0256	lower 95% Mean	3.612
median	50.0%	3.6530	N	1000.000
quartile	25.0%	3.2756	Sum Wgts	1000.000
	10.0%	2.9938		
	2.5%	2.7944		
	0.5%	2.6286		
minimum	0.0%	2.5775		

Figure 8. ANN Predictions with 1,000 randomly generated samples

As a check or augmentation to the test of hypothesis $H3_0$ (the MVLR and the ANN predicted scores are equal) a two sample test (*z* statistic) was conducted on the statistics of the 1,000 sample MVLR data set predictions and the 1,000 sample ANN data set predictions. The two sample test (z statistic) as used here, specifically tests whether the sample means can be stated to be from the same population for 30 or greater samples (Miller and Freund 1977). The z statistic value obtained was 2.294. Since the value is greater than the critical value 1.96 for an α=0.05 level of significance, it cannot be stated that these two samples are of the same population.

Findings

The findings of the comparison between the MVLR and ANN models are as follows:

- The MVLR predicted scores and the derived scores were not statistically different for an α=0.05 level of significance.
- The ANN predicted scores and the derived scores were not statistically different for an α=0.05 level of significance.
- The MVLR predicted scores and the ANN predicted scores were not statistically different for an α=0.05 level of significance.
- Neither model exhibited an unreasonable sensitivity to exceed the boundaries of the relevant range when subjected to 1,000 samples representative of the full solution space.
- The MVLR predicted scores of a 1,000 sample simulated data set and the ANN predicted scores of a 1,000 sample simulated data set were significantly different for an α=0.05 level of significance.

For the case study chosen, there was no statistical evidence that one model paradigm was better than the other. However, because the ANN model produced a smaller MSE by less than half of the MVLR model, was less dependent on knowing the mathematical relationships between the factors, was able to simulate performance within the bounds of the relevant range, and was recommended over MVLR in the literature review, the ANN model was preferred (Lykins and Chance 1992, Eksioglu et al 1996).

Refine Model and Perform Analysis

Before the utility of developing a knowledge team process model with ANN is demonstrated, the model should be refined to remove as much experimental bias as possible. This section demonstrates the value of the developed model as a tool for process analysis.

The screening and pruning did not eliminate the two input factors, Practical Exercise and Evaluator, that captured the experimental limitations. These two factors showed too much influence in the derived scoring results to eliminate during the pruning process. The IPB Study showed that although the scores for teams in PE 1 were higher than for those in PE 7, the differences were not statistically significant (Deliman et al. 1997). Further investigation revealed the teams in PE 7 should have higher scores than those in PE 1 because the PE 7 teams had already performed the process before in PE 1. Thus, the effect of this factor does not need to be normalized across the complete data set.

The complete data sample was too small to explore the influence of the evaluator scoring the team, thus the data set for the sensitivity analysis was used to explore the bias and compute the normalization multipliers for derived scores of the complete data set. Figure 9 provides a graph of the predicted scores by evaluator of the sensitivity analysis data set. The graph clearly shows a difference in the predicted scores by the evaluator.

To test the statistical significance of this bias, a two sample test (z statistic) was performed with the two samples having the greatest difference between the means, Evaluator 1 and Evaluator 7. The z statistic value obtained was 3.33. Since the value is greater than the critical value 1.96 for an $\alpha=0.05$ level of significance, it cannot be stated

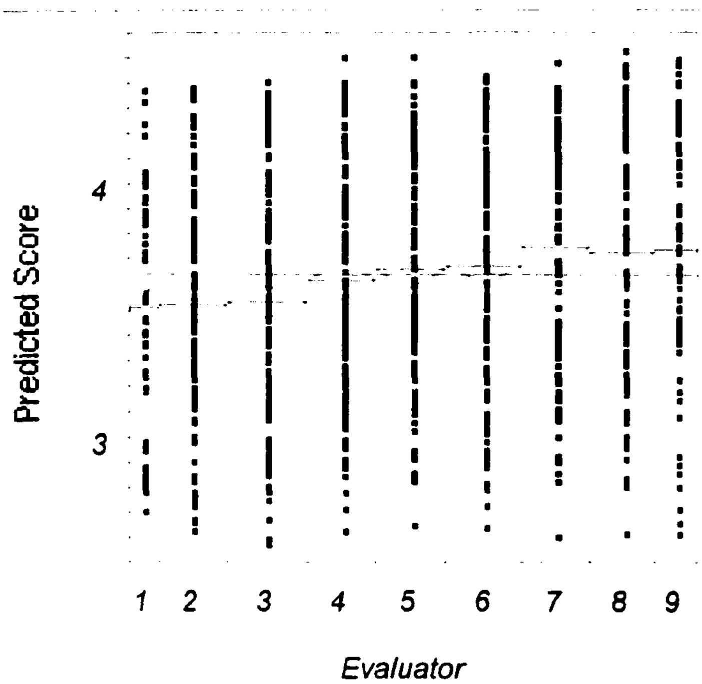


Figure 9 Distribution of predicted scores and means (diamonds) by evaluators for 1,000 simulations.

that these two samples are of the same population and thus the bias is significant. Had these two samples not been shown to be different, pairwise comparisons would need to be performed between the other evaluator data sets to determine any significant bias.

Table 10 provides a table of the mean predicted scores by evaluator and the normalization factor required to remove the bias. Table G5 in Appendix G provides the normalized derived scores along with the refined training set for the final ANN model build.

Table 10. Computation of Normalizing Factors

Evaluator	Mean	Normalization Factor*
1	3.512	1.037
2	3.518	1.035
3	3.530	1.031
4	3.623	1.005
5	3.666	0.993
6	3.680	0.989
7	3.755	0.969
8	3.730	0.976
9	3.745	0.972
All	3.640	
*Computed from division of Evaluator mean by the mean of all evaluator means.		

The emphasis of this thesis research was to develop a procedure for modeling team performance and not the analysis of the performance per se. The literature review cited references regarding process improvement and value added analysis. However, to demonstrate the utility of the developed procedure for predicting team performance, a cursory analysis of the high and low performing teams was performed and is discussed further in this section.

After the derived scores were normalized, the ANN was retrained on the complete data set given in Table G5 Appendix G excluding the evaluator input factor for 5,000 training cycles. The sensitivity analysis data set that was previously discussed excluding the evaluator input factor was used to simulate 1,000 teams with the refined ANN model. Table 11 provides the means of the factors and predicted scores for the high or top 10 percent performing teams and low or bottom 10 percent performing teams, categorized by the PE for that simulation. The data provided in Table 11 results in the following observations:

- The team characteristic factors Gender, Rank, Years of Service, Officer Transition Course (OTC) Training, IPB experience, and number of National Training Center (NTC) Rotations (rounds to 2), showed very little variation between the high performing teams and low performing teams regardless of the PE, as defined by the means and variances of the sample data. Therefore, team mix considerations cannot be explored.

- The team performance factors on whether a DSS was used or not (CAS or MMS), the number of team members developing the MCOO, and the number of team members exploring what-ifs for TPLs showed a significant difference in predicted score, regardless of the PE. Furthermore, the results showed that 87 percent ($1 - 0.13 \times 100$ percent) of the top performing teams used the DSS in the IPB Study regardless of the PE, and 93 percent to 91 percent of the bottom performing teams did not use the DSS in their respective PEs. The top performing teams used three to four members on the MCOO development and six to seven on the what-ifs for TPLs versus the bottom performing teams that used six members on the MCOO development and two to three people on the What-ifs for TPLs.

- The team performance factors Time Phase Line (TPL) development time and Situational Map (SITMAP) development time also showed very little variation between the top and bottom performers regardless of the PE, as also defined by the means and variances of the sample data. Therefore, analyzing the variation of these factors would probably not yield much.

Table 11. Mean Factor Values and Predicted Scores for the Top 10 Percent and Bottom 10 Percent Performing Teams

		Team Characteristics							Team Performance				
	PE	GENDER	RANK	YRS	OTC Train	IPB Exper	NTC Rota	CAS MMS	#MEM MCOO	#MEM Whatifs	TPL Time	SIT MAP Time	Predic Score
TOP 10%	7	1.79	2.72	1.27	1.56	2.78	1.87	0.13	3.81	6.85	0.59	1.99	4.34
TOP 10%	1	1.82	2.73	1.27	1.58	2.75	1.88	0.13	3.55	6.42	0.56	2.10	4.22
BOT 10%	7	1.78	2.72	1.28	1.55	2.79	2.20	0.93	6.20	2.25	0.61	1.89	3.17
BOT 10%	1	1.79	2.73	1.25	1.58	2.80	2.28	0.91	5.91	2.74	0.58	2.04	2.99

Based on the analysis, the DSS did add value to the process. Moreover, three to four members should develop the MCOO, and six to seven members should explore what-ifs for the TPLs.

CHAPTER IV

CONCLUSION

Summary of Findings

As previously stated in the *Problem Statement*, what is needed is a procedure to relate the in-process elements of team performance to the end-product while taking into account the variability of team interpersonal actions. More specifically stated in the section *Conclusions of Review*, a procedure is needed to model the performance of a knowledge team process while taking in account the variability of SME judgement for value added analysis and process improvement.

This thesis showed there was no statistical difference in the derived scores and the MVLR predicted scores, nor was there a statistical difference in the derived scores and the ANN predicted scores. While a paired *t* test between the MVLR predictions and the ANN predictions for the 7 trials were not significantly different, a two sample test of 1,000 simulations of each model revealed a significant statistical difference. Therefore, because of the results of the two sample test, smaller MSE for the 7 trials, ease of model development, and recommendations from similar studies by Lykins and Chance (1992) and Eksioglu et al (1996) the ANN was the preferred model to complete the case study.

This thesis has systematically provided a procedure to model a team process by combining a modified AHP with a modeling paradigm such as MVLR or ANN. A proofof-principle has been demonstrated by the case problem results and should have utility for value added analysis and process improvement. Moreover, the procedure is general enough to be applied to other team processes.

Summary of Contributions

In this age of information and eagerness to improve efficiency, lower cost, and to improve quality, tools and procedures are needed to analyze processes. Improving knowledge team effectiveness is a difficult, but an important endeavor. The procedure resulting from this thesis provides a method to create a model of a team process for developing a recommendation to management for improving team makeup and performance. With such a model, an analyst can perform sensitivity analysis by varying the team characteristics, members performing a task, and time performing a task within the process. This type of analysis is useful to determine ways to improve the process and eliminate process elements that do no add value.

Lessons Learned and Future Research

The procedure developed in this thesis provides a good approach for modeling a team process. Insufficiently structured processes or a lack of field data could yield problems in modeling the process. However, an insufficiently structured process does not necessarily indicate that it is a bad process, only that it will be difficult to model and make improvements. Data inadequacy may be caused by any of the factors or combination of the following factors: (1) insufficient quantity of samples replicating the process,

(2) insufficient quantity of data factors describing the process, (3) lack of pertinence of data factors describing the process, or (4) poor data collection methods.

Further research should be explored to compare MVLR and ANN model paradigms for modeling the team process to determine which is better overall or for particular processes. Different processes with larger sample sizes would be advisable. Expected enhancements to the procedure may develop by exploring: (1) screening methods for MVLR, (2) pruning methods for ANN, (3) different ANN architectures, and (4) refinement of the Early Stopping Method.

Analysts using the procedure described herein are advised to validate their model with field experiments and test proposed design improvements prior to making recommendations to management. Furthermore, as with the case study, the focus of a DSS development project should be maximizing the value added contributions provided by the knowledge worker utilizing the DSS (Robinson 1991).

REFERENCES

Amari, S., Murata, N., Muller, K. R., Finke, M., and Yang, H. (1996). "Statistical theory of overtraining— Is cross validation asymptotically effective?" Advances in Neural Information Processing Systems, Vol. 8, pp. 176-182, Cambridge, MA: MIT Press.

Anastusi, Ann (1976). "Reliability," Psychological Testing, 4th ed., MacMillan, New York, New York, p 109.

Baker, D. P. (1992). "Principles for Measuring Teamwork Skills," *Human Factors* 34(4), 469-475.

Beaudet, D. B., and Williges, R. C. (1988). "The Role of Screening Studies in Sequential Research Designs," *Proceedings of the Human Factors Society--32nd Annual Meeting.*

Bedworth, David D., and Bailey, James E. (1982). *Integrated Production Control Systems.* John Wiley & Sons, Inc., New York, New York.

Bittner, A. C. (1992). "Robust Testing and Evaluation of Systems: Framework, Approaches, and Illustrative Tools," *Human Factors* 34(4), 477-484.

Boza, L. B., Hsu, J. C., and Wyndrum, R. W., Jr. (1990). "The Role of Process Engineering in Product and Service Realization," *IEE/ISHM 1990 IEMT Symposium.* Italy

Bradyhouse, R. G. (1982). "Value Added Analysis - Major Key to Productivity Gains," *Manager Value Engineering Systems.* Black and Decker Manufacturing Company, Towson, Maryland.

