University of Alberta

A Fuzzy Expert System for Design Performance Prediction and Evaluation

By

Zhuo Sun

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Master of Science

In

Construction Engineering and Management

Department of Civil and Environmental Engineering

Edmonton, Alberta

Spring 2000

0-612-60182-X

University of Alberta

Library Release Form

Name of Author: Zhuo Sun

Title of Thesis: A Fuzzy Expert System for Design Performance Prediction and Evaluation

Degree: Master of Science

Year this Degree Granted: 2000

Sun Zhuo

Zhuo Sun
7-1-1, No. Yi 1
Huangsi Street
DongCheng District
Beijing, 100011
P.R. China

Dated: Jan 26/ 2000

University of Alberta

Faculty of Graduate Studies and Research

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled **A Fuzzy Expert System for Design Performance Prediction and Evaluation** submitted by **Zhuo Sun** in partial fulfillment of the requirement for the degree of **Master of Science** in **Construction Engineering and Management**.

A Robinson

Dr. A. Robinson

Dr. W. Pedrycz

Dr. S. T. Ariaratnam

Date: Jan 24/2000

ABSTRACT

The objective of this research is to develop a model for use in evaluating and predicting a design firm's performance using fuzzy set theory and fuzzy logic techniques. The final fuzzy expert system consists of a series of membership functions, a fuzzy rulebase, a fuzzy inference system and a defuzzification module.

Fuzzy membership functions were developed based on actual survey data using a proposed new technique, which made use of the limited data set. The generated membership functions were tested twice to prove the feasibility and accuracy of this new technique. Before generating If-Then rules, correlation analysis was used to reduce the number of factors in the model. A new technique was developed and used to generate If-Then rules, based on the frequency of actual data and the correlation analysis. The model was tested with the actual data collected from the survey. Sensitivity analysis was conducted to explore the effect of modifying components of the fuzzy expert system.

ACKNOWLEDGEMENTS

I would like to gratefully acknowledge my supervisor, Dr. A. Robinson for her thorough and patient guidance, valuable advice and consistent support throughout my study in University of Alberta, especially during the research period. I have benefited a lot from her intelligence and hard work.

I also wish to thank Dr. W. Pedrycz for his helpful discussions and useful advice. He has provided me with valuable suggestions during the research. I am grateful to K. Gue, who worked in the summer to conduct the survey and compiled the data for this research. Dr. Allouche also provided the author with some good ideas. The author would like to thank him as well.

I would sincerely like to thank all of the participating companies in the survey. Their patience and knowledge made this research possible. Amongst the participating companies are: Chatwin Engineering Ltd., Colt Engineering, GPEC Consulting Ltd., H.A. Simons Ltd., Hydroconsult En3 Services Ltd., Infrastructure Systems Ltd., Maxim Engineering Inc, Reid Crowther & Partners Ltd., Reid Crowther & Partners Ltd., Shaflik Engineering Ltd., SNC-Lavalin Inc. (Edmonton), SNC-Lavalin Inc., (Calgary), Techna-West Engineering Ltd., UMA Engineering Ltd., Westmar Consultants Inc. (Edmonton), Westmar Consultants Inc.(North Vancouver).

Finally, I would like to thank all of my good friends in Edmonton who made my life in Canada very colorful and meaningful. Special thanks to my dearest parents far away in China, who always encourage and support me.

Table of Contents

List of Tables

List of Figures

CHAPTER 1 INTRODUCTION

1.1 Background and Problem Statement

In the construction industry, productivity is receiving more and more attention. It is something that everyone wants to improve, but to understand productivity and know how to improve it is, in practice, a complex subject. This is particularly true for an engineering organization. Historically, researchers have focused their attention more on addressing the management of the construction phase than on the management of the engineering phase of a project. "The reason perhaps is that the cost associated with the engineering phase is only 3-10% of the total project cost" (Eldin, 1991). However, this relatively small percentage of the total cost can amount to a large sum of money on a very large project. Moreover, most construction costs are fixed by design. Subsequently, the design greatly affects construction costs. In the lifecycle of a project, making changes in the early design phase requires the least effort, whereas if they are made in the construction phase, the resources needed are often huge. The management of the engineering phase should therefore be given more time and attention in order to reduce overall construction costs.

According to the Construction Industry Institute (CII, 1986), design, which occurs in the early stages of a project, is a collection of activities including those associated with creative thinking, sophisticated engineering calculations, and the translation of ideas into drawings, specifications, and procurement of major items for construction. These activities involve many contributors including owners, the various designer groups, vendors, and perhaps construction representatives.

When evaluating design productivity, the author would like to use terminology other than productivity, and so the term "design performance" will be used in this thesis. According to studies carried out by CII (1986, 1987), the first term (design productivity) refers to the design process itself. Even though man-hours, computer time, and design costs and design schedules are all proper parameters for measuring design productivity, the measurement parameters for design performance involves total project costs, total project schedules, design quality, constructability and plant start-up, etc. Since we are interested not only in the design process itself but also want to see the overall effect of the design on the project, design performance is a good term to use in this case. However, so far, "there is no reliable, cost-effective method for measuring design performance in design organizations" (Thomas, et. al., 1999). How should we evaluate the performance of an architectural and engineering organization? This is still a problem upon which little research has been done.

Imprecision and vagueness in engineering design are natural phenomena that arise in any human conceptualization activity. The lack of quantitative information is often cited as a serious deficiency. The difficulty in measuring the performance of an engineering organization always lies in the fact that it is difficult to find suitable variables to assess and quantify the contributions of both the input and output in such an organization. Unlike manufacturing organizations, the input and output of a design organization is information. Traditional methods of measuring productivity, such as simplistic measurements of cost per drawing or work hours per drawing, have obvious limitations because of variations in drawing size and content. Subsequently, a new method needs to be found to evaluate the performance of the design organization.

So far, few people have tried to apply fuzzy set theory, which deals best with linguistic problems with fewer data, to this area. This thesis develops a new method of evaluating the performance of a design organization based on the techniques of fuzzy set theory. This method could prove to be a good technique for this research area.

1.2 Purpose of the Study and Expected Contributions

The main objective of this research is to develop a model for use in evaluating and predicting a design firm's performance using fuzzy set theory and fuzzy logic techniques. In addition, this research also involves the following sub-objectives:

- First, to compile a thorough list of factors that could potentially affect design performance and a list of factors to measure design performance.
- Second, to develop and test a new technique for generating membership functions used in fuzzy set theory based on objective data.
- Third, to develop a fuzzy expert model for use in predicting design performance. Once this has been accomplished, we can use pre-construction information to predict the design performance – that is, the timeliness, cost performance, and accuracy of the design phase of a project.
- Fourth, to evaluate the design effectiveness of a completed project using the fuzzy model, and to translate numeric information into linguistic descriptors that are easy to understand.

The research is expected to make a contribution because of its attempts to:

- Use fuzzy set theory for design performance prediction and evaluation, which will cast a new light on this area of research.
- Experiment with a new technique for generating membership functions based on limited data, which may yield an accurate method of generating and evaluating membership functions for different contexts.
- Test and experiment with the use of statistical analysis as an intermediate step for generating expert rules.
- Experiment with rule-based fuzzy expert systems to provide a reasoning framework for design performance prediction and evaluation.

1.3 Research Methodology

The objective of this research is to measure the performance of the design phase of a construction project. The first step involved a thorough literature review to identify the criteria by which to evaluate and measure design performance.

Based on these evaluation criteria, the author identified the factors that influence design (i.e. input factors) and the factors that measure design performance (i.e. output factors). Because most of these factors are subjective in nature, each factor was divided into several sub-factors, each of which can be expressed more objectively. This structured list of factors forms the basis of the model for predicting and evaluating design performance.