Buede, D. M., and Maxwell, D. T. (1995). "Rank Disagreement: a Comparison of Multi-criteria Methodologies," *Journal of Multi-Criteria Decision Analysis* 4, 1-21.

Campion, M. A. (1991). "Meaning and Measurement of Turnover: Comparison of Alternative Measures and Recommendations for Research," *Journal of Applied Psychology* 76(2), 199-212.

Charlton, S. G. (1992). "Establishing Human Factors Criteria for Space Control Systems," *Human Factors* 34(4), 485-501.

Deliman, Niki C., Baylot, E. Alex, Williamson, Jeffrey L., and Bunch, Laura S. (1997). *Intelligence Preparation of the Battlefield (IPB) Process Value Added via Automated Geo-Referenced Mobility Algorithms and Products*, Final Draft, USAEWES, Vicksburg, Mississippi.

Eksioglu, Mahmut; Fernandez, Jeffrey E.; Twomey, Janet M. (1996). "Predicting peak pinch strength: Artificial neural networks vs. regression," *International Journal of Industrial Ergonomics*, volume 18, pp. 431-441.

Escover, J. L. (1994). "The Value Measure," *Management Decision* 32(1), 12-14.

Haykin, Simon. (1999). *Neural Networks, A Comprehensive Foundation, Second Edition*, Prentice Hall, Upper Saddle River, New Jersey.

Headquarters, Department of the Army. (1994). *Intelligence Preparation of the Battlefield, Field Manual 34-130*, Washington D.C.

Johnson, Richard A., and Wichern, Dean W. (1988). *Applied Multivariate Statistical Analysis, Second Edition*, Prentice Hall, Englewood Cliffs, New Jersey.

Kantowitz, B. H. (1992). "Selecting Measures for Human Factors Research," *Human Factors* 34(4), 387-398.

Klimasauskas, C. C. (1992). "Neural Networks: an Engineering Perspective," *IEEE Communications Magazine*, September 1992, 50-53.

Kosko, Bart. (1992). *Neural Networks and Fuzzy Systems*, Prentice-Hall, Englewood Cliffs, New Jersey.

Lykins, Sue, and Chance, David. (1992). "Comparing Artificial Neural Networks and Multiple Regression for Predictive Application." *Proceedings of the Eighth Annual Conference on Applied Mathematics*, pp. 155-169, Edmond, Oklahoma.

Mason, D. E. (1995). *Identifying Measures of Effectiveness for Marine Corps C^4I Systems*, M.S. thesis, Naval Postgraduate School, Monterey, California.

Meister, D. (1985). *Behavioral Analysis and Measurement Methods,* John Wiley and Sons, New York, New York.

Meister, D. (1986). *Human Factors Testing and Evaluation*, Elsevier Science Publishers, New York, New York.

Mikaloff, Robert S. (1996). "Mobility Analysis for the Digitized Brigade," *Armor*, September-October, 41-44.

Miller, Irwin, and Freund, John E. (1977). *Probability and Statistics for Engineers*, 2nd Edition, Prentice-Hall, Inc., Englewood Cliffs, New Jersey.

Muckler, F. A., and Seven, S. A. (1992). "Selecting Performance Measures: "Objective" Versus "Subjective" Measurement," *Human Factors* 34(4), 441-455.

Neter, John, Wasserman, William, and Kutner, Michael. (1989). *Applied Linear Regression Models, Second Edition*, Richard D. Irwin, Inc., Homewood, Illinois.

NeuralWare. (1995a). *Reference Guide: Software Reference for Professional II/PLUS and NeuralWorks Explorer*, NeuralWare, Inc., Pittsburgh, Pennsylvania.

NeuralWare. (1995b). *Using NeuralWorks: A Tutorial for NeuralWorks Professional II/PLUS and NeuralWorks Explorer*, NeuralWare, Inc., Pittsburgh, Pennsylvania.

Osborne, W. M., and Rosenthal, L. (1985). "Metrics and Techniques to Measure Microcomputer Productivity," *Journal of Research of the National Bureau of Standards*, Vol. 90, July/August, 305-17.

Pasmore, W. A., and Purser, R. E. (1993). "Designing Work Systems for Knowledge Workers," *Journal for Quality and Participation* 16(4), 78-84.

Robinson, M. (1991). "Measuring the Bottom Line Impact of Decision Support Systems," *Information Executive* 4(2), 28(5).

Saaty, Thomas L. (1990). *The Analytic Hierarchy Process*, Second Edition, RWS Publications, Pittsburgh, Pennsylvania, Vol. 1.

SAS Institute, Inc. (1995). *JMP Introductory Guide Version 3.1*, Cary, North Carolina.

Scheaffer, Richard L., William Mendenhall, and Lyman Ott. (1986). *Elementary Survey Sampling*, Third Edition, Duxbury Press, Boston, Maryland, 324 pp.

Stone, M. (1974). "Cross-validatory choice and assessment of statistical predictions," *Journal of the Royal Statistical Society*, Vol. B36, pp. 111-133.

U.S. Army. (1989). *Intelligence Preparation of the Battlefield, FM 34-130*, Department of the Army, Washington, DC.

U.S. Army Training and Doctrine Command. (1996). *Military Operations Future Operational Capability*, Final Draft Pamphlet 525-66, Fort Monroe, Virginia.

Vokurka, R., and Fliedner, G. (1995). "Measuring Operating Performance: a Specific Case Study," *Production and Inventory Management Journal*, 38-43.

Williges, R. C., Williges, B. H., and Han, S. H. (1992). "Developing Quantitative Guidelines Using Integrated Data from Sequential Experiments," *Human Factors* 34(4).

Zurada, Jacek M. (1992). *Artificial Neural Systems*, West Publishing Company, St. Paul Minnesota.

BIBLIOGRAPHY

Driskell, J. E., and Salas, E. (1991). "Group Decision Making under Stress," *Journal of Applied Psychology* 76(3), 473-478.

Dunks, Capt J. J. (1995). *Development of a Metric Evaluation Instrument for U.S. in Measuring Organizational Effectiveness*, M.S. thesis, Air University Air Force Institute of Technology, Ohio.

Green, C., and Brown, W. (1993). "Well-Designed Performance Measurement is the Strongest Tool for Controlling Your Business Objectives," *Canadian Manager/ Manager Canadien* 18(4), 24-27.

Han, Sung, H., Williges, Robert, C., and Williges, Beverly H. (1990). "Empirical Models Based on Data from Several Experiment," *Human Factors Society Annual Meeting*, 1224-1228.

Hill, S. G., Iavecchia, H. P., Byers, J. C., Bittner, A. C., Jr., Zaklad, A. L., and Christ, R. E. (1992). "Comparison of Four Subjective Workload Rating Scales," *Proceedings of the Human Factors Society 34(4)*, 429-439.

Koury, LTC R. R, and Loerch, LTC A. G. (1992). *Army Program Value Added Analysis 94-99*, Technical Report CAA-SR-92-10, U.S. Army Concepts Analysis Agency, Washington, District of Columbia.

Loerch, LTC A. G., Koury, LTC R. R., and Maxwell, MAJ D. T. (1995). *Value Added Analysis for Army Equipment Modernization*, Research Paper CAA-RP-95-3, U.S. Army Concepts Analysis Agency, Bethesda, Maryland.

Meyer, Dr. M. D. (1995). *Alternative Performance Measures for Transportation Planning: Evolution Toward Multimodal Planning*, Technical Report FTA-GA-27-7000-95-1, Transportation Research and Education Center, Atlanta, Georgia.

Miller, T. A. (1995). *Developing an Integrated Model for Measuring Organizational Effectiveness in Light of the Government Performance and Results Act of 1993*, M.S. thesis, Air University Air Force Institute of Technology, Ohio.

Morgan, B.B., Glickman, A.S., Woodward, E. A., Blaiwes, A.S., and Salas, E. (1986). *Measurement of Team Behavior in a Navy Environment (Tech Report TR-86-014)*. Naval Training Systems Center, Human Factors Division, Orlando, Florida.

Pfleeger, S. L. (1995). "Maturity, Models and Goals: How to Build a Metrics Plan," *J. Systems Software* 31, pp 143-155.

Poore, J. H. (1988). "Derivation of Local Software Quality Metrics (Software Quality Circles)," *Software—Practice and Experience* 18(11), pp 1017-1027.

Ramey, J. (1995). "What Technical Communicators Think about Measuring Value Added: Report on a Questionnaire," *Technical Communication*, pp 40-51.

Sharman, P. (1992). "A Tool Kit for Continuous Improvement," *CMA Magazine* 66(4), pp 17-20.

Tan, Charles Y., and Iglewicz, Boris. (1999). "Measurement-Methods Comparisons and Linear Statistical Relationship," *Technometrics* Vol. 41, No. 3., pp 192-200.

Turban, Efraim, and Aronson, Jay E. (1998). *Decision Support Systems and Intelligence Systems*, Fifth Edition, Prentice Hall, Englewood Cliffs, New Jersey.

U.S. Army Command and General Staff College. (1994). *Command and Staff Decision Processes*, Student Text 101-5, Fort Leavenworth, Kansas.

APPENDIX A

CASE PROBLEM DESCRIPTION

Background

Unit movement assessments in the Intelligence Preparation of the Battlefield (IPB) process at the brigade and below organization levels are time intensive procedures typically performed manually under critical time constraints. The brigade intelligence officer (S2, Staff 2nd function) must depend largely on reconnaissance of available materials to address the commander's needs; however, the reconnaissance per se is really not sufficient for these requirements (Mikaloff 1996). Automated procedures that evaluate ground vehicle mobility exist but are not readily available at brigade and lower levels. Automation potentially offers increased quality of products, and analyses as well as time savings to the analyst. Consequently, it is important to evaluate the value added by incorporating these types of automated analyses into interactive and geographically - referenced systems that can be utilized in the IPB process at the brigade and below. Definitions of IPB and other related terms that concern this study are included in Appendix B.

Value added assessments of automated analysis can provide practical insight regarding the impact of automation on users and on processes and can reveal system development needs. Although many resources have been allocated to support digitization of the battlefield efforts, few studies aimed at systematically quantifying resultant value added via automation have been conducted. These types of studies constitute an important component in the overall digitization of the battlefield scheme as they aid in focusing future automation efforts which support system development and future force design.

The U.S. Army Engineer Research and Development Center, with over 40 years of research and development experience in ground vehicle mobility, conducted this study as part of the FY96 Army Study Program. This study was sponsored by the Office of the Deputy Chief of Staff for Intelligence (ODCSINT), Headquarters, Department of the Army. The study provides a prospective framework for evaluating the impact of decision support technologies and battlefield digitization on performance.

Study Objectives

The objectives of this study were to identify mobility-related IPB functions that can be performed using automated geographically-referenced means at brigade and below and to report resultant value-added to the IPB process, if any. The study incorporated formal hypotheses testing, analytical rigor, and soldier involvement. Time to conduct tasks, quality of products, insight into analyses, and perceived importance/priority regarding automation were assessed. The study focused on the value of using mobility assessment tools and applications in the IPB process which are tightly coupled with mobility assessments (i.e., line-of-sight from unit to destination or enemy unit).

This report is not an evaluation of any particular computer-based software system, nor does it provide a review of all computer-based software systems containing mobility assessment products. Emphasis is on mobility-related IPB functions rather than on software systems employed during the experiment. Over 200 Army officers participated in the study through controlled experiments and questionnaires.

The tasks involved in accomplishing the study objectives included the following: (1) formulating the problem, (2) developing questionnaires for obtaining information on intelligence users attitudes concerning automation of a broad range of mobility-related components of the IPB process, (3) designing experiments for generating pertinent data related to performance of staffs with and without automated IPB capabilities, and (4) analyzing and reporting results. The study was conducted using two approaches to obtain results from a representative cross-section of the Military Intelligence (MI) community.

Designed experiments were used in the first approach to compare staff performance based on automated versus manual IPB functionality. Experiments were carried out in conjunction with the Military Intelligence Officer Advanced Course (MIOAC) Brigade Operations and Intelligence (BOI) Section as identified in preliminary stages of the study. The second approach involved utilizing questionnaires to obtain information regarding perceived value in automating mobility-related IPB functions. Questionnaires were disseminated to participants in the MIOAC and active S2 staff throughout the Army.

The problem formulation, consisted of three sub-tasks: (1) becoming familiar with the IPB process, (2) identifying mobility-related components of the IPB process that can be automated, and (3) determining procedures for measuring value added when these components are automated. The determination of procedures to measure value added involved identification of characteristic measures and statistical methods for evaluation. As part of the problem formulation, test populations and environments were selected.