A four-month survey among industrial contractors in Alberta and British Columbia was conducted to obtain case studies of actual design projects. From this survey, both objective numbers and subjective linguistic variables were obtained to describe each factor and its sub-factors. These data were used to develop and test the fuzzy membership functions and the expert rules.

The data collected from the survey were plotted to obtain the frequency distribution for the values of each factor. Because of the limited amount of data received, the existing methods for generating membership functions could not be applied. Furthermore, because of the complexity of the model, existing methods of generating fuzzy expert rules were infeasible. Consequently, a new technique for generating membership functions and fuzzy expert rules were developed. To verify the method of generating membership functions, development and testing of membership functions were conducted twice.

Generating membership functions was the first step in the development of a fuzzy expert system. Because of the complexity of the model, a correlation analysis of the data was used to simplify the model, before generating expert rules.

A complete and consistent rulebase was developed using the data from the survey. The accuracy of the expert rulebase was also tested.

Finally, a sensitivity analysis of the fuzzy expert system was conducted to test its sensitivity to fuzzy operators, implication methods, aggregation methods, and defuzzification methods. This sensitivity analysis also helped to improve the model's accuracy.

1.4 Thesis Organization

Chapter 2 contains a literature review. It examines the application of fuzzy set theory in the construction industry. It also covers the previous research on design performance evaluation.

Chapter 3 describes the model of factors that influence and measure design performance. A list of factors was developed that incorporates all aspects of the design process.

Chapter 4 describes the survey used for data collection. It includes a description of how the survey was developed, how the data were gathered, and how the data were processed for use in the next step.

Chapter 5 describes how the membership functions were developed using the new technique. The validation of this method is also presented.

Chapter 6 describes how correlation analysis was used to eliminate factors and simplify the model.

Chapter 7 describes the procedures used for generating fuzzy if-then rules and how these rules were tested with the survey data. This chapter also presents the results of the sensitivity analysis.

Chapter 8 presents the conclusions of this research and recommendations for future development.

CHAPTER 2 LITERATURE REVIEW

2.1 Introduction

In recent years, increasing attention has been given to studying productivity and productivity improvement in industrialized nations. In the construction area, Herbsman et al. (1990), for example, identified certain influential factors affecting construction productivity. They proposed a statistical model to relate these factors to construction productivity. In this model, previous data accumulated from past projects have to be collected to fit this statistical model. The purpose of such a model is to estimate productivity rates with a high degree of accuracy for future projects.

Another example is Portas and AbouRizk (1997), who developed a neural network model for estimating construction productivity. The method they proposed can be used to estimate construction productivity for certain construction tasks. It is a good method for utilizing artificial intelligence techniques to solve a real-world problem. However, this method requires highly precise and extensive data.

The factors that affect construction productivity and the research on construction productivity have been studied extensively. However, the research on design productivity has not received enough attention. According to CII (1986), "Measurement of design productivity is perhaps more difficult than measuring productivity in the construction phase". This chapter describes previous studies on design productivity. It also discusses previous research on the techniques used in this thesis to model design performance.

2.2 Previous Studies of Design Performance

Design is such a subjective process that it is influenced by many factors. Until now, the most thorough research was conducted by the Construction Industry Institute (CII), the Bureau of Engineering Research, at the University of Texas at Austin (1986, 1987). In evaluating design effectiveness, the CII Design Task Force did its research from both the input perspective and from the point of view of the output or products of the design process. First, they identified a list of input and output factors. Then the Objectives Matrix method was employed to deal with the subjective factors inherent in the model.

The Objectives Matrix, developed by Riggs and others (CII, 1986), can be used for productivity evaluation and also for the measurement of design effectiveness. The four major components of this method are the criteria, the weights, the performance scale, and the performance index. This method tried to translate subjective judgement into countable numbers, illustrating that subjective judgement in such a model is unavoidable.

The CII research failed to combine these input and output variables, as each was studied separately. The difficulty CII research encountered was that no matter what the input or output variables are, they are factors that affect the design productivity at different stages. Subsequently, to best evaluate design effectiveness, the researchers can not treat these factors separately; these factors have to be considered as a whole. Moreover, as CII noted, most evaluation factors are subjective. How to quantify them is a major concern when they are to be incorporated in a model.

Thomas et al. (1999), McGeorge (1989), and Armentrout (1985) have also done research in the area of design productivity. However, none of this research thoroughly examined the factors that affect design productivity. Armentrout's work focused more on design performance measurement, a qualitative management method. No factors were identified and no quantitative analysis was provided. McGeorge, similarly, did not relate certain factors to design productivity. He discussed design problems in terms of the whole project lifecycle, proposing a management procedure to improve design productivity. Thomas and others did not get beyond the limitations of the traditional productivity formula. They still used the ratio of input and output to evaluate design productivity. To meet the requirements of this formula, they tried to convert all the output to one unified unit so that they could apply the time spent on the project as input over the unified output to get a ratio that could serve as a criteria to evaluate design productivity. It is obvious that this method has several shortcomings. First, for a large project, how to break down all the tasks? After all, converting is difficult. Second, how to determine the conversion factor? This is a serious issue for diverse projects.

No research has proposed a suitable model for dealing with the subjective factors encountered in design effectiveness evaluation.

2.3 Application of Fuzzy Set Theory in Construction

A fuzzy set approach, pioneered by Zadeh (1965), was developed specifically to deal with uncertainties that are not statistical in nature. The fuzzy set approach has been widely applied to represent the uncertainties of real-life situations. The past few years have witnessed a rapid growth in the number and variety of applications of fuzzy set theory. In the construction field, the applications in various disciplines can be traced back to the early 1980s in a number of countries. For example, fuzzy set theory has been used as a general assessment technique that has been applied in risk assessment during construction (Ross et al., 1994; Duckstein et al., 1990; Thiel et al., 1985; Dahab et al., 1998; Lee et al., 1995). Beer et al. (1998) have also used this method for safety assessment.

A fuzzy algorithm has been used to estimate unmeasured variables, such as influent substrate and recyclable biomass concentrations in the operation and control of a high-purity oxygen-activated sludge process (Yin et al., 1999). Another important application of fuzzy set theory in construction is in classification, for which it is extremely suited. Zhang et al. have applied this approach to CPT soil classification (1999). Fuzzy pattern recognition models have also provided a new methodology for diagnosing engineering problems (Chao et al., 1998).

Moreover, fuzzy set theory and fuzzy logic have been used in virtually all types of analysis problems encountered in civil engineering work. Ayyub et al. (1984) used fuzzy set and system theory to estimate the duration of a project activity, when a probabilistic method could not be used for linguistic expressions; Chao and Skibniewski (1998) used fuzzy logic to evaluate alternative construction technologies; Fayek (1998) applied fuzzy set theory to the markup size selection; and, Russel et al. (1994) utilized fuzzy logic for developing an automated corrective action selection system.