Survey Process

Overview

One aspect of assessing value added resulting from automating mobility- related components of the IPB process involved gauging perceptions of potential users in the MI community. Measurements were obtained using questionnaires. A questionnaire contains elements arranged in a fixed order and format that seek information regarding judgments, comparisons, and opinions of the target population (Meister 1985 or Scheaffer et al. 1986). Questionnaires were selected over an interview because they could be administered to a group of persons simultaneously, thus saving time and decreasing administrator requirements. In addition, they could be mailed or administered via telephone, allowing for expansion of the number and type of respondents. Statistical analysis of results (i.e. Sign Test, Cluster, Analysis of Variance) was conducted to evaluate responses and develop a prioritized list of IPB process elements ordered by relative importance for automation.

Target Population

The target population sampled consisted of respondents familiar with conducting the IPB process at the brigade organization level and below. The respondents were selected from two sources and were identified with sponsor guidance: (1) active Army S2 staff and (2) students participating in the Military Intelligence Officer Advanced Course

(MIOAC), Brigade Operations and Intelligence (BOI) section, at the U.S. Army Intelligence School, Fort Huachuca, Arizona. Sample size and selection were constrained by time available for study, MIOAC schedule and enrollment, sponsor guidelines regarding contact of Divisions, and operational requirements of Divisions. It was not feasible to obtain a list of the entire population of S2s, Assistant S2s, and MIOAC students from which to randomly sample respondents. Questionnaires were administered to the entire group of class participants in conjunction with controlled experiments during MIOAC1 (December-January), MIOAC3 (March-April), and MIOAC4 (May-June); MIOAC2 students participated in special assignments supporting Prairie Warrior exercises and, as such, were inaccessible. Further discussion of questionnaire respondents outside the MIOAC class is beyond the scope of this thesis.

Development and Design

Information gathered in the questionnaire pertained to attitudes toward automating components of the IPB process in general, perceptions regarding importance of automating specific IPB elements, and characteristics of respondents. General comments and identification of parts of the IPB process not specifically identified on the questionnaire were solicited as well. Additional questions specific to experimental roles were added to questionnaires administered to MIOAC students since they participated in controlled experiments utilizing automated IPB functions. Appendix D contains the MIOAC questionnaire form. The form was designed with input from the MI community and followed standard development practices such as those described in Meister (1985).

An important consideration in questionnaire design dealt with response errors. The Army Research Laboratory - Huachuca provided expertise by aiding in developing and reviewing the questionnaires with this consideration in mind.

In questionnaires for the MIOAC, background information related to branch of service, branch within service, experience, computer literacy, and other factors were obtained so as to investigate significant differences between subgroups regarding preferences for automated functionality. For example, respondents with more experience may place a higher importance on automating certain IPB functions when compared to respondents with limited experience. Distribution of response characteristics define the profile of the typical respondent. The questionnaire responses from each member of the teams were averaged to characterize the team member make-up of each team for this thesis research.

Administration Procedures

The questionnaire was administered by the study team to MIOAC participants at the conclusion of Practical Exercise (PE) 1 and again after PE 7 to investigate differences in responses regarding value added or importance of automating various IPB functions. As anticipated, not all those who returned questionnaires after PE 1 did so after PE 7 due to absences, lack of interest, or other considerations. The MIOAC3 and MIOAC4 students participated and provided 106 questionnaires. Each student was given a questionnaire form and instructions regarding its completion. Participants were asked to complete questionnaires in the classroom and return them to administrators. Returned questionnaires were not always filled-out completely. That is, some respondents did not answer one or more questionnaire items. Non-response data items were not included in the analysis.

Designed Experiments

Overview

Two experiments were conducted (in two phases each): one experiment for PE 1, *Mission Analysis*, and one for PE 7, *Production of Operations Order*. The *Production of Operations Order* included *Mission Analysis* through *Courses of Action Development/Analysis and Decision Brief*, hereafter referred to as *COA Analysis and Decision Brief*. Both phases of each experiment were carried out with students in MIOAC section 3 (MIOAC3) and section 4 (MIOAC4). The experiment involving PE 1 was concerned with determining whether automating certain mobility-related functions improved scores achieved in conducting *Mission Analysis*. The experiment involving PE 7 dealt with determining whether automating certain mobility-related functions improved the scores for *COA Analysis and Decision Brief*. In order to construct statistically robust Multi-Variant Linear Regression Model and Artificial Neural Network as required by this thesis, and an analysis of data requirements, revealed that the 12 samples taken from this *COA Analysis and Decision Brief* were insufficient to answer questions posed by this thesis research. Thus, only data from PE 1 and the *Mission Analysis* portion of PE 7 were used.

A randomized block design was chosen for each experiment. The blocking factors were brigade/regiment assignments (area of maneuver operations) and by course section (instruction differences). Blocking factors are used to partition out variability that may add noise to the measures of interest. Experimental teams or squads were grouped to compare data in like classifications, specifically the same area of operation and instruction emphasis. Computer-Assisted Squads (CAS) and Manual Means Squads (MMS) were randomly assigned within blocks.

The data collected during the experiments was of two forms, in-process elemental and final outcome. In-process elemental refers to data items collected during the conduct of the practical exercises. In-process data collection was used to gain insight into systematic differences that exist between CAS and MMS. Process observers trained in data collection methods collected information pertaining to time for task completion, effort put into product development, proportion of time spent on various tasks, thoroughness of analysis, and completeness of products. Table A1 provides a listing of these elemental data requirements.

The end product was measured using the *Mission Analysis* evaluation forms, or "cut-sheets," scored by "outside" evaluators (Appendix C). The MIOAC uses the cut-sheets to score the performance of a squad for each phase of the IPB process on the quality, completeness, and accuracy of the products. If a squad does not make the "cut" for a particular phase, it must execute that phase again until it "makes-the-cut." Outside evaluators were used to rate squad performance so as to not introduce bias regarding the utility of automation in the IPB process. Outside is defined as being independent of and not involved in the design and conduct of the experiments. Military rank of the evaluators ranged from senior Captains to Colonels (retired) and all were experienced with the IPB process. A given outside evaluator scored both squads within the block to obtain appropriate relative comparisons between CAS and MMS. The in-process data is given in Appendix E, Tables E1 and E2. The end product data is given in Appendix F, Table F18.

Table A1. Data Noted by Process Observers During Experiments

DATA REQUIREMENTS
Duration for producing final MCOO for each squad
During production of MCOO, duration for CAS computing and projecting severely-restricted/restricted areas for MCOO basis
During production of MCOO, duration for CAS to trace severely-restricted/restricted areas for MCOO basis
Color photograph of all other squads' in-process MCOO at the time of completion of the first MCOO
Color photograph of all completed MCOOs
Duration for producing Time Phase Lines for every squad
During production of Time Phase Lines, duration for computing and projecting Time Phase Lines for CAS
During production of Time Phase Lines, duration for CAS tracing Time Phase Lines
Number of what-ifs considered for Time Phase Lines for every squad
Color photograph of final Time Phase Lines products for every squad
Duration for every squad to create situation template
Number of queries for weapon fans made by CAS in determining weapon systems placement
Note any prolonged absence of a squad member (more than 30 minutes) and whether his or her assigned task is given to another member
Note squad's use of other (optional) computer assisted products; computer operators must notify data collectors when other products are being generated
Note the number of persons within the squad that work on each manual or computer assisted activity
Note the proportion of time involved persons contribute to any particular computer assisted activity
Start and stop times, including breaks, for the process of creating the products noted above
Number of Threat and Friendly Courses of Action (COA) developed
Time spent on COA development and analysis (PE 7 only)
Number of persons involved in Threat / Friendly COA development and analysis (not wargaming)
Time spent wargaming each COA (PE 7 only)
Number of persons involved in wargaming each COA (PE 7 only)

Data Collection Methods

For the in-process elements, a stopwatch was used to record time to the nearest second. Photographs were taken of MCOO and Time Phase Lines when completed. Placards were used to identify wall clock time, squad number, brigade assignment, and PE number for photographed products. Cut-sheets were developed with MIOAC instructors for *Mission Analysis* and *COA Analysis and Decision Brief* based on versions used by previous instructors.

Site Layout

The experiments were conducted in two large rooms connected by a doorway on a common wall; each room measured approximately 24 feet by 48 feet. See Figure A1. The two rooms provided the seven simulated tactical operation centers (TOC) needed for the practical exercises. One room housed the automated squads, the other housed the manual squads. Each room was divided into four roughly equal sized cells. Squads occupied each of the four cells in the manual room. Three cells within the automated room were each used by a CAS; a large area adjacent to the door connecting the two rooms was left vacant for use by the instructors and to allow a large walkway between the two rooms. A smaller walkway was mirrored in the MMS classroom.

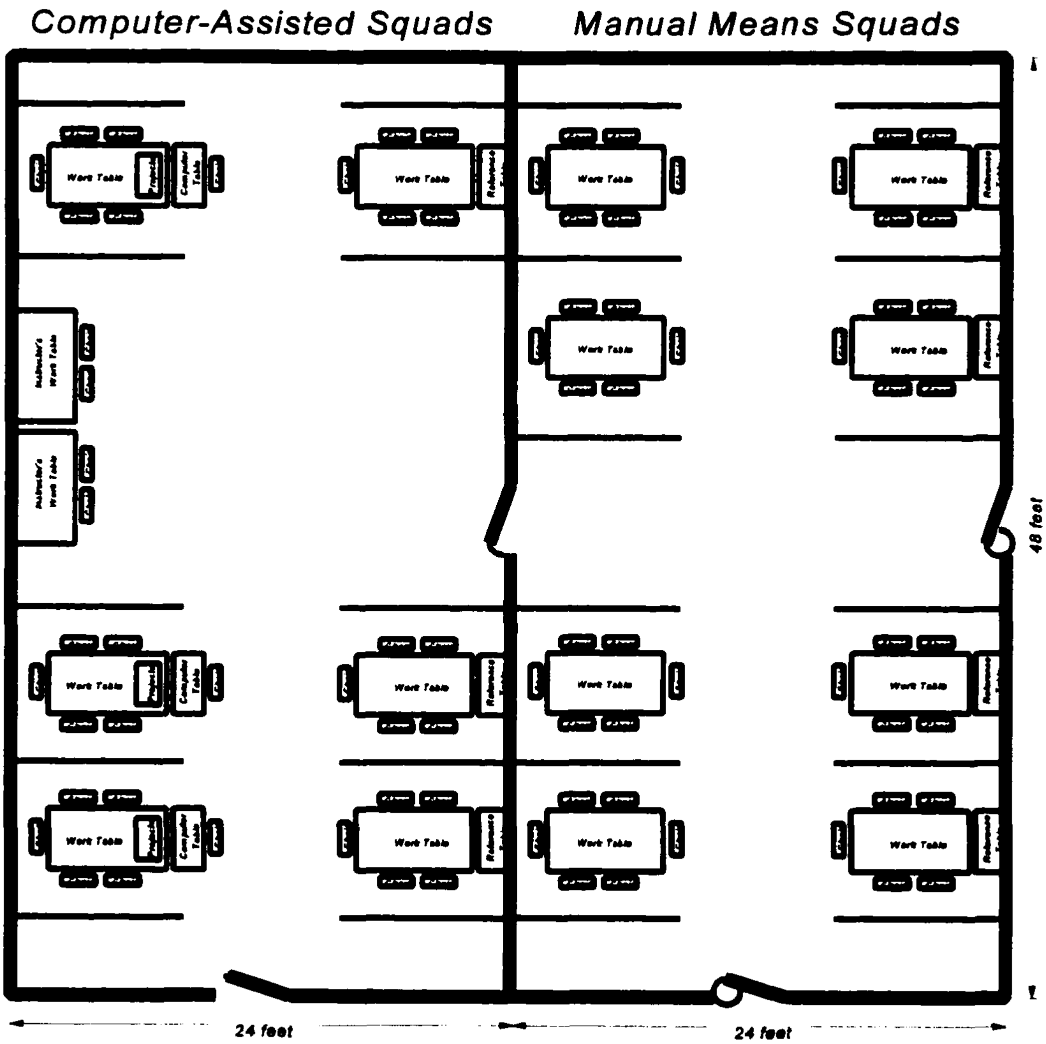


Figure A1. Site Layout

APPENDIX B

DEFINITIONS

The following are definitions of IPB and other related terms. Definitions marked with an * were obtained from FM 34-130.

a. *Area of interest (AI). The geographical area from which information and intelligence are required to permit planning or successful conduct of the command's operation. The AI is usually larger than the command's AO and battle space; it includes any threat forces or characteristics of the battlefield environment that will significantly influence accomplishment of the command's mission.

b. *Area of operations (AO). That portion of an area of conflict necessary for military operations. AOs are geographical areas assigned to commanders which they have responsibility and in which they have the authority to conduct military operations.

c. *Avenues of approach (AA). An air or ground route of an attacking force of a given size leading to its objective or key terrain in its path.

d. *Battlefield Operating System (BOS). The major functions performed by the force on the battlefield to successfully execute Army operations in order to accomplish military objectives. BOS forms a framework for examining complex operations in terms of functional operating systems. The systems include maneuver, fire support, air defense, command and control, intelligence, mobility and survivability, and combat service support.

e. Comprehensive Analysis Mobility Model - Developmental (CAMMS-D). Mobility model application which was developed at Waterways Experiment Station and consists not only of NRMM, but also a GIS user interface and graphical display mechanisms necessary to create useful automated decision tools.

f. *Course of Action (COA). A possible plan open to an individual or commander that would accomplish or is related to accomplishment of the mission. A COA is initially stated in broad terms with the details determined during staff wargaming. To develop COAs, the staff must focus on key decisions. COAs include five elements: WHAT (the type of operation), WHEN (the time the action will begin), WHERE (boundaries, axis, etc.), HOW (the use of assets), and WHY (the purpose of desired and state).

g. *Decision point (DP). The point in space and time where the commander or staff anticipated making a decision concerning a specific friendly COA. DPs are usually associated with threat force activity or the battlefield environment and are therefore associated with one or more NAIs. DPs also may be associated with the friendly force and the status of ongoing operations.

h. *Decision Support Template (DST). A graphic record of wargaming. The DST depicts DPs, time lines associated with movement of forces and the flow of the operation, and other key items of information required to execute a specific friendly COA.

i. *Doctrinal template. A model based on postulated threat doctrine. Doctrinal templates illustrate the disposition and activity of threat forces and assets (HVTs) conducting a particular operation unconstrained by the effects of the battlefield environment. They represent the application of threat doctrine under ideal conditions. Ideally, doctrinal templates depict the threat's normal organization for combat, frontages, depth, boundaries, and other control measures, assets available from other commands, objective depths, engagement areas, battle positions, and so forth. Doctrinal templates are usually scaled to allow ready use on a map background. They are one part of a threat model.

j. *Electronic warfare. Consists of three subcomponents: electronic attack (EA), electronic warfare support (ES), and electronic protection (EP).

k. *Event matrix. A description of the indicators and activity expected to occur in each NAI. It normally cross-references each NAI and indicator with the times they are expected to occur and the COAs they will confirm or deny. There is no prescribed format.

l. *High-payoff target (HPT). Target whose loss to the threat will contribute to the success of the friendly COA.

m. *High-value target (HVT). Assets that the threat commander requires for the successful completion of a specific COA.

n. *Information requirement. An intelligence requirement of lower priority than the PIR of lowest priority.

o. *Intelligence preparation of the battlefield (IPB). The systematic, continuous process of analyzing the threat and environment in a specific geographic area. IPB is designed to support the staff estimate and military decision making process. Most intelligence requirements are generated as a result of the IPB process and its interrelation with the decision making process.

p. *Intelligence requirement (IR). A requirement for intelligence to fill a gap in the command's knowledge and understanding of the battlefield or threat forces. Intelligence requirements are designed to reduce the uncertainties associated with successful completion of a specific friendly COA; a change in the COA usually leads to a change in intelligence requirements. Intelligence requirements that support decisions which affect the overall mission accomplishment (such as choice of a COA, branch, or sequel) are designated by the commander as PIR. Less important intelligence requirements are designated as IR.

q. *Key terrain. Any locality or area the seizure, retention, or control of which affords a marked advantage to either combatant.

r. *Lines of Communication (LOC). All the routes (land, water, and air) that connects an operating military force with one or more bases or operations and along which supplies and military forces move. Note that not all roads and rails are LOCs; some are unsuited, others may be suitable but not used. Note also that in this context, a communications center is an area where LOCs converge, such as transshipment points or hub-pattern cities.

s. *Mobility Corridor (MC). Areas where a force will be canalized due to terrain restrictions. They allow military forces to capitalize on the principles of mass and speed and are therefore relatively free of obstacles.

t. *Modified combined obstacle overlay (MCOO). A product used to depict the battlefield's effects on military operations. It is normally based on a product depicting all obstacles to mobility, modified to also depict the following, which are not prescriptive nor inclusive.

* Cross-country mobility restrictions (such as RESTRICTED).
* Objectives.
* AAs and Mobility corridors.
* Likely locations of counter-mobility obstacle.
* Defensible terrain.
* Likely engagement areas.
* Key terrain.

u. *Named Area of Interest (NAI). The geographical area where information that will satisfy a specific information requirement can be collected. NAIs are usually selected to capture indications of threat COAs but also may be related to conditions of the battlefield.

v. NATO Reference Mobility Model (NRMM). A computer-based collection of equations and algorithms designed to predict the steady-state operating capability of a given vehicle operating in a given terrain. (Ahlvin et al.)

w. *OCOKA. A commonly used acronym and mnemonic for the military aspects of terrain. The acronym does not dictate the order in which the factors are evaluated; use the order best suited to the situation at hand. The military aspects of terrain are observation and fields of fire, concealment and cover, obstacles, key terrain, and avenues of approach.

x. Outside Evaluators. Evaluators who were independent of and not involved in the design and conduct of the experiments in the MIOAC.

y. *Phase line. A line used for control and coordination of military operations. It is usually a recognizable terrain feature extending across the zone of action. Units normally report PLs, but do not halt unless specifically directed. PLs often are used to prescribe the timing of delay operations.

z. *Priority Intelligence Requirement (PIR). An intelligence requirement associated with a decision that will affect the overall success of the command's mission. PIR are a subset of intelligence requirements of a higher priority than information requirements. PIR are prioritized among themselves and may change in priority over the course of the operation's conduct. Only the commander designates PIR.

aa. *Restricted. A classification indicating terrain that hinders movement. Little effort is needed to enhance mobility through restricted terrain but units may have difficulty maintaining preferred speeds, moving in combat formations or transitioning from one formation to another. A force can generally use administrative or march formations through restricted terrain with only minimal delay.

bb. Risk-based mobility predictions. Predictions which factor risk (conservative estimates vs. non-conservative estimates) into perspectives of restricted/severely restricted areas by evaluating best-case, worst-case, and in-between cases. Risk-based methodology is based on the fact that there is a range in vehicle speed outcomes rather than an exact outcome for a given vehicle (or unit) operating over a specified areas.

cc. *Severely restricted. A classification indicating terrain that severely hinders or slows movements in combat formations unless some effort is made to enhance mobility. Severely restricted terrain includes manmade obstacles, such as minefields and cities, as well as natural barriers. Severely restricted terrain generally slows or impedes administrative and march formations.

dd. *Situation template. Depictions of assumed threat dispositions, based on threat doctrine and the effects of the battlefield, it the threat should adopt a particular COA. In effect, they are the doctrinal templates depicting a particular operation modified to account for the effects of the battlefield environment and the threat's current situation (training and experience levels, logistic status, losses, dispositions). Normally, the situation template depicts threat units two levels of command below the friendly force as well as the expected locations of HVTs. Situations templates use TPLs to indicate movement of forces and the expected flow of the operation. Usually, the situation template depicts a critical point in the COA. Situation templates are one part of a threat COA model. Models may contain more than one situation template.

ee. *Target area of interest (TAI). The geographical area where HVTs can be acquired and engaged by friendly forces. Not all TAIs will form part of the friendly COA; only TAIs associated with HPTs are of interest to the staff. These are identified during staff planning and wargaming. TAIs differ from engagement areas in degree. Engagement areas plan for the use of all available weapons; TAIs might be engaged by a single weapon.

ff. Time contour analysis. This analysis produces overlays that depict the area that can be covered by a vehicle or group of vehicles starting at a given point in specified time intervals such as hourly intervals. Time contour analysis are derived from vehicle speed predications.

gg. *Time phase line (TPL). A line used to represent the movement of forces or the flow of an operation over time. It usually represents the location of forces at various increments of time, such as lines that show unit locations at 2-hour intervals. TPLs should account for the effects of the battlefield environment and the anticipated effects of contract with other forces. For example, TPLs depicting threat movement through an area occupied by friendly forces should use movement rates based on a force in contact with the enemy rather than convoy movement speeds.

hh. *Unrestricted. A classification indicating terrain that is free of restrictions to movement.

APPENDIX C

MIOAC QUESTIONNAIRE

IPB PROCESS VALUE-ADDED QUESTIONNAIRE

STUDENT ID ___________ (last 6 digits in Social Security Number)
DATE ___________
INSTRUCTIONS: Please respond to the following questions by circling the correct response or filling in the blanks.

1. What age group are you in?
   (1) 20-25 (2) 26-30 (3) 31-35 (4) 36-40 (5) 41-45 (6) over 45

2. What is your gender? _______

3. Are you a U.S. officer? _____ If not, what country?_____________

4. What is your rank? (1) 2LT (2) 1LT (3) CPT (4) MAJ (5) LTC (6) COL
   Number of years at your current rank? _______________

5. Do you have prior enlisted service? ______
   If yes, how many years? _________ What MOS? ___________

6. How many years of service do you have (excluding prior enlisted service)?
   (1) 1-5 (2) 6-10 (3) 11-15 (4) 16-20 (5) over 20

7. What is your branch of service? (1) Army (2) Navy (3) Marines
   (4) Air Force
   What is your branch within your service?
   (1) CHEM (2) ADA (3) IN (4) AR (5) MI (6) FA (7) AVN
   (8) Other ________

8. Are you an OTC graduate? _____
   If yes, what is your most recent graduation date? ___________
   If yes, what was your previous branch?
   (1) CHEM (2) ADA (3) IN (4) AR (5) MI (6) FA (7) AVN
   (8) Other _________

9. What is your highest level of education?
   (1) high school (2) bachelor's degree (3) master's degree
   (4) doctoral degree
   How many hours have you completed for credit beyond your highest degree? ______
   What was your field(s) of study in college? ___________________________

10. List any military automated systems that you have used (i.e., WARRIOR, ASAS, etc.). ____

11. How comfortable are you with using computers?
    (1) uncomfortable (2) not very comfortable (3) somewhat comfortable
    (4) comfortable

12. What is your squad (this course)? __________ Which PE # have you just completed?_________

    What brigade were you in for this PE? (1) 2nd (2) 3rd (3) 9th (4) AVN (5) Other

    Position for this PE: (1) S2 (2) S3 (3) A/S2 (4) A/S3 (5) XO (6) FS0 (7) ADO (8) ENG (9) Other _____

13. Were you in the group that participated in the computer assisted IPB process during this PE? ______

14. Outside of this class, have you served as a battalion S2 before? ______
    Conducted the manual IPB process? _____ If yes, how many times within the last 2 years?______

15. What level of expertise in the IPB process do you feel you have?
    (1) inexperienced (2) not very experienced (3) somewhat experienced
    (4) experienced

16. Below is a list of IPB process components.
    A. Rate the importance of automating each component. State reasons for your response.
    B. Uniquely rank ALL of the following components in the order of importance.
    (1=First, 18=Last) Rank importance If Very or Not, why?
    Very Somewhat NOT (computer trust, human decision, etc)

Rank	Component	Very	Somewhat	Not	If Very or Not, why? (computer trust, human decision, etc.)
	Key terrain				
	Avenues of approach				
	Ground mobility corridors				
	Lines of communication				
	Areas of concealment				
	Areas of observation				
	Restricted areas				
	Named areas of interest				
	*Risk-based mobility				
	*Risk-based mobility estimates				
	Areas to recon				
	Intervisibility lines				
	Weapons range fans				
	Electronic warfare				
	Line of sight				
	Route selection				
	Time phase lines				
	Enemy order of battle				
	3-D terrain views				

* Risk-based refers to best-case/worst-case scenarios.

17. Below is a list of IPB products.
    A. Rate the importance of automating each component. State reasons for your response.
    B. Uniquely rank ALL of the following components in the order of importance. (1=First, 4=Last)

Rank	**IPB Product**	**Very**	**Somewhat**	**Not**	**If Very or Not, why? (computer trust, human decision, etc.)**
	Modified Combined Obstacle Overlay				
	Situation templates				
	Doctrinal templates				
	Event templates				

19. Regarding your ability to perform IPB in an effective and timely manner, automation (computer assistance) of IPB functions listed in 16 & 17 will:
    (1) help (2) hinder (3) not make a difference.

20. Have you been to National Training Center before? ________ If yes, how many rotations and/or assignments have you completed? ___________

21. General comments.____
    ____
    ____
    ____
    ____
    ____

APPENDIX D

MISSION ANALYSIS EVALUATION FORM

CUT SHEET: MISSION ANALYSIS

Evaluators, score each element according to the following (scores of 2 & 4 permitted):
1 = nogo 3 = marginal go 5 = go

S-2 BRIEF

____ {B}1. Oriented commander to the terrain.

____ {B}2. Briefed/explained brigade's area of operation, area of interest, and battlespace.

____ {B}3. Briefed weather and light data and effects on friendly and enemy COAs.

____ {B}4. Briefed military aspects of terrain using the avenue, box, or belt technique; reflected proper amount of detail. (OCOKA - observation and fields of fire, concealment and cover, obstacles, key terrain, avenues of approach)

____ {B}5. Accurately identified restricted vs. severely restricted terrain in the proper amount of detail.

____ {D}6. Identified key terrain and explained its significance.

____ {D}7. Identified and categorized mobility corridors/As from TAA (tactical assembly area) to OBJ (objective) and enemy counterattack routes.