2.4 Methods of Generating Membership Functions

According to the literature, the author summarized the most commonly used methods for constructing membership functions as follows:

1) Horizontal method of membership estimation

According to W. Pedrycz et.al. (1998), the main idea behind the horizontal method is to gather enough responses regarding the membership values of each element in the universe of discourse for a certain concept. The steps when applying this method are as follows:

a. First, a group of testees must be available. The number in this group should meet the requirements of statistical theory, say, at least 30.

b. According to this group's response, each element in the set will receive a response related to a linguistic descriptor. For example, in the data set, the number 1 will be designated a response related to the linguistic descriptor "small" for a certain frequency. So the estimated grade of 1 belonging to "small" is the frequency over the total number of responses. The estimated value of the membership function at X_I is simply the ratio of the number of positive replies $P(X_I)$ to the total number N of responses, $A(X_I) = P(X_I) / N$, $I=1, 2, \ldots, n$. Once the membership value for each element is determined, the plot of the membership functions for all the linguistic variables can be done.

c. This method is very simple and easy to understand. However, when applying this method to our model, one of the limitations is the determination of the several selected elements. This method requires the universe of discourse to be highly unionized. If the units of the numbers from the responses differ greatly, then this method will have some difficulty in handling this.

d. Another difficulty is the number of samples, since in the survey, there is no way to ask the testees what they think of each element in the set. So if the number of replies is not large enough, enough information needed to build the three membership functions for each factor can not be collected. Moreover, since the testees were not asked about what they thought about each element in the set, some elements may get zero responses. In that case, it would be difficult to build membership functions properly. Furthermore, the survey conducted to collect data yielded a sample size of 18, not sufficiently large enough to apply this method.

e. However, due to its simplicity, it is still a powerful method for generating membership functions.

2) Vertical method of membership estimation

The procedure behind the vertical method is quite similar to that of the horizontal method. However, it deals with a range instead of just a single element in the universe of discourse, since the membership function in this method is decided by a group of α-cuts. The steps involved in using this vertical method, according to Pedrycz and Gomide (1998), are as follows:

a. First, a group of α-cut levels is selected.

b. Second, the testees are asked to identify the corresponding subset of X whose elements belong to A to a degree not less than α.

c. Then the membership function is constructed piling up these successive α -cuts.

d. The main advantage of horizontal and vertical methods lies in their conceptual clarity and simplicity.

e. However, as we can see, the results of these two methods are scattered and discrete. This is the major problem with these two methods. The experts are always exposed to a single element either in the universe of discourse or the membership scale, making the experiment quite discrete.

f. In our model, the major difficulty in utilizing the vertical method results from the complexity of the problem to be solved. The range of each variable in the model may differ greatly, depending on the different circumstances. Unless the range of the X-axis is fixed, this method cannot be properly used.

3) Membership estimation as a problem of parametric optimization

In this method, a rough membership function is already given. However, the parameters of the function still need to be tuned to fit a standard curve. The parameters are estimated as follows (according to Pedrycz and Gomide, 1998):

a. First, let us assume a parameterized membership function A(x;p) where $x \in X$, and p is a vector of its parameters in the appropriate parameter space P and where element x_k and its membership value $M(x_k)$ are also given.

b. To estimate the parameter p, the mean squared error method can be utilized. Therefore, the problem changes to:

$$\min_{p \in P} \sum_{k=1}^{N} [M(x_k) - A(x_k; p)]^2 \qquad (2\text{-}1)$$

which can be solved by an appropriate iterative, gradient-based algorithm.

Another method that can be used for parameter estimation is that of using expert knowledge, which was first proposed by Bobrowicz et al. (1990). In this method, the parameters of the membership function are determined by two sources of knowledge: one is the knowledge of semantic links joining the three predicates (such as small, average, large) the researchers want to represent in the same universe of discourse, and the other one is the expert knowledge (heuristics) about the meaning that the expert wants to give to each predicate.

c. The advantage of such a method is that it gives a smooth function rather than a rough shape.

d. The difficulty in using this method in our model lies in the fact that it is hard to initially determine the membership value of each element in the universe of discourse. If the shape of the membership function can not be predefined, then Bobrowicz's method can not be applied.

4) Membership estimation via fuzzy clustering

According to Pedrycz and Gomide (1998), "Fuzzy clustering forms another important class of membership estimation methods and is algorithmic in nature". The difference between fuzzy clustering and conventional clustering lies in the fact that fuzzy clustering can have an object belonging to one cluster to a certain degree, whereas in the conventional clustering method, an object can either belong to one cluster or not. Fuzzy clustering is more suitable for real-world problem solving.

Generally speaking, the fuzzy clustering technique simply partitions a set of numerical data into several clusters. The degree to which an object belongs to one cluster can be interpreted as the membership value. Various algorithms for fuzzy clustering have been developed by a large number of researchers. Some of the most widely used are the fuzzy C-means algorithm (Bedzek, 1981) and the fuzzy isodata (Dunn, 1974). Recently, the possibilistic algorithm (Raghu, 1994), and the Gaussian clustering algorithm (Li et al., 1999) have been discussed by several researchers.

5) Constructing membership functions using statistical data

Civanlar and Trussell (1986) proposed a method to derive membership functions from the probability density function (pdf). According to their research, "for any pdf, the method is capable of generating membership functions in accordance with the possibility -probability consistency principle". The foundation of such a method is that there is a relationship between statistically-based sets and the fuzzy sets. To generate membership functions from pdf is the most natural idea. Heuristically, an element, that is the one most likely to happen in a statistical set, should have a high membership value in the fuzzy set. Based on this theorem, membership functions can be derived from probability density functions.

This method gives us a hint on how to generate membership functions. However, the author lacked enough data to plot the pdf. Simply for this reason, this method can not be used.

6) Constructing membership functions using interpolation and measurement theory

According to Chen and Otto, "Measurement theory provides a mathematically axiomatic method to construct membership values" (1995). In this method, the membership value of the real-valued variables can be initially decided by measurement theory. The user of this method has to identify the bound value of the points in X, that is, the least and the greatest value of the elements in X. Then for each of the rest of the elements that belong to X, the user must answer questions such as the following:

"On a scale of zero to one, what is your belief μ that you are indifferent between:

1) receiving the objective performance provided by x_i; or

2) receiving the objective performance provided by x_{best} with certainty μ and receiving the objective performance provided by x_{worst} with certainty $(1-\mu)$?"

The method simply provides a comparison of the rest elements in X (except one largest value and one least value) with the two boundaries. So only the finite sets of the whole fuzzy set can be obtained. Then this set has to be interpolated. In such a case, constrained interpolation theory has to be used. The method proposed by Chen and Otto overcomes two serious problems encountered by traditional interpolation theory: the monotony and convexity of the membership function and a boundary restricted to [0,1].

In this method, the Measurement theory is used to build the framework of the membership functions based on the subjective preferences. Then the revised interpolation theory is used to determine the remaining membership values so that the finite set can be extended to a smooth infinite membership function. The advantage of this method is that it does not need a lot of data (only the two boundary values are the most important information). However, the big disadvantage of this method is the value in x-axis has to have a limit, or else, this method cannot be applied. But in many cases, the limit of the problem cannot be decided subjectively. For example, one factor is the number of employees in the design firm. How to set a limit for this problem to get a membership value of 1 is hard to implement in reality.

7) Induction of fuzzy rules and membership functions from training examples

Hong and Lee (1996) proposed a method that can automatically derive membership functions and fuzzy if-then rules from a set of training data. This method is easy to carry implement, and it can significantly reduce the time and effort needed for developing a fuzzy expert system. However, this method runs into a dilemma when too many input variables exist, because the decision table will become too complex to solve.