____ {D}8. How well were time phase lines applied to terrain?

____ {D}9. Identified air avenues of approach (fixed and rotary).

____ {A}10. MCOO (modified combined obstacle overlay) was neat, legible, and clear and reflected the proper amount of detail.

____ {D}11. Was line of sight and terrain utilized in placement of enemy unit locations?

____ {E}12. Developed two enemy courses of action in detail. Identified HVTs (high value targets), center of gravity culmination point and decisive point and defeat mechanism. Developed by enemy BOS (battlefield operating system). Developed time of proposed enemy action.

_____ {F}13. Addressed enemy capabilities and vulnerabilities.

_____ {F}14. Developed initial list of information requirements and prioritized them in order of importance.

S-3 BRIEF

_____ {F}15. Addressed specified, implied and mission essential task.

_____ {F}16. Address task and purpose in restated mission statement. Who, What, When, Where, and Why.

_____ {C}17. Address limitations and restrictions.

_____ {F}18. Recommend command and control, friendly force, and initial intelligence requirements.

_____ {F}19. Determine broad C2 considerations.

_____ {F}20. Propose acceptable risk for the commander.

_____ {C}21. Determine critical facts and assumptions.

Note: Hierarchical label {A-F} inserted before question number to help with explanation of other data and cross-referencing.

APPENDIX E

TEAM CHARACTERISTICS AND PERFORMANCE DATA

Table E1. Mission Analysis Squad Average Characteristics (Selected)

Identifiers			Characteristics Considered							
Month	Squad	BD	Age 20-25 yrs: 1 26-30 yrs: 2 31-35 yrs: 3	Gender Female: 1 Male:2	Rank 2Lt: 1 1Lt: 2 CPT: 3	Years of Service 1-5 : 1 6-10 : 2	OTC Grad No: 1 Yes: 2	Computer Skill Level	IPB Experience	NTC Rotations
March	3	3	2.1	1.8	2.9	1.1	1.8	3.8	2.5	1.4
March	2	3	2.3	1.7	2.9	1.4	1.6	3.3	2.4	1.0
March	4	2	2.0	1.6	2.8	1.6	1.6	3.2	3.0	0.5
March	6	2	2.1	1.8	2.9	1.2	1.8	4.0	2.6	0.8
March	5	9	2.4	1.9	3.0	1.4	1.8	3.5	2.9	1.9
March	1	9	1.9	1.8	2.7	1.0	1.6	3.2	2.7	4.3
April	1	3	1.9	1.8	2.8	1.0	1.4	3.2	2.7	5.9
April	5	3	2.3	1.9	3.0	1.6	1.4	3.3	2.7	0.3
April	2	3	2.1	1.8	2.9	1.2	1.6	3.4	3.0	1.2
April	7	3	2.1	2.0	3.0	1.0	1.7	4.0	2.9	2.6
April	6	2	2.1	1.8	2.8	1.4	1.4	4.1	2.9	1.3
April	4	2	2.1	1.6	2.8	1.8	1.4	3.4	3.0	5.2
May	6	3	2.1	1.9	2.7	1.1	1.6	3.6	2.9	1.1
May	5	3	2.4	1.9	2.6	1.3	1.8	3.6	2.7	4.6
May	2	9	2.3	1.6	3.0	1.1	1.6	2.9	3.0	0.9
May	1	9	1.7	1.6	2.4	1.3	1.8	3.4	2.7	1.0
May	4	2	2.0	1.9	2.4	1.4	1.6	3.3	3.0	1.0
May	3	2	2.1	1.9	2.4	1.0	1.8	3.1	2.7	0.4
June	3	3	2.1	1.9	2.4	1.0	1.1	3.0	2.4	1.0
June	4	3	2.0	1.9	2.4	1.4	1.4	3.3	3.0	1.3
June	1	2	1.7	1.7	2.5	1.3	1.5	3.5	2.4	1.3
June	2	2	2.2	1.7	3.0	1.0	1.7	2.8	3.0	1.5
June	5	2	2.4	1.9	2.6	1.3	1.4	3.6	2.7	1.6
June	6	2	2.1	2.0	2.9	1.1	1.3	3.4	3.0	1.0

Note: Grayed rows indicate data records removed from field data set.

IPB — Intelligence Preparation of the Battlespace, NTC — National Training Center

See Appendix C for cross referencing of column headings with questions.

Table E2. Mission Analysis In-Process Performance Parameters and Values

IDENTIFIERS			PERFORMANCE PARAMETERS CONSIDERED							
Month	Squad	BD	CAS (0) MMS (1)	Members on MCOO (ea)	MCOO Time (hrs)	Members on TPL (ea)	Whatifs for TPL (ea)	TPL Time (hrs)	Members on SITMAP (ea)	SITMAP Time (hrs)
March	3	3	0	4	2.98	2	1	0.27	3	3.04
March	2	3	1	1	2.27	1	1	0.55	1	2.75
March	4	2	0	4	3.36	1	2	0.38	1	2.03
March	6	2	1	3	2.62	2	2	1.08	2	0.80
March	5	9	0	2	1.62	1	2	0.50	2	1.41
March	1	9	1	2	4.90	2	1	0.57	1	0.91
April	1	3	0	1	1.55	1	2	0.28	2	1.35
April	5	3	1	3	1.60	1	2	1.20	3	1.40
April	2	3	0	2	5.90	1	6	0.60	1	3.10
April	7	3	1	2	2.26	1	2	0.58	1	1.53
April	6	2	0	1	1.50	1	1	0.40	2	1.38
April	4	2	1	4	1.87	3	3	1.46	3	3.77
May	6	3	0	2	2.12	1	1	0.30	2	1.75
May	5	3	1	3	3.31	2	1	0.45	3	2.03
May	2	9	0	3	4.73	1	1	0.42	1	2.50
May	1	9	1	3	2.20	2	1	0.11	1	0.67
May	4	2	0	2	3.82	1	1	1.13	1	1.25
May	3	2	1	4	2.93	0	0	0.00	4	3.09
June	3	3	0	3	1.49	1	2	0.35	2	1.25
June	4	3	1	1	1.20	1	1	0.22	2	2.65
June	1	2	0	3	0.77	2	2	0.46	3	0.89
June	2	2	1	2	2.45	2	2	0.50	2	2.75
June	5	2	0	2	2.86	2	1	1.71	3	1.10
June	6	2	1	2	2.22	1	2	0.28	2	4.65

Note: Grayed rows indicate data records removed from field data set.

BD — Brigade, CAS — Computer Aided Squad, MMS — Manual Means Squad,
MCOO — Modified Combined Obstacle Overlay, TPL — Time Phased Line, SITMAP — Situational Map
See Appendix B for definitions of column headings.

APPENDIX F

EVALUATION QUESTIONS AND SCORES

Full Description of cut-sheet question groupings for Mission Analysis

{A} Developing neat, legible, and clearly written documentation and map overlays. Includes MCOO, terrain briefing, and intelligence estimate.

{B} Orienting Commander to the terrain. Includes explanation of brigade's battlespace, weather/light conditions and effects on possible friendly and enemy COAs. Used avenue, box, or belt technique with proper amount of detail for briefing terrain (OCOKA). Articulated and correctly identified restricted and severely restricted terrain.

{F} Orienting Commander to the mission and presenting recommendations. Includes addressing specified, implied, essential tasks, and purpose. Recommending intelligence requirements, C2 considerations, and acceptable risk factors.

{C} Presenting assumptions, limitations, critical facts, and restrictions.

{D} Identifying objectives, mobility, and terrain utilization. Includes avenues of approach, mobility corridors, key terrain, counterattack routes, air avenues of approach (fixed and rotary), time phase lines, and using line-of-sight for locating likely battle positions.

{E} Developing two enemy courses of action. Includes identifying HVTs, center of gravity culmination point, decisive point, and defeat mechanism. Addressing enemy capabilities, vulnerabilities, and BOS. Developing event matrix and template.

-9 Base factor overwhelmingly less important than other factor

-7 Base factor relative importance very weak

-5 Base factor essentially less important than other factor

-3 Base factor moderately less important than other factor

1 Base factor roughly equivalent in importance to other factor

3 Base factor moderately more important than other factor

5 Base factor essentially more important than other factor

7 Base factor relative importance very strong

9 Base factor overwhelmingly more important than other factor

Note: Hierarchical label {A-F} inserted before question number in Appendix D to help with explanation of other data and cross-referencing. The form used for the pairwise comparisons is as follows.

Other Factor

Base Factor	Developing neat, legible, and clearly written documentation & map overlays	Orienting Comman-der to the terrain	Orienting Commander to the mission and presenting recommendations	Presenting assumptions, limitations, critical facts, and restrictions	Identifying objectives, mobility, and terrain utilization	Developing two enemy courses of action
Developing neat, legible, and clearly written documentation and map overlays	1					
Orienting Commander to the terrain		1				
Orienting Commander to the mission and presenting recommendations			1			
Presenting assumptions, limitations, critical facts, and restrictions				1		
Identifying objectives, mobility, and terrain utilization					1	
Developing two enemy courses of action						1

Table F1. SME 1 (Relative Rank)

	Paired Question					
Base Question	A	B	F	C	D	E
A	1.00	9.00	3.00	1.00	9.00	-3.00
B		1.00	-3.00	-3.00	-3.00	-7.00
F			1.00	7.00	1.00	3.00
C				1.00	1.00	-7.00
D					1.00	-3.00
E						1.00

Table F2. SME 1 (Transform)

	Paired Question					
Base Question	A	B	F	C	D	E
A	1.00	8.00	2.00	0.00	8.00	-2.00
B		1.00	-2.00	-2.00	-2.00	-6.00
F			1.00	6.00	0.00	2.00
C				1.00	0.00	-6.00
D					1.00	-2.00
E						1.00

Table F3. SME 2 (Relative Rank)

	Paired Question					
Base Question	A	B	F	C	D	E
A	1.00	3.00	5.00	5.00	3.00	7.00
B		1.00	3.00	3.00	1.00	5.00
F			1.00	3.00	1.00	5.00
C				1.00	-3.00	5.00
D					1.00	7.00
E						1.00

Table F4. SME 2 (Transform)

	Paired Question					
Base Question	A	B	F	C	D	E
A	1.00	2.00	4.00	4.00	2.00	6.00
B		1.00	2.00	2.00	0.00	4.00
F			1.00	2.00	0.00	4.00
C				1.00	-2.00	4.00
D					1.00	6.00
E						1.00

Table F5. SME 3 (Relative Rank)

	Paired Question					
Base Question	A	B	F	C	D	E
A	1.00	-7.00	9.00	1.00	1.00	1.00
B		1.00	9.00	5.00	5.00	5.00
F			1.00	1.00	1.00	1.00
C				1.00	1.00	1.00
D					1.00	3.00
E						1.00

Table F6. SME 3 (Transform)

	Paired Question					
Base Question	A	B	F	C	D	E
A	1.00	-6.00	8.00	0.00	0.00	0.00
B		1.00	8.00	4.00	4.00	4.00
F			1.00	0.00	0.00	0.00
C				1.00	0.00	0.00
D					1.00	2.00
E						1.00

Table F7. SME 4 (Relative Rank)

	Paired Question					
Base Question	A	B	F	C	D	E
A	1.00	-5.00	1.00	-7.00	5.00	5.00
B		1.00	1.00	-3.00	-3.00	1.00
F			1.00	-3.00	1.00	1.00
C				1.00	-3.00	3.00
D					1.00	1.00
E						1.00

Table F8. SME 4 (Transform)

	Paired Question					
Base Question	A	B	F	C	D	E
A	1.00	-4.00	0.00	-6.00	4.00	4.00
B		1.00	0.00	-2.00	-2.00	0.00
F			1.00	-2.00	0.00	0.00
C				1.00	-2.00	2.00
D					1.00	0.00
E						1.00

Table F9. SME 5 (Relative Rank)

	Paired Question					
Base Question	A	B	F	C	D	E
A	1.00	-5.00	1.00	1.00	-5.00	-9.00
B		1.00	3.00	3.00	1.00	-3.00
F			1.00	1.00	-3.00	-5.00
C				1.00	-3.00	-5.00
D					1.00	-3.00
E						1.00

Table F10. SME 5 (Transform)

Paired Question						
Base Question	A	B	F	C	D	E
A	1.00	-4.00	0.00	0.00	-4.00	-8.00
B		1.00	2.00	2.00	0.00	-2.00
F			1.00	0.00	-2.00	-4.00
C				1.00	-2.00	-4.00
D					1.00	-2.00
E						1.00

Table F11. SME 6 (Relative Rank)

Paired Question						
Base Question	A	B	F	C	D	E
A	1.00	1.00	-3.00	1.00	1.00	-3.00
B		1.00	-3.00	1.00	1.00	-3.00
F			1.00	3.00	1.00	5.00
C				1.00	3.00	5.00
D					1.00	1.00
E						1.00

Table F12. SME 6 (Transform)

Paired Question						
Base Question	A	B	F	C	D	E
A	1.00	0.00	-2.00	0.00	0.00	-2.00
B		1.00	-2.00	0.00	0.00	-2.00
F			1.00	2.00	0.00	4.00
C				1.00	2.00	4.00
D					1.00	0.00
E						1.00

Table F13. SME 7 (Relative Rank)

	Paired Question					
Base Question	A	B	F	C	D	E
A	1.00	-5.00	-7.00	-7.00	-3.00	-5.00
B		1.00	-5.00	-5.00	3.00	-5.00
F			1.00	3.00	5.00	3.00
C				1.00	-5.00	-3.00
D					1.00	-5.00
E						1.00

Table F14. SME 7 (Transform)

	Paired Question					
Base Question	A	B	F	C	D	E
A	1.00	-4.00	-6.00	-6.00	-2.00	-4.00
B		1.00	-4.00	-4.00	2.00	-4.00
F			1.00	2.00	4.00	2.00
C				1.00	-4.00	-2.00
D					1.00	-4.00
E						1.00

Table F15. SME Transform Average

	Paired Question					
Base Question	A	B	F	C	D	E
A	1.00	-1.14	0.86	-1.14	1.14	-0.86
B		1.00	0.57	0.00	0.29	-0.86
F			1.00	1.43	0.29	1.14
C				1.00	-1.14	-0.29
D					1.00	0.00
E						1.00

Table F16. AHP Matrix

	Paired Question					
Base Question	A	B	F	C	D	E
A	1.00	0.47	1.86	0.47	2.14	0.54
B	2.14	1.00	1.57	1.00	1.29	0.54
F	0.54	0.64	1.00	2.43	1.29	2.14
C	2.14	1.00	0.41	1.00	0.47	0.78
D	0.47	0.78	0.78	2.14	1.00	1.00
E	1.86	1.86	0.47	1.29	1.00	1.00
Sum	8.15	5.74	6.08	8.32	7.18	6.00

Table F17. Normalized by Question

	Paired Question								
Base Question	A	B	F	C	D	E	AVG	No. of Question	Question Weight
A	0.12	0.08	0.31	0.06	0.3	0.09	0.16	1	0.1589
B	0.26	0.17	0.26	0.12	0.18	0.09	0.18	5	0.0362
F	0.07	0.11	0.16	0.29	0.18	0.36	0.19	6	0.0325
C	0.26	0.17	0.07	0.12	0.06	0.13	0.14	2	0.0683
D	0.06	0.14	0.13	0.26	0.14	0.17	0.15	5	0.0295
E	0.23	0.32	0.08	0.15	0.14	0.17	0.18	2	0.0907
							1.00	21	

Table F18. Subject Matter Expert Scores by Question

MONTH	SQUAD	BG	SUBJECT MATTER EXPERT SCORES BY QUESTION																					AVG
			B1	B2	B3	B4	B5	D6	D7	D8	D9	A10	D11	E12	E13	F14	F15	F16	C17	F18	F19	F20	C21	
March	3	3RD	5	5	5	5	5	5	5	1	5	5	5	3	5	1	5	5	3	3	3	5	5	4.238
March	2	3RD	1	3	5	3	3	5	1	1	3	3	5	1	3	1	5	5	5	1	1	5	5	3.095
March	4	2ND	3	3	1	5	3	3	5	3	5	5	1	3	1	3	5	3	5	3	3	3	3	3.286
March	6	2ND	3	1	3	1	1	3	3	1	1	3	5	3	5	3	5	3	3	3	3	3	3	2.810
March	5	9TH	5	5	5	5	5	5	5	1	5	5	1	5	5	5	5	5	5	1	1	1	5	4.048
March	1	9TH	5	3	5	3	5	5	5	5	5	5	3	5	5	5	5	1	5	3	1	1	1	3.857
April	1	3RD	4	3	4	3	3	3	4	3	4	3	3	4	3	3	4	5	3	4	2	2	3	3.333
April	5	3RD	2	2	3	4	3	4	3	3	4	4	3	3	4	2	4	5	4	1	1	3	4	3.143
April	2	3RD	3	2	2	2	2	1	3	1	1	2	1	2	2	1	3	3	3	1	1	3	4	2.048
April	7	3RD	3	3	3	2	4	2	3	3	2	4	2	4	2	1	2	0	3	0	0	2	2	2.611
April	6	2ND	4	5	4	4	4	3	4	5	5	4	5	4	3	4	4	4	4	5	4	5	5	4.238
April	4	2ND	4	4	3	3	4	4	2	1	2	2	1	1	0	0	4	4	4	0	4	4	3	3.000
May	6	3RD	5	5	3	5	5	5	5	5	5	5	3	3	5	3	5	5	3	3	3	3	5	4.238
May	5	3RD	5	5	3	3	1	3	3	1	1	1	3	3	3	3	5	5	3	5	3	3	3	3.095
May	2	9TH	3	3	3	3	5	3	3	3	3	5	3	3	3	3	3	3	3	5	3	5	3	3.381
May	1	9TH	3	3	3	3	3	5	3	1	1	3	3	1	1	3	5	3	5	3	3	3	3	2.905
May	4	2ND	5	3	3	1	3	1	5	1	5	5	1	1	1	3	3	3	1	3	1	3	5	2.714
May	3	2ND	3	3	1	1	1	3	3	1	5	1	1	3	1	3	3	3	1	1	1	1	5	2.143
June	3	3RD	5	4	4	5	4	5	5	5	4	4	4	5	5	4	4	4	4	5	5	3	4	4.381
June	4	3RD	5	2	4	4	3	4	4	5	4	5	1	5	3	3	5	5	2	5	4	3	3	3.762
June	1	2ND	5	5	5	5	4	5	5	4	4	5	4	4	5	5	5	5	5	3	2	2	2	4.238
June	2	2ND	4	4	4	3	2	4	4	2	3	5	3	3	3	5	4	4	4	5	5	5	5	3.857
June	5	2ND	5	5	3	5	5	5	5	5	4	5	4	3	5	3	5	5	5	3	1	3	5	4.238
June	6	2ND	5	5	3	5	5	5	5	3	4	5	5	5	5	5	5	5	5	5	3	5	5	4.667

Note: Questions with a value of "0" (no response) were excluded from computations and questions taken in March and May were converted to a 5 point scale. Column headings in this table cross-reference to the evaluation questions in Appendix D.

Table F19. Subject Matter Expert Scores Normalized by Number of Questions and Weight of Question

MONTH	SQUAD	BG	SUBJECT MATTER EXPERT SCORES NORMALIZED BY NUMBER OF QUESTIONS AND WEIGHT OF QUESTION																					SUM
			B1	B2	B3	B4	B5	D6	D7	D8	D9	A10	D11	E12	E13	F14	F15	F16	C17	F18	F19	F20	C21	
March	3	3RD	0.18	0.18	0.18	0.18	0.18	0.15	0.15	0.03	0.15	0.79	0.15	0.27	0.45	0.03	0.16	0.16	0.20	0.10	0.10	0.16	0.34	4.304
March	2	3RD	0.04	0.11	0.18	0.11	0.11	0.15	0.03	0.03	0.09	0.48	0.15	0.09	0.27	0.03	0.16	0.16	0.34	0.03	0.03	0.16	0.34	3.092
March	4	2ND	0.11	0.11	0.04	0.18	0.11	0.09	0.15	0.09	0.15	0.79	0.03	0.27	0.09	0.10	0.16	0.10	0.34	0.10	0.10	0.10	0.20	3.397
March	6	2ND	0.11	0.04	0.11	0.04	0.04	0.09	0.09	0.03	0.03	0.48	0.15	0.27	0.45	0.10	0.16	0.10	0.20	0.10	0.10	0.10	0.20	2.971
March	5	9TH	0.18	0.18	0.18	0.18	0.18	0.15	0.15	0.03	0.15	0.79	0.03	0.45	0.45	0.16	0.16	0.16	0.34	0.03	0.03	0.03	0.34	4.374
March	1	9TH	0.18	0.11	0.18	0.11	0.18	0.15	0.15	0.15	0.15	0.79	0.09	0.45	0.45	0.16	0.16	0.03	0.34	0.10	0.03	0.03	0.07	4.068
April	1	3RD	0.14	0.11	0.14	0.11	0.11	0.09	0.12	0.09	0.12	0.48	0.09	0.36	0.27	0.10	0.13	0.16	0.20	0.13	0.07	0.07	0.20	3.287
April	5	3RD	0.07	0.07	0.11	0.14	0.11	0.12	0.09	0.09	0.12	0.64	0.09	0.27	0.36	0.07	0.13	0.16	0.27	0.03	0.03	0.10	0.27	3.344
April	2	3RD	0.11	0.07	0.07	0.07	0.07	0.03	0.09	0.03	0.03	0.32	0.03	0.18	0.18	0.03	0.10	0.10	0.20	0.03	0.03	0.10	0.27	2.153
April	7	3RD	0.13	0.13	0.13	0.08	0.17	0.07	0.10	0.10	0.07	0.74	0.07	0.42	0.21	0.04	0.08	0.00	0.24	0.00	0.00	0.08	0.16	3.010
April	6	2ND	0.14	0.18	0.14	0.14	0.14	0.09	0.12	0.15	0.15	0.64	0.15	0.36	0.27	0.13	0.13	0.13	0.27	0.16	0.13	0.16	0.34	4.138
April	4	2ND	0.17	0.17	0.13	0.13	0.17	0.14	0.07	0.03	0.07	0.37	0.03	0.11	0.00	0.00	0.15	0.15	0.32	0.00	0.15	0.15	0.24	2.744
May	6	3RD	0.18	0.18	0.11	0.18	0.18	0.15	0.15	0.15	0.15	0.79	0.09	0.27	0.45	0.10	0.16	0.16	0.20	0.10	0.10	0.10	0.34	4.291
May	5	3RD	0.18	0.18	0.11	0.11	0.04	0.09	0.09	0.03	0.03	0.16	0.09	0.27	0.27	0.10	0.16	0.16	0.20	0.16	0.10	0.10	0.20	2.831
May	2	9TH	0.11	0.11	0.11	0.11	0.18	0.09	0.09	0.09	0.09	0.79	0.09	0.27	0.27	0.10	0.10	0.10	0.20	0.16	0.10	0.16	0.20	3.520
May	1	9TH	0.11	0.11	0.11	0.11	0.11	0.15	0.09	0.03	0.03	0.48	0.09	0.09	0.09	0.10	0.16	0.10	0.34	0.10	0.10	0.10	0.20	2.780
May	4	2ND	0.18	0.11	0.11	0.04	0.11	0.03	0.15	0.03	0.15	0.79	0.03	0.09	0.09	0.10	0.10	0.10	0.07	0.10	0.03	0.10	0.34	2.831
May	3	2ND	0.11	0.11	0.04	0.04	0.04	0.09	0.09	0.03	0.15	0.16	0.03	0.27	0.09	0.10	0.10	0.10	0.07	0.03	0.03	0.03	0.34	2.030
June	3	3RD	0.18	0.14	0.14	0.18	0.14	0.15	0.15	0.15	0.12	0.64	0.12	0.45	0.45	0.13	0.13	0.13	0.27	0.16	0.16	0.10	0.27	4.375
June	4	3RD	0.18	0.07	0.14	0.14	0.11	0.12	0.12	0.15	0.12	0.79	0.03	0.45	0.27	0.10	0.16	0.16	0.14	0.16	0.13	0.10	0.20	3.855
June	1	2ND	0.18	0.18	0.18	0.18	0.14	0.15	0.15	0.12	0.12	0.79	0.12	0.36	0.45	0.16	0.16	0.16	0.34	0.10	0.07	0.07	0.14	4.320
June	2	2ND	0.14	0.14	0.14	0.11	0.07	0.12	0.12	0.06	0.09	0.79	0.09	0.27	0.27	0.16	0.13	0.13	0.27	0.16	0.16	0.16	0.34	3.949
June	5	2ND	0.18	0.18	0.11	0.18	0.18	0.15	0.15	0.15	0.12	0.79	0.12	0.27	0.45	0.10	0.16	0.16	0.34	0.10	0.03	0.10	0.34	4.362
June	6	2ND	0.18	0.18	0.11	0.18	0.18	0.15	0.15	0.09	0.12	0.79	0.15	0.45	0.45	0.16	0.16	0.16	0.34	0.16	0.10	0.16	0.34	4.774

Note: Questions with a value of "0" (no response) were excluded from computations