The algorithm for this method is as follows:

a) A set of training data should be available first. Also, both the input and the desired output should be known.

b) For a m-dimensional input space, the ith training example can then be described as $(X_{i1}, X_{i2}, X_{i3}, \dots X_{im}, Y_i)$, where X_{ir} $(1 \le r \le m)$ is the rth attribute value of the ith training example and Y_I is the output value of the ith training example.

c) Cluster and fuzzify the output data. In this step, the membership function for the output is derived.

c1) Sort the output values of the training instances in ascending order. The modified order after sorting is then: $Y_1^{'}, Y_2^{'} \cdots Y_n^{'}$, where $Y_i^{'} \le Yi+_1^{'}$ (for I=1,..., n-1).

c2) Find the difference between adjacent data. The difference between adjacent data provides information about the similarity between them.

c3) Find the value of the similarity between adjacent data. In order to obtain the value of the similarity between adjacent data, we convert each distance to a real number, S_I, between 0 and 1, according to the following formula:

$$s_i = \begin{Bmatrix} 1 - \dfrac{diff_i}{C^* \sigma_s} fordiff_i \le C^* \sigma_s, \\ 0, otherwise \end{Bmatrix}, \tag{2-2}$$

where σ_s is the standard deviation of difference and C is a control parameter deciding the shape of the membership functions of similarity. A larger C causes a greater similarity.

c4) Cluster the training instances according to similarity. Here, the α-cut is used as a threshold for two adjacent bits of data to be thought of as belonging to the same class. A larger α will have a smaller number of groups. If $S_I < \alpha$, then divide the two adjacent bits of data into different groups. Otherwise, put them into the same group. After the above operation, we can obtain the result formed as ($Y_i^{'}$, R_j), meaning that the ith output data will be clustered into the R_j, , where R_j means the jth produced fuzzy region.

c5) Determine membership functions of the output space. For simplicity, only triangular membership functions are used here for each linguistic variable. A triangular membership function can be defined by a triad (a, b, c). Here, b is the central-vertex-point: $a = b_j - \dfrac{b_j - y_i^{'}}{1 - \mu_j(y_i^{'})}$. (2-3)

$$c = b_j + \frac{y_i^{'} - b_j}{1 - \mu_j(y_k^{'})}. \qquad (2\text{-}4)$$

c6) Find the membership value belonging to the derived group for each instance. From the above membership functions, we can obtain the fuzzy value of each output datum formed as ($Y_i^{'}$, R_j, μ_{ij}), referred to as the ith output data, which has the fuzzy value μ_{ij} to the cluster R_j. Each training instance is then transformed as

(x1, x2, …, xm; (R_1, μ1), (R_2, μ_2), …, (R_k, μ_k)), simply using (x11, x12) to take the place of y1.

At this point, the membership functions for the output have been decided.

d) Construct initial membership functions for input attributes.

We assign each input attribute an initial membership function, which is assumed to be a triangle (a, b, c) with b-a=c-b=the smallest predefined unit. For example, if three values of an attribute are 10, 15 and 20, then the smallest unit is chosen to be 5.

e) Construct the initial decision table. In this step, we build a multi-dimensional decision table (each dimension represents a corresponding attribute) according to the initial membership functions.

f) In order to simplify the initial decision table, take certain merging operations to eliminate redundant and unnecessary cells.

g) Rebuild the membership functions. When we execute the above operations, corresponding membership functions for the dimension are derived at the same time. The parameters of the new membership functions are calculated in accordance with the different operations of the previous step.

h) This method can also derive the if-then rules automatically. After the previous step, we can derive the decision rules from the decision table. Each cell in the decision table is used to derive a rule:

If input 1=d_1, input 2=d_2, ..., and input m=d_m, then output =R_i.

i) Once we get the membership functions from the above steps, the fuzzy inference system is also built.

j) The disadvantage of this method is that the decision table and the initial membership functions are complex if there are many attributes (input variables) or if the predefined unit is small. In our model, the number of some factors' input variables are much greater than 4, making the use of this method very difficult.

8) Finding relevant attributes and membership functions

Hong and Chen (1997) developed an improved version of Hong and Lee's method, which overcomes the difficulty of the previous method and makes it possible to address more complex situations.

This improved method first selects relevant attributes by eliminating unimportant input variables to simplify the inputs; then it can build appropriate initial membership functions. These attributes and membership functions are then used in a decision table to derive final fuzzy if-then rules and membership functions.

The learning algorithm for this method includes three main parts:

I. Finding the relevant attributes

The fact that this method can find the relevant attributes is the major improvement of this method over Hong and Lee's method. In this step, we are going to decide which attributes are suitable for distinguishing the classes. In a word, this is a procedure that simplifies the input variables.

1a) Sort each attribute value of each class appearing in training instances in ascending order.

1b) For each attribute value A_{ij}, count how many instances of A_{ij} belong to the same class.

1c) Sum how many instances of A_I with attribute values belong to only one class. An attribute value that belongs to only one class is certainly useful in distinguishing among classes. An attribute with many such attribute values is then a relevant attribute. Sum the total number of instances whose attribute values of A_I belong to only one class. Let t_I represent this sum.

1d) Calculate the fitness degree of each attribute. The fitness degree f_l of attribute A_l is: $f_l = t_l/n$.

1e) Assume the original attribute order is $A_1, A_2, \ldots A_m$. After step 1d, the sorted order will be $A'_1, A'_2, \ldots A'_m$.

1f) Subjectively, decide a threshold and then select relevant attributes to meet this threshold. At this point, the relevant attributes have been selected.

II. Building initial membership functions

2a) find the initial default group number of each relevant attribute

2b) find the range of each attribute

2c) find the group interval of each attribute

2d) extend the possible minimum attribute value of the relevant attribute

2e) divide the possible range of each attribute into G groups

2f) find the point b, for each initial membership function. For simplicity, only triangular membership functions are used here to represent the fuzziness of each interval. A triangular membership function can be defined by a triad (a, b, c):

$$b_{ij} = \sum_{s=1}^{r_{ij}} A'_{ij}(I_{ijs}) / r_{ij}, \tag{2-5}$$

where $A'_{ij}(I_{ijs})$ represents the attribute value of instance I_{ijs} in A'_{ij}.

2g) find the points a and c for each membership function:

$$a_{ij} = b_{i(j-1)} \text{ and } c_{ij} = b_{i(j+1)} \tag{2-6}$$

III. Deriving decision rules

The steps that can be followed to derive the decision rules are almost the same as those used in Hong and Lee's method. After step 2, the initial membership functions will be modified to be the final membership functions. The decision rules can be derived from the decision table.

IV. Discussion of this method

Since this technique is essentially an improvement of Hong and Lee's method, it will work well for more than 4 variables. The assumption of this method, which eliminates certain insignificant variables, is that there is no correlative relationship between these variables. This means that when we simplify our input variables, we also lose some information. It is a trade-off: the more information we keep, the more complex our calculation will be. But if we use this method properly in our model, we will obtain good results.

2.5 Methods of Inducing Fuzzy Expert Rules

1) General

The problem of generating fuzzy expert rules is one of the most important issues in the development of fuzzy expert system models. This problem is an example of a more general knowledge engineering problem called "knowledge acquisition." Yager (1991) described a very general framework for the formulation of rule bases for a fuzzy system model. Yager suggested that this formulation could be accomplished by a two-step procedure. The first step is the generation of the rules. The second step is tuning, which can be accomplished by either of two methods: parameter modification or rule weighting. However, it is the first step that is the most difficult to automate, as discussed by Yager (1991).

Many methods have been developed for rules generation, such as Hong and Lee's method and Hong and Chen's method, and the Mountain Clustering Method.

2) Description of Methods

a) Hong and Lee's method and Hong and Chen's method

As described in the previous section, these two method will generate membership functions for all the attributes and then automatically produce the If-Then rules. Once we decide to use this method, we have to use it for both membership function generation and if-then rules induction. The description of this method has been stated previously.

b) Mountain clustering technique

Yager and Filev (1994) used the mountain-clustering technique to develop fuzzy expert systems. In their model, first of all, the structure and the initial estimates of the parameters of the rule base of the fuzzy system model are obtained through the mountain clustering method. Then the back-propagation algorithm is used to tune the model. The strength of such a method is that it presents a better representation of input regions where the data are sparse. Another desirable feature is that it allows multiple rules to fire in a given region.

In applying this method to the proposed model, the complexity of the model and its extensive number of input variables make this method infeasible.

2.6 Summary

Many methods have been used for design performance evaluation. All have tried to solve one common problem, namely how to quantify linguistic inputs into numeric data that is easy to handle. Some researchers tried to simplify the model to a few easily quantified factors, and some just discarded the problem and solved it qualitatively. Previous research illustrates that design performance evaluation is subjective in nature. Traditional quantitative methods can never meet this challenge.

Fuzzy set theory is a generalization of set theory. It was developed specially for dealing with uncertainties and vagueness that are not statistical in nature. It provides a good way to quantify linguistic terms. It also provides the ability to relate multi-inputs to the output, especially when the reasoning relationship is difficult to express in numeric equations or involves linguistic judgment. Fuzzy set theory provides the ability to deal with problems involving natural language, subjective judgments, complex relationships, and limited data.

Previous research on generating fuzzy membership functions and fuzzy expert rules, although not entirely suitable for handling the proposed model, has provided a background for the methods developed in this thesis.

This thesis applies fuzzy set theory to the evaluation of design performance. Fuzzy set theory is chosen because it suits the nature of the problem to be solved, because of the problem's qualitative, subjective, and imprecise nature. This research has also been inspired by the desire to apply fuzzy set theory to a new application area.

CHAPTER 3 A MODEL OF FACTORS INFLUENCING DESIGN PERFORMANCE

3.1 Identifying Factors that Influence Design Performance

1) General

Design is such a complex process that no single factor can be used to predict nor evaluate its performance. First, design involves contributions from a number of groups, including engineers, clients, and contractors. Although this research is focused on design performance, the design problem can not be considered in isolation. One has to look at this problem from a wider perspective to consider all of the involved groups.

Second, design is a dynamic process that begins when a client proposes a preliminary plan, which continues throughout the entire construction life cycle until the contractor finishes the project. To evaluate design performance, one cannot use one simple factor such as the number of drawings per hour, day, or week, since this is a static variable. There are many factors that affect design performance before design even begins, and once the design work is complete a new set of variables can be used to evaluate its performance. On that note, we have to find a systematic and complete set of factors to evaluate design performance.

Third, design performance may vary depending on different situations or project contexts. Different project locations, different climates, different types of owners, etc, will all have a great effect on design performance. However, these factors are relatively stable for a given area. They may not seem significant in their effect on design performance, but if we compare two projects in two very different contexts, these factors will have a great effect on design performance.

In a word, design performance needs a complete, dynamic, and comprehensive set of criteria for evaluation. A thorough literature review was conducted to compile a comprehensive list of factors affecting design performance (CII, 1986 and APEGGA, 1990a, 1990b, 1998a, 1998b). These factors can be classified according to three groups, based on their different functions, described next.

2) Classes of Factors

- Context variables.

  The factors in this group are used to describe a given project context. They influence the design performance before any physical design work starts. These factors are qualitative. They are used to classify design projects into similar groups. Once these factors of a design project are specified, they can be considered as fixed variables, i.e., they will not vary for a given project.

  The variables in this group are listed in Table 3-1:

Table 3-1: Context Variables

Variables	Categories
Type of project	Oil-gas: pipeline Oil-gas: refinery, compressor station Chemical processing or extraction plant Mining Pulp and paper mill Power plant Water treatment plant Other
Type of design contract	Lump-sum Unit price Cost plus (cost reimbursable) Guaranteed maximum price Negotiated Other
Scope of design contract	Design only Design and manage (construction or project management) Design and build Other
Type of construction contract	Lump-sum Unit price Cost plus (cost reimbursable) Guaranteed maximum price Negotiated Other
Scope of construction contract	Construct only Design and build Management (construction or project management) Other
Method of tendering for the construction contract	Open Prequalified Other
Project priorities	Cost Schedule Quality Safety Aesthetics Environmental impact Constructability Potential for future development or expansion

- Input variables.

The factors in this category can be considered as the design firm's input to the design work, and they come into play before construction starts. Each of these factors is variable, i.e., they can vary from project to project in a given context. The factors in this stage can be used to predict the output of the design firm's performance for a given project. They are listed in Table 3-2. Each factor can be described using a linguistic term and a numerical measure, as listed in Table 3-2.

In reality, there are many factors that affect the final design performance. The following aspects are taken into consideration:

1) The surrounding economic conditions, as reflected by factor 2 and 14.
2) The profile of the owner, the designer, and the prime vendors, as reflected by factors 1, 3, 10 and 11.
3) Other pre-design factors, such as project function complexity (factor 7), size of the design contract (factor 4), continuity of manhour commitment (factor 5), design scope definition clarity (factor 6), and project environmental conditions (factor 9).
4) The dynamic factors that happen in the project life cycle. Those factors include the complexity of the design process (factor 8), the complexity of construction tendering process (factor 12), and the complexity of the construction process (factor 13).

Table 3-2: Input Factors Influencing Design Performance

Factor No.	Name of Factor	Linguistic Descriptors	Numerical Scale
Input #1	Overall size of the design firm	Small, average, large	1-10 rating
Input #1.1	Number of employees	Small, average, large	Real numbers
Input #1.2	Annual volume of work	Small, average, large	Real Numbers
Input #1.3	Number of projects held	Small, average, large	Real numbers
Input #2	Competition level in the market	Low, average, high	1-10 rating
Input #2.1	Number of similar design firms	Small, average, large	Real numbers
Input #2.2	Number of projects available in the market	Small, average, large	Real numbers
Input #3	Design firm overall quality	Poor, average, good	1-10 rating
Input #3.1	Scope of the project	Small, average, large	1-10 rating
Input #3.2	Number of designers involved	Small, average, large	Real numbers
Input #3.3	Senior to junior designers ratio	Small, average, large	Real numbers
Input #3.4	Level of design team skill	Poor, average, good	1-10 rating
Input #3.5	Average experience of design team	Small, average, large	Real numbers (years)
Input #3.6	Skills of the design team supervisor	Poor, average, good	1-10 rating
Input #3.7	Number of years of supervisor's experience	Small, average, large	Real numbers (years)
Input #3.8	Number of personnel changes	Small, average, large	Real numbers
Input #3.9	Level of familiarity with CAD	Low, average, high	1-10 rating
Input #4	Size of the design contract	Small, average, large	1-10 rating
Input #4.1	Total cost of design	Small, average, large	Real numbers (dollars)
Input #4.2	Duration of the design process	Short, average, long	Real numbers (months)
Input #4.3	Number of man-hours expended on the design	Small, average, large	Real numbers (hours)
Input #5	Continuity of the manhour commitment for the project	Small, average, large	1-10 rating

Table 3-2: Input Factors Influencing Design Performance (continued)

Factor No.	Name of Factor	Linguistic Descriptors	Numerical Scale
Input #5.1	Number of manhours per week, per designer, on the project	Small, average, large	Real numbers (manhours)
Input #5.2	Total manhours on the project	Small, average, large	Real numbers (manhours)
Input #6	Level of clarity of scope definition	Low, average, high	1-10 rating
Input #6.1	Clarity of project definition	Low, average, high	1-10 rating
Input #6.2	Description of alternatives considered	Low, average, high	1-10 rating
Input #6.3	Percent of data available before design	Small, average, large	Percent (%)
Input #6.4	Amount of information from previous projects	Small, average, large	1-10 rating
Input #7	Complexity of project functions	Low, average, high	1-10 rating
Input #7.1	Repetition of design	Small, average, large	Percent (%)
Input #7.2	Percent of new design	Small, average, large	Percent (%)
Input #7.3	Percent of innovation	Small, average, large	Percent (%)
Input #7.4	Percent of specified system special requirements	Small, average, large	Percent (%)
Input #7.5	Special consideration of the building envelope	Small, average, large	Percent (%)
Input #8	Complexity of the design process	Low, average, high	1-10 rating
Input #8.1	Number of design contracts involved	Small, average, large	Real numbers
Input #8.2	Number of locations of project	Small, average, large	Real numbers
Input #8.3	Number of owners or stakeholders involved	Small, average, large	Real numbers
Input #8.4	Number of review authorities involved	Small, average, large	Real numbers

Table 3-2: Input Factors Influencing Design Performance (continued)

Factor No.	Name of Factor	Linguistic Descriptors	Numerical Scale
Input #8.5	Average length of time for reviewing	Short, average, long	Real numbers (days)
Input #8.6	Number of environmental assessment reviews	Small, average, large	Real numbers
Input #9	Complexity of project conditions	Low, average, high	1-10 rating
Input #9.1	Insufficient working area	Small, average, large	Real numbers
	Magnitude of this problem	Small, average, large	1-10 rating
Input #9.2	Restricted access to site	Small, average, large	Real numbers
	Magnitude of this problem	Small, average, large	1-10 rating
Input #9.3	In-situ soil conditions	Small, average, large	Real numbers
	Magnitude of this problem	Small, average, large	1-10 rating
Input #9.4	Air temperature	Small, average, large	Real numbers
	Magnitude of this problem	Small, average, large	1-10 rating
Input #9.5	Amount of precipitation	Small, average, large	Real numbers
	Magnitude of this problem	Small, average, large	1-10 rating
Input #9.6	Lack of services available to the site	Small, average, large	Real numbers
	Magnitude of this problem	Small, average, large	1-10 rating
Input #9.7	Compatibility with land use zoning	Small, average, large	Real numbers
	Magnitude of this problem	Small, average, large	1-10 rating
Input #9.8	Disposal of contaminated materials	Small, average, large	Real numbers
	Magnitude of this problem	Small, average, large	1-10 rating
Input #10	Quality of the owner's profile	Poor, average, good	1-10 rating
Input #10.1	Time used by owner to make a decision	Short, average, long	Real numbers (days)
Input #10.2	Number of times owner changed or interfered	Small, average, large	Real numbers
Input #10.3	Number of changes of the owner's personnel	Small, average, large	Real numbers

Table 3-2: Input Factors Influencing Design Performance (continued)

Factor No.	Name of Factor	Linguistic Descriptors	Numerical Scale
Input #10.4	Working experience of the owner's representative	Small, average, large	Real numbers (years)
Input #10.5	Owner's attitude toward risk	Risk averse, average, risk prone	1-10 rating
Input #11	Quality of the vendor's profile	Poor, average, good	1-10 rating
Input #11.1	Length of receiving certified information	Short, average, long	Real numbers (days)
Input #11.2	Completeness of certified information	Small, average, large	Percent (%)
Input #11.3	Number of errors found	Small, average, large	Real numbers
Input #12	Complexity of the tendering process	Low, average, high	1-10 rating
Input #12.1	Number of work packages involved	Small, average, large	Real numbers
Input #12.2	Percentage of non-standard forms and conditions for the project	Small, average, large	Percent (%)
Input #13	Complexity of construction process	Low, average, high	1-10 rating
Input #13.1	Completeness of design before construction	Small, average, large	Percent (%)
Input #13.2	Number of construction phases	Small, average, large	Real numbers
Input #13.3	Number of prime contractors involved	Small, average, large	Real numbers
Input #13.4	Number of sub-contractors involved	Small, average, large	Real numbers
Input #13.5	Design consultant site-visiting frequency	Small, average, large	Real numbers /week, month
Input #13.6	Percentage of renovations or additions	Small, average, large	Percent (%)
Input #13.7	Early occupation required by the owner	Small, average, large	Real numbers (days)
Input #14	Economic (market) conditions	Unfavorable, average, favorable	1-10 rating

- Output variables.

Design performance evaluation can not be considered as complete unless construction is complete. At this stage, all of the design outcome information becomes available, therefore, evaluating design performance from the output point of view also becomes possible.

The output factors used to measure design performance are listed in Table 3-3 together with their linguistic and numerical measures. They can be classified according to three categories:

1) Level of performance against the cost of design;
2) Level of performance against the schedule for the design project;
3) Level of accuracy of design performance (i.e., quality).

Other measures of design performance that were not used in the model are:

1) level of usability of design documents
2) level of constructability
3) level of economy of design
4) level of ease of start-up.

Table 3-3: Output Factors Influencing Design Performance

Factor No.	Name of Factor	Linguistic Descriptors	Numerical Scale
Output #1	Level performance against cost	Poor, average, good	1-10 rating
Output #1.1	Percentage of changes of manhours (actual/budgeted)	Small, average, large	Percent (%)
Output #1.2	Manhours due to change orders	Small, average, large	Percent (%)
Output #1.3	Rework manhours	Small, average, large	Percent (%)
Output #1.4	Percentage of changes of cost (actual/budgeted)	Small, average, large	Percent (%)
Output #1.5	Design cost due to change orders	Small, average, large	Percent (%)
Output #1.6	Design cost/total construction cost	Small, average, large	Percent (%)
Output #2	Level of performance of schedule	Poor, average, good	1-10 rating
Output #2.1	Percentage of changes of duration (actual/schedules)	Small, average, large	Percent (%)
Output #2.2	Design document release deadline missed	Small, average, large	Percent(%)
Output #3	Level of accuracy of design documents	Low, average, high	1-10 rating
Output #3.1	Number of changes	Small, average, large	Real numbers
Output #3.2	Total cost of approved changes	Small, average, large	Real numbers (days)
Output #3.3	Impact of design changes on cost	Small, average, large	Percent (%)
Output #3.4	Total number of design rework manhours	Small, average, large	Real numbers
Output #3.5	Design accuracy	Small, average, large	Real numbers

3.2 Identifying Measures for each Factor Affecting Design Performance

Even though a thorough list of factors that affect design performance was developed, most are still too abstract to be quantified. Each higher-level factor was therefore broken down into a series of sub-factors, which are easier to quantify.

The functions of these sub-factors are as follows:

1) To transfer subjective measures into objective ones.

Each higher-level factor is subjective in nature and is therefore difficult to measure in a universal fashion. Sub-factors for each factor were required in order to provide objective measures of each factor. By defining the sub-factors objectively (i.e., with a number), one can ensure that the definition of the higher-level factor has the same meaning to all users of the model. For example, input factor 1 is the size of the design firm. In addition to a subjective judgment (i.e., a scale from 1 to 10), this factor can be broken down into 3 sub-factors that can be easily quantified. The three sub-factors are the number of people employed, the average annual volume of work of the firm, and the number of projects that the design firm currently has. All of the higher-level subjective factors have a detailed description in terms of their sub-factors. Both the input and output factors are described in terms of sub-factors.

2) To transfer an abstract concept into a specific concept.

Another function of the sub-factor is to describe the concept that is being considered. Each factor involves several aspects, and sub-factors illustrate the factor in a more comprehensive way. As a result, the meaning of each factor can be understood easily and clearly.

3) To provide for special considerations for Input factor 9 (project condition complexity).

For this factor, the author not only developed a series of sub-factors, but also specified two types of sub-factors. One is the number of occurrences of each problem, the other is the magnitude of this problem. In this model, only the overall effect of each problem as mapped on the project condition complexity was concerned. This effect should incorporate both the frequency of each problem occurring, as well as the magnitude of each occurrence of the problem.

The numerical scales used for the measure each factor and sub-factor are listed in Table 3-2 and Table 3-3, together with their corresponding linguistic descriptors.

3.3 Structure of the Model of Factors

At this point, a complete set of criteria for design performance evaluation was developed based on the input and output factors identified. The framework of this model of factors is shown in Figure 3-1.

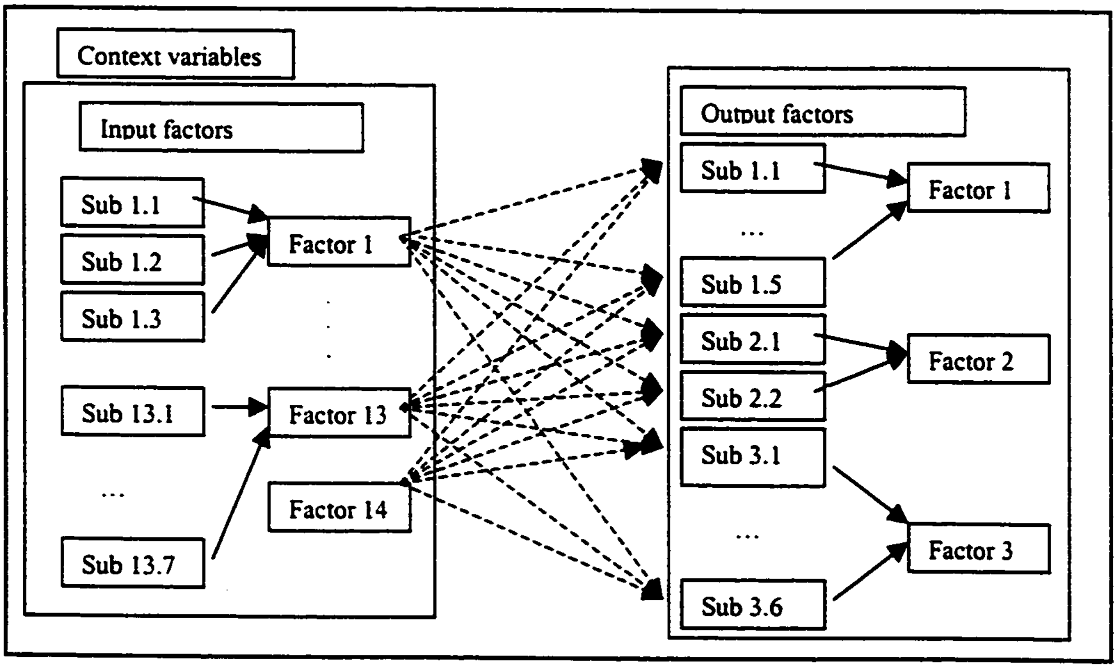


Figure 3-1: Structure of the Model of Design Performance Evaluation Factors

3.4 Model Description

Figure 3-1 shows the structure of the complete design performance evaluation model. Based on this structure, the model works as follows:

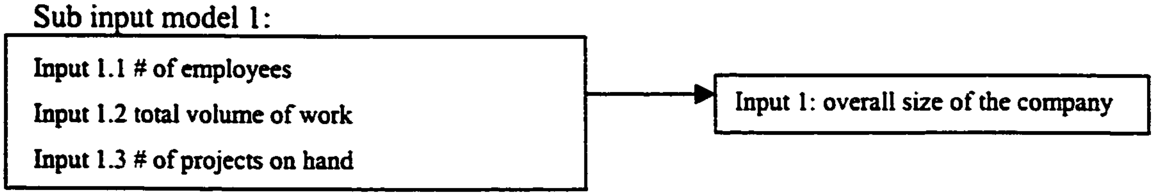
- Input factor section: the input sub-factors are used to determine each input factor in each sub-input model. These relationships are shown in solid lines in Figure 3-1.
- Output factor section: the output sub-factors are used to determine each output factor in each sub-output model. These relationships are shown in solid lines in Figure 3-1.
- All the higher-level input factors are used to predict each output sub-factor. These relationships are shown in dashed lines in Figure 3-1.

The model can be used in the following ways:

- To predict design performance. Based on the characteristics of the project and its environment (i.e., based on the input sub-factors), the higher-level input factors can be determined. The higher-level input factors can then be used to predict design performance, which is described by the output sub-factor.
- To describe design performance. The predicted (or actual) values for the output sub-factors can be used to describe the overall design performance, in terms of the three higher-level output factors.

The complete models are shown as follows.

Sub input model 1:



Sub input model 2:

Input 2.1 # of similar firms in the market
Input 2.2 # of projects available in the market
→ Input 2: competition level

Sub input model 3:

Input 3.1: design within normal scope extent
Input 3.2: # of designers involved
Input 3.3: senior to junior designer ratio
Input 3.4: skill level
Input 3.5: average experience (in years)
Input 3.6: motivation and leadership skill
Input 3.7: supervisor experience (in years)
Input 3.8: # of design personnel changes
Input 3.9: familiarity of design software usage
→ Input 3: design firm profile

Sub input model 4:

Input 4.1: total design cost
Input 4.2: design process duration
Input 4.3: total manhours expended
→ Input 4: size of design contract

Sub input model 5:

Input 5.1: manhour per week for each designer
Input 5.2: average manhour expended per week by the design team
→ Input 5: continuity of design

Sub input model 6:

Input 6.1: clarity of design definition
Input 6.2: clarity of design alternative definition
Input 6.3: percentage of available design data prior to detailed design
Input 6.4: amount of available similar design information
→ Input 6: scope definition level

Sub input model 7:

Input 7.1: percentage of design features repetition
Input 7.2: percentage of unique or new design features on the project
Input 7.3: percentage of upgrades or changes to the existing facility
Input 7.4: percentage of specialized design requirement
Input 7.5: percentage of special considerations on the project for the building envelope.

→ Input 7: project function complexity

Sub input model 8:

Input 8.1: number of design contracts involved
Input 8.2: number of engineered locations
Input 8.3: number of owners or stakeholders involved
Input 8.4: number of review authorities involved
Input 8.5: review duration
Input 8.6: number of environmental assessment reviews

→ Input 8: design process complexity

Sub input model 9:

Input 9.1: effect of insufficient working area
Input 9.2: effect of restricted access to site
Input 9.3: effect in-situ soil conditions
Input 9.4: effect of air temperature
Input 9.5: effect of precipitation
Input 9.6: effect of lack of services available to site
Input 9.7: effect of requirement of compatibility with land use zoning
Input 9.8: effect of requirements of contaminated materials disposal

→ Input 9: project condition complexity

Sub input model 10:

Input 10.1: duration of the owner to make a decision
Input 10.2: number of the changes of the owner
Input 10.3: number of the owner's personnel changes
Input 10.4: owner representative's experience (in years)
Input 10.5: owner's attitude toward risk

→ Input 10: quality of owner's profile

Sub input model 11:

Input 11.1: length to receive the certified information
Input 11.2: percentage of completeness of certified information
Input 11.3: number of errors of certified

→ Input 11: quality of vendor's profile

Sub input model 12:

Input 12.1: number of construction packages
Input 12.2: percentage of non-standard forms and conditions

→ Input 12: complexity of tendering process for construction

Sub input model 13:

Input 13.1: percentage of completed design prior to construction
Input 13.2: number of construction phases
Input 13.3: number of prime contractors involved
Input 13.4: number of sub-contractors involved
Input 13.5: frequency of site visits by the design consultant
Input 13.6: percentage of renovation or addition work
Input 13.7: days of early occupied requirement by the owner

→ Input 13: complexity of construction process

Sub output model 1:

Output 1.1: actual design manhours over budgeted one
Output 1.2: percentage of design manhours due to owner changes
Output 1.3: percentage of design manhours due to rework
Output 1.4: actual design cost over budgeted
Output 1.5: design cost over the cost due to the owner changes
Output 1.6: design cost over total construction cost

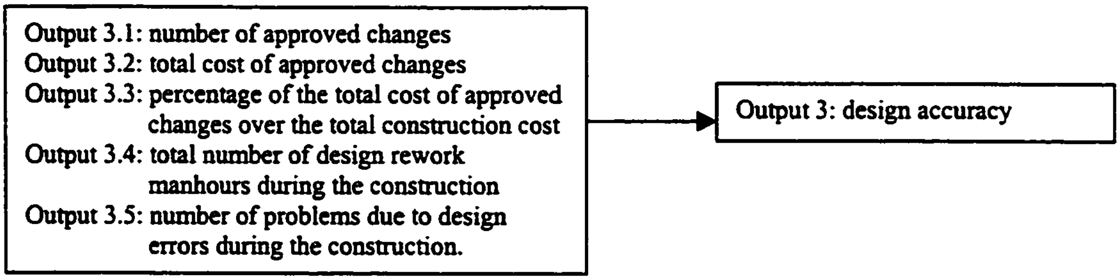
→ Output 1: cost performance

Sub output model 2:

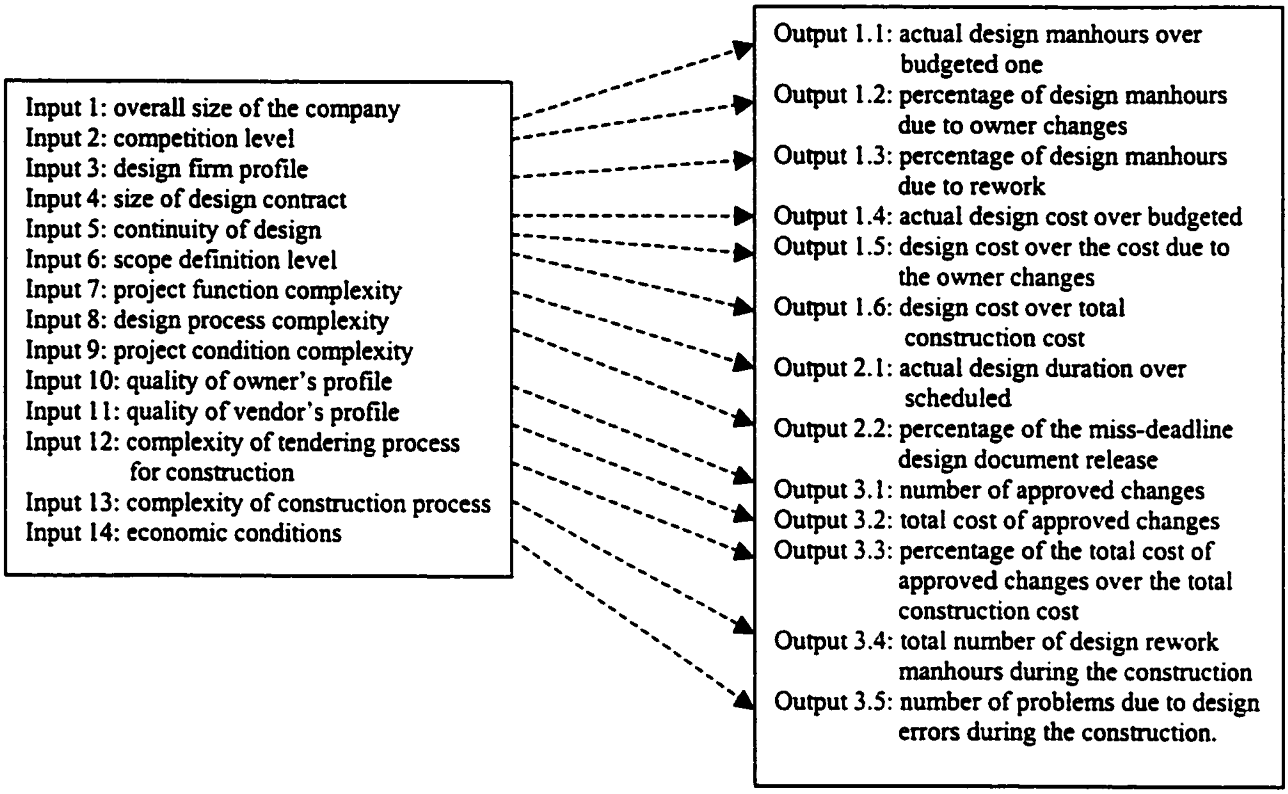
Output 2.1: actual design duration over scheduled
Output 2.2: percentage of the miss-deadline design document release

→ Output 2: schedule performance

Sub output model 3:



Input factors to each sub output factor



3.6 Summary

The model presented in this chapter provides the basis for the fuzzy expert system used to predict and evaluate design performance. The next step in the development of the fuzzy expert system is the collection of data to refine and test the system. The data collection process is described in Chapter 4.

CHAPTER 4 DATA COLLECTION AND PROCESSING

4.1 Introduction

A survey was designed to collect the data needed to implement and test the design performance prediction and evaluation model described in Chapter 3. The main goals of this survey are:

- To collect actual data from completed projects by design firms for the purposes of modeling.
- To identify the profiles of industrial design firms in Alberta and British Columbia.

4.2 Survey Methodology

Based on the model of factors stated in Chapter 3, a survey questionnaire was developed. The questions were specifically designed to obtain an accurate, representative value for each of the factors and subfactors in the model. In addition to the numerical value for each factor, the questions also required a linguistic description of each factor. The collection of both numerical and linguistic values for each factor was specifically for the purpose of generating membership functions, described in Chapter 5. The survey questionnaire is contained in Appendix 1.

Before the questionnaires were mailed to companies, a pilot study was conducted. The purposes of this pilot study were to refine the survey questions to meet actual industry practices and to check the appropriateness of the survey in terms of meeting the data collection requirements. The two persons contacted for the pilot study were Mr. Paul Veldoen of Colt Engineering (Calgary) and Dr. Witold Pedrycz (a fuzzy logic researcher) in the Electrical Engineering Department of the University of Alberta, both of whom are well qualified to comment on the survey questionnaire. The preliminary survey questionnaire was discussed with them for their suggestions. They both provided valuable information about the survey questionnaire. Based on discussions with them, the following modifications were made to the questionnaire. First, respondents were allowed to circle more than one linguistic term and second, to provide a range of numerical values, rather than be restricted to a single value. They also provided general comments on the questionnaire, such as focusing it on a particular sector of the construction industry, to conduct an initial small-scale survey rather than a widespread survey to determine likely response rate and the quality of the responses. Consequently, the survey was mailed out only in Alberta and British Columbia and only the industrial sector was chosen for this research.

The survey was carried out in two phases. First, to test the effectiveness and response rate, the survey was initially targeted at industrial design consultants in Alberta. This allowed the typical response rate to be determined and a review of the quality of responses to be completed before a large-scale survey was conducted. The list of registered design consultants in Alberta was obtained from the Consulting Engineers of Alberta Internet site. Each of these firms was contacted first to determine whether the firm's profile included industrial design and, if so, the name of a contact with whom correspondence could be made. The survey was mailed to the companies on this list at the beginning of June 1999. Follow-up calls were made at the end of June to each company that had not yet