Table F20. Summary of Evaluation

MONTH	SQUAD	BD	CAS?	EVALUATOR	Combined Measures Linear Average	Combined Measures AHP Derivation	Combined Measures Relative Difference
March	3	3RD	Yes	A	4.24	4.30	-1.6%
March	2	3RD	No	A	3.10	3.09	0.1%
March	4	2ND	Yes	B	3.29	3.40	-3.4%
March	6	2ND	No	B	2.81	2.97	-5.7%
March	5	9TH	Yes	C	4.05	4.37	-8.1%
March	1	9TH	No	C	3.86	4.07	-5.5%
April	1	3RD	YES	C	3.33	3.29	-1.4%
April	5	3RD	NO	C	3.14	3.34	6.0%
April	2	3RD	YES	B	2.05	2.15	4.9%
April	7	3RD	NO	B	2.61	3.01	13.2%
April	6	2ND	YES	D	4.24	4.14	-2.4%
April	4	2ND	NO	D	3.00	2.74	-9.3%
May	6	3RD	Yes	D	4.24	4.29	-1.2%
May	5	3RD	No	D	3.10	2.83	8.5%
May	2	9TH	Yes	E	3.38	3.52	-4.1%
May	1	9TH	No	E	2.90	2.78	4.3%
May	4	2ND	Yes	F	2.71	2.83	-4.3%
May	3	2ND	No	F	2.14	2.03	5.3%
June	3	3RD	YES	G	4.38	4.37	-0.1%
June	4	3RD	NO	G	3.76	3.86	2.4%
June	1	2ND	YES	H	4.24	4.32	1.9%
June	2	2ND	NO	H	3.86	3.95	2.3%
June	5	2ND	YES	I	4.24	4.36	2.9%
June	6	2ND	NO	I	4.67	4.77	2.3%
					Mean	3.53	

APPENDIX G

CORRELATION OF PARAMETERS AND REFINED SCORES

Table G1. Correlation Matrix 1 of 4

	CHARACTERISTICS CONSIDERED								
	Age	Gender	Rank	Years of Service	OTC Grad	Computer Skill Level	IPB Experience	NTC Rotations	Practical Exercise
Age	1.000	0.292	0.461	0.118	0.064	-0.002	0.125	-0.059	-0.068
Gender	0.292	1.000	-0.089	-0.331	-0.154	0.283	0.009	-0.055	0.212
Rank	0.461	-0.089	1.000	0.031	0.159	0.163	0.219	0.069	0.046
Years of Service	0.118	-0.331	0.031	1.000	-0.104	0.175	0.193	-0.028	0.031
OTC Grad	0.064	-0.154	0.159	-0.104	1.000	0.200	0.013	-0.053	-0.718
Computer Skill Level	-0.002	0.283	0.163	0.175	0.200	1.000	-0.101	0.055	0.016
IPB Experience	0.125	0.009	0.219	0.193	0.013	-0.101	1.000	0.051	0.079
NTC Rotations	-0.059	-0.055	0.069	-0.028	-0.053	0.055	0.051	1.000	0.168
Practical Exercise	-0.068	0.212	0.046	0.031	-0.718	0.016	0.079	0.168	1.000

Note: Values are correlation coefficients as a goodness of linear fit between given factors. Greyed cells indicates redundance.

Table G2. Correlation Matrix 2 of 4

	CHARACTERISTICS CONSIDERED									
	CAS or MMS	Members on MCOO	MCOO Time	Members on TPL	Whatifs for TPL	TPL Time	Members on SITMAP	SITMAP Time	Evaluator	Derived Score
Age 26-30 yrs: 2 31-35 yrs: 3	0.030	-0.047	0.168	-0.098	-0.046	0.352	0.221	0.260	-0.091	0.043
Gender	0.048	-0.334	-0.173	-0.419	-0.143	0.016	0.223	0.025	0.184	0.235
Rank	0.057	-0.095	0.079	0.027	0.407	0.136	-0.185	0.224	-0.459	0.189
Years of Service	0.031	0.162	-0.219	0.256	0.309	0.500	0.060	0.017	-0.185	-0.203
OTC Grad	0.220	0.282	0.373	0.102	-0.318	-0.174	-0.045	-0.093	-0.432	-0.437
Computer Skill Level	-0.078	-0.109	-0.237	0.156	0.026	0.161	0.080	-0.248	-0.378	0.064
IPB Experience	0.053	-0.130	0.203	-0.119	0.125	0.066	-0.257	0.335	0.162	-0.038
NTC Rotations	0.120	-0.132	0.051	0.443	0.294	0.132	0.068	0.008	-0.197	-0.145
Practical Exercise	0.046	-0.292	-0.618	0.094	0.500	0.229	0.256	0.105	0.477	0.323

Note: Values are correlation coefficients as a goodness of linear fit between given factors.

Table G3. Correlation Matrix 3 of 4

	CHARACTERISTICS CONSIDERED								
	Age	Gender	Rank	Years of Service	OTC Grad	Computer Skill Level	IPB Experience	NTC Rotations	Practical Exercise
CAS or MMS	0.030	0.048	0.057	0.031	0.220	-0.078	0.053	0.120	0.046
Members on MCOO (ea)	-0.047	-0.334	-0.095	0.162	0.282	-0.109	-0.130	-0.132	-0.292
MCOO Time	0.168	-0.173	0.079	-0.219	0.373	-0.237	0.203	0.051	-0.618
Members on TPL (ea)	-0.098	-0.419	0.027	0.256	0.102	0.156	-0.119	0.443	0.094
Whatifs for TPL	-0.046	-0.143	0.407	0.309	-0.318	0.026	0.125	0.294	0.500
TPL Time	0.352	0.016	0.136	0.500	-0.174	0.161	0.066	0.132	0.229
Members on SITMAP (ea)	0.221	0.223	-0.185	0.060	-0.045	0.080	-0.257	0.068	0.256
SITMAP Time	0.260	0.025	0.224	0.017	-0.093	-0.248	0.335	0.008	0.105
Evaluator	-0.091	0.184	-0.459	-0.185	-0.432	-0.378	0.162	-0.197	0.477
Derived Score	0.043	0.235	0.189	-0.203	-0.437	0.064	-0.038	-0.145	0.323

Note: Values are correlation coefficients as a goodness of linear fit between given factors.

Table G4. Correlation Matrix 4 of 4

	CHARACTERISTICS CONSIDERED									
	CAS or MMS	Members on MCOO	MCOO Time	Members on TPL	Whatifs for TPL	TPL Time	Members on SITMAP	SITMAP Time	Evaluator	Derived Score
CAS or MMS	1.000	0.023	0.024	0.177	0.035	0.023	0.049	0.305	-0.047	-0.450
Members on MCOO	0.023	1.000	0.234	0.258	0.120	0.055	0.392	0.167	-0.105	-0.304
MCOO Time	0.024	0.234	1.000	0.049	-0.411	0.079	-0.332	0.029	-0.167	-0.216
Members on TPL	0.177	0.258	0.049	1.000	0.385	0.459	0.127	-0.064	0.015	0.067
Whatifs for TPL	0.035	0.120	-0.411	0.385	1.000	0.346	-0.037	0.067	-0.047	0.163
TPL Time	0.023	0.055	0.079	0.459	0.346	1.000	0.146	-0.165	0.049	-0.078
Members on SITMAP	0.049	0.392	-0.332	0.127	-0.037	0.146	1.000	0.241	0.259	-0.024
SITMAP Time	0.305	0.167	0.029	-0.064	0.067	-0.165	0.241	1.000	0.148	0.006
Evaluator	-0.047	-0.105	-0.167	0.015	-0.047	0.049	0.259	0.148	1.000	0.332
Derived Score	-0.450	-0.304	-0.216	0.067	0.163	-0.078	-0.024	0.006	0.332	1.000

Note: Values are correlation coefficients as a goodness of linear fit between given factors. Greyed cells indicates redundance.

Table G5. Refined Complete Data Set After Normalization and Elimination of Evaluator Bias on Derived Score

Gender	Rank	Years of Service	OTC Grad	IPB Experience	NTC Rotations	Practical Exercise	CAS or MMS	Members on MCOO	MCOO Time	TPL Time	SITMAP Time	Derived Score	Norm Score
1.860	2.429	1.000	1.143	2.400	1.000	1	0	3	1	0.350	1.250	4.375	4.240
2.000	3.000	1.000	1.714	2.857	2.600	1	1	2	1	0.580	1.530	3.010	3.114
1.880	3.000	1.375	1.750	2.875	1.875	0	0	2	1	0.500	1.410	4.374	4.510
1.570	3.000	1.143	1.571	3.000	0.857	0	0	3	1	0.420	2.500	3.520	3.495
1.860	2.429	1.429	1.571	3.000	1.000	0	0	2	1	1.130	1.250	2.831	2.800
1.670	2.500	1.333	1.500	2.400	1.333	1	0	3	2	0.460	0.890	4.320	4.215
1.860	2.429	1.000	1.800	2.714	0.429	0	1	4	0	0.000	3.090	2.030	2.008
1.750	2.875	1.250	1.750	2.625	0.750	0	1	3	2	1.080	0.800	2.971	3.073
2.000	2.857	1.143	1.286	3.000	1.000	1	1	2	1	0.280	4.650	4.774	4.641
1.860	2.571	1.286	1.750	2.714	4.571	0	1	3	2	0.450	2.030	2.831	2.845
1.630	2.750	1.625	1.625	3.000	0.500	0	0	4	1	0.380	2.030	3.397	3.515
1.860	3.000	1.571	1.429	2.714	0.286	1	1	3	1	1.200	1.400	3.344	3.448
1.750	2.875	1.125	1.750	2.500	1.375	0	0	4	2	0.270	3.040	4.304	4.461
1.570	2.429	1.286	1.750	2.714	1.000	0	1	3	2	0.110	0.670	2.780	2.760
1.860	2.571	1.286	1.429	2.714	1.571	1	0	2	2	1.710	1.100	4.362	4.241
1.670	3.000	1.000	1.667	3.000	1.500	1	1	2	2	0.500	2.750	3.949	3.854
1.630	2.750	1.750	1.429	3.000	5.250	1	1	4	3	1.460	3.770	2.744	2.756
1.800	2.700	1.000	1.600	2.700	4.300	0	1	2	2	0.570	0.910	4.068	4.195
1.860	2.714	1.143	1.625	2.857	1.143	0	0	2	1	0.300	1.750	4.291	4.311
1.800	2.800	1.400	1.429	2.857	1.300	1	0	1	1	0.400	1.380	4.138	4.157
1.670	2.889	1.444	1.556	2.444	1.000	0	1	1	1	0.550	2.750	3.092	3.205
1.860	2.429	1.429	1.429	3.000	1.286	1	1	1	1	0.220	2.650	3.855	3.737
1.750	2.750	1.000	1.429	2.714	5.875	1	0	1	1	0.280	1.350	3.287	3.389

APPENDIX H

EXAMPLE MVLR AND ANN MODELS FOR CASE STUDY

MVLR Model for Training Set 1

Table H1. Cross Reference of MVLR Variables to Factors

Age	Gender	Rank	Years of Service	OTC Grad	Computer Skill Level	
X_1	X_2	X_3	X_4	X_5	X_6	
IPB Experience	NTC Rotations	Practical Exercise	CAS or MMS	Members on MCOO	MCOO Time	
X_7	X_8	X_9	X_{10}	X_{11}	X_{12}	
Members on TPL	Whatifs for TPL	TPL Time	Members on SITMAP	SITMAP Time	Evaluator	Score
X_{13}	X_{14}	X_{15}	X_{16}	X_{17}	X_{18}	f
Note: values in parenthesis below variable is the standard error						

$$
\begin{array}{llllllllllll}
f = & -4.70 + & -0.54X_1 & + & 4.14X_2 & + & 2.28X_3 & + & 0.79X_4 & + & 2.21X_5 \\
 & (4.80) & (0.98) & & (1.90) & & (1.64) & & (1.08) & & (1.26) \\
 & + & -0.48X_6 & + & -0.037X_7 & + & -0.20X_8 & + & -0.33X_9 & + & -0.72X_{10} \\
 & & (0.58) & & (0.89) & & (0.17) & & (0.60) & & (0.28) \\
 & + & -0.22X_{11} & + & 0.083X_{12} & + & 1.17X_{13} & + & -0.10X_{14} & + & -1.21X_{15} \\
 & & (0.35) & & (0.38) & & (0.34) & & (0.64) & & (0.55) \\
 & & 0.118X_{16} & + & -0.09X_{17} & + & 0.06X_{18} & & & & \\
 & & (0.38) & & (0.16) & & (0.13) & & & &
\end{array}
$$

ANN Model for Training Set 1

Table H2. Cross Reference of ANN Variables to Factors

Age	Gender	Rank	Years of Service	OTC Grad	Computer Skill Level	
Y[0]	Y[1]	Y[2]	Y[3]	Y[4]	Y[5]	
IPB Experience	NTC Rotations	Practical Exercise	CAS or MMS	Members on MCOO	MCOO Time	
Y[6]	Y[7]	Y[8]	Y[9]	Y[10]	Y[11]	
Members on TPL	Whatifs for TPL	TPL Time	Members on SITMAP	SITMAP Time	Evaluator	Score
Y[12]	Y[13]	Y[14]	Y[15]	Y[16]	Y[17]	YOUT[0]

```
/* Recall-Only Run-time for <untitled> */
/* Control Strategy is: <backprop> */

#if __STDC__
#define  ARGS(x) x
#else
#define  ARGS(x) ()
#endif /* __STDC__ */

/* --- External Routines --- */
extern double exp   ARGS((double));
/* *** LINK IN MATH LIBRARIES *** */

#if __STDC__
int NN_Recall( void *NetPtr, float Yin[18], float Yout[1] )
#else
int NN_Recall( NetPtr, Yin, Yout )
void  *NetPtr; /* Network Pointer (not used) */
float Yin[18], Yout[1]; /* Data */
#endif /* __STDC__ */
{
  float  Xout[30], Xsum[30]; /* work arrays */
  long   ICmpT; /* temp for comparisons */
```

```
/* *** WARNING: Code generated assuming Recall = 0 *** */

/* Read and scale input into network */
Xout[2] = Yin[0] * (1.4285713) + (-2.4285713);
Xout[3] = Yin[1] * (2.5000001) + (-4.0000003);
Xout[4] = Yin[2] * (1.6666669) + (-4.0000008);
Xout[5] = Yin[3] * (1.2500001) + (-1.2500001);
Xout[6] = Yin[4] * (1.4285716) + (-1.5714288);
Xout[7] = Yin[5] * (0.7692308) + (-2.1538462);
Xout[8] = Yin[6] * (1.6666669) + (-4.0000008);
Xout[9] = Yin[7] * (0.17857143) + (-0.05357143);
Xout[10] = Yin[8];
Xout[11] = Yin[9];
Xout[12] = Yin[10] * (0.33333333) + (-0.33333333);
Xout[13] = Yin[11] * (0.24213074) + (-0.18644067);
Xout[14] = Yin[12] * (0.33333333);
Xout[15] = Yin[13] * (0.33333333);
Xout[16] = Yin[14] * (0.58479531);
Xout[17] = Yin[15] * (0.33333333) + (-0.33333333);
Xout[18] = Yin[16] * (0.25125628) + (-0.16834171);
Xout[19] = Yin[17] * (0.125) + (-0.125);
LAB107:

/* Generating code for PE 0 in layer 3 */
Xsum[20] = (float)(0.057325989) + (float)(-0.05944255) * Xout[2] +
  (float)(-0.15127972) * Xout[3] + (float)(-0.15764004) * Xout[4] +
  (float)(-0.06546925) * Xout[5] + (float)(-0.18807848) * Xout[6] +
  (float)(-0.026239607) * Xout[7] + (float)(0.17631157) * Xout[8] +
  (float)(-0.024073586) * Xout[9] + (float)(0.084314264) * Xout[10] +
  (float)(-0.17475054) * Xout[11];
Xsum[20] += (float)(0.24513395) * Xout[12] +
  (float)(-0.07570865) * Xout[13] + (float)(-0.12373302) * Xout[14] +
  (float)(0.0059926445) * Xout[15] + (float)(0.077416979) * Xout[16] +
  (float)(0.19747657) * Xout[17] + (float)(-0.04104767) * Xout[18] +
  (float)(-0.22089522) * Xout[19];

/* Generating code for PE 1 in layer 3 */
Xsum[21] = (float)(0.07531447) + (float)(0.17684428) * Xout[2] +
  (float)(0.27229318) * Xout[3] + (float)(0.69451112) * Xout[4] +
  (float)(-0.50822979) * Xout[5] + (float)(-0.24830633) * Xout[6] +
  (float)(-0.04657457) * Xout[7] + (float)(-0.211982) * Xout[8] +
  (float)(0.11900024) * Xout[9] + (float)(0.19074948) * Xout[10] +
  (float)(-0.80056012) * Xout[11];
```

```
Xsum[21] += (float)(-0.41344786) * Xout[12] +
   (float)(-0.020667097) * Xout[13] + (float)(0.2284407) * Xout[14] +
   (float)(0.16431913) * Xout[15] + (float)(-0.29333016) * Xout[16] +
   (float)(0.011174412) * Xout[17] + (float)(0.23810488) * Xout[18] +
   (float)(0.43168399) * Xout[19];

/* Generating code for PE 2 in layer 3 */
Xsum[22] = (float)(-0.067809112) + (float)(-0.0053591505) * Xout[2] +
   (float)(0.12483589) * Xout[3] + (float)(0.28350881) * Xout[4] +
   (float)(-0.12579216) * Xout[5] + (float)(-0.0044687763) * Xout[6] +
   (float)(0.10161842) * Xout[7] + (float)(-0.14851837) * Xout[8] +
   (float)(-0.065371864) * Xout[9] + (float)(0.25380874) * Xout[10] +
   (float)(-0.57840532) * Xout[11];
Xsum[22] += (float)(-0.34029365) * Xout[12] +
   (float)(0.20504066) * Xout[13] + (float)(0.37463167) * Xout[14] +
   (float)(0.14478722) * Xout[15] + (float)(-0.30285701) * Xout[16] +
   (float)(-0.062908396) * Xout[17] + (float)(-0.058646932) * Xout[18] +
   (float)(0.15244944) * Xout[19];

/* Generating code for PE 3 in layer 3 */
Xsum[23] = (float)(0.021712821) + (float)(0.27444381) * Xout[2] +
   (float)(0.230147) * Xout[3] + (float)(0.21939589) * Xout[4] +
   (float)(-0.019622164) * Xout[5] + (float)(-0.29551747) * Xout[6] +
   (float)(-0.14718355) * Xout[7] + (float)(-0.2816847) * Xout[8] +
   (float)(-0.10503756) * Xout[9] + (float)(0.15717499) * Xout[10] +
   (float)(-0.53041983) * Xout[11];
Xsum[23] += (float)(-0.14455) * Xout[12] + (float)(-0.22182758) * Xout[13]
    + (float)(0.057297759) * Xout[14] + (float)(-0.052005358) * Xout[15]
    + (float)(-0.24818622) * Xout[16] + (float)(-0.26810849) * Xout[17] +
   (float)(0.20770641) * Xout[18] + (float)(0.45264933) * Xout[19];

/* Generating code for PE 4 in layer 3 */
Xsum[24] = (float)(0.18779555) + (float)(-0.20402764) * Xout[2] +
   (float)(-0.27189803) * Xout[3] + (float)(-0.77492291) * Xout[4] +
   (float)(0.26990998) * Xout[5] + (float)(0.57972693) * Xout[6] +
   (float)(-0.22543038) * Xout[7] + (float)(0.39375758) * Xout[8] +
   (float)(0.56010318) * Xout[9] + (float)(-0.47224388) * Xout[10] +
   (float)(1.1507142) * Xout[11];
Xsum[24] += (float)(0.49003279) * Xout[12] +
   (float)(-0.066723406) * Xout[13] + (float)(-0.40883002) * Xout[14] +
   (float)(-0.24952975) * Xout[15] + (float)(0.40514562) * Xout[16] +
   (float)(0.071232222) * Xout[17] + (float)(-0.39010376) * Xout[18] +
   (float)(-0.89807367) * Xout[19];
```

```
/* Generating code for PE 5 in layer 3 */
Xsum[25] = (float)(-0.089726642) + (float)(-0.034636367) * Xout[2] +
   (float)(0.22548361) * Xout[3] + (float)(0.20260413) * Xout[4] +
   (float)(-0.0032939625) * Xout[5] + (float)(-0.26527044) * Xout[6] +
   (float)(0.19396503) * Xout[7] + (float)(-0.37838534) * Xout[8] +
   (float)(-0.019882374) * Xout[9] + (float)(0.020910552) * Xout[10] +
   (float)(-0.68942577) * Xout[11];
Xsum[25] += (float)(-0.087543815) * Xout[12] +
   (float)(0.24406624) * Xout[13] + (float)(0.17339514) * Xout[14] +
   (float)(-0.062970318) * Xout[15] + (float)(-0.23300278) * Xout[16] +
   (float)(-0.11845291) * Xout[17] + (float)(-0.022264367) * Xout[18] +
   (float)(0.52231717) * Xout[19];

/* Generating code for PE 6 in layer 3 */
Xsum[26] = (float)(-0.12712111) + (float)(0.25038403) * Xout[2] +
   (float)(-0.10328344) * Xout[3] + (float)(0.1291703) * Xout[4] +
   (float)(-0.015075168) * Xout[5] + (float)(-0.053444475) * Xout[6] +
   (float)(-0.047624577) * Xout[7] + (float)(-0.072006024) * Xout[8] +
   (float)(-0.19506989) * Xout[9] + (float)(0.29203054) * Xout[10] +
   (float)(-0.41286999) * Xout[11];
Xsum[26] += (float)(-0.31649864) * Xout[12] +
   (float)(-0.17260231) * Xout[13] + (float)(0.29197583) * Xout[14] +
   (float)(0.27842873) * Xout[15] + (float)(-0.040052824) * Xout[16] +
   (float)(-0.20874429) * Xout[17] + (float)(0.28167006) * Xout[18] +
   (float)(0.26585534) * Xout[19];

/* Generating code for PE 7 in layer 3 */
Xsum[27] = (float)(0.12371042) + (float)(-0.29842502) * Xout[2] +
   (float)(-0.038114056) * Xout[3] + (float)(-0.43036178) * Xout[4] +
   (float)(0.21957934) * Xout[5] + (float)(0.2878828) * Xout[6] +
   (float)(-0.18084572) * Xout[7] + (float)(0.16747692) * Xout[8] +
   (float)(0.28615183) * Xout[9] + (float)(-0.40030488) * Xout[10] +
   (float)(0.69505513) * Xout[11];
Xsum[27] += (float)(0.28732643) * Xout[12] +
   (float)(0.043565128) * Xout[13] + (float)(-0.58580804) * Xout[14] +
   (float)(-0.23211473) * Xout[15] + (float)(0.38700625) * Xout[16] +
   (float)(0.27268511) * Xout[17] + (float)(-0.27426323) * Xout[18] +
   (float)(-0.43431687) * Xout[19];

/* Generating code for PE 8 in layer 3 */
Xsum[28] = (float)(0.026486266) + (float)(0.039567832) * Xout[2] +
   (float)(0.27878395) * Xout[3] + (float)(0.3890374) * Xout[4] +
   (float)(-0.37737322) * Xout[5] + (float)(-0.34974959) * Xout[6] +
```

```
    (float)(0.011750641) * Xout[7] + (float)(-0.33800283) * Xout[8] +
    (float)(-0.17442487) * Xout[9] + (float)(0.205642) * Xout[10] +
    (float)(-0.48824683) * Xout[11];
Xsum[28] += (float)(-0.4204118) * Xout[12] + (float)(0.17671473) * Xout[13]
    + (float)(0.37403974) * Xout[14] + (float)(-0.052312553) * Xout[15] +
    (float)(-0.053863544) * Xout[16] + (float)(-0.058087837) * Xout[17] +
    (float)(0.19172211) * Xout[18] + (float)(0.50510061) * Xout[19];

/* Generating code for PE 0 in layer 3 */
Xout[20] = 1.0 / (1.0 + exp( -Xsum[20] ));

/* Generating code for PE 1 in layer 3 */
Xout[21] = 1.0 / (1.0 + exp( -Xsum[21] ));

/* Generating code for PE 2 in layer 3 */
Xout[22] = 1.0 / (1.0 + exp( -Xsum[22] ));

/* Generating code for PE 3 in layer 3 */
Xout[23] = 1.0 / (1.0 + exp( -Xsum[23] ));

/* Generating code for PE 4 in layer 3 */
Xout[24] = 1.0 / (1.0 + exp( -Xsum[24] ));

/* Generating code for PE 5 in layer 3 */
Xout[25] = 1.0 / (1.0 + exp( -Xsum[25] ));

/* Generating code for PE 6 in layer 3 */
Xout[26] = 1.0 / (1.0 + exp( -Xsum[26] ));

/* Generating code for PE 7 in layer 3 */
Xout[27] = 1.0 / (1.0 + exp( -Xsum[27] ));

/* Generating code for PE 8 in layer 3 */
Xout[28] = 1.0 / (1.0 + exp( -Xsum[28] ));

/* Generating code for PE 0 in layer 4 */
Xsum[29]  = (float)(-0.063403845) + (float)(-0.13670824) * Xout[20] +
    (float)(0.82744712) * Xout[21] + (float)(0.49447986) * Xout[22] +
    (float)(0.53418481) * Xout[23] + (float)(-1.4181669) * Xout[24] +
    (float)(0.55539185) * Xout[25] + (float)(0.42739907) * Xout[26] +
    (float)(-0.92722863) * Xout[27] + (float)(0.69787592) * Xout[28];
Xout[29] = 1.0 / (1.0 + exp( -Xsum[29] ));
```

```
   /* De-scale and write output from network */
   Yout[0] = Xout[29] * (2.74) + (2.03);

   /* Generating code for PE 0 in layer 4 */
   return( 0 );
}
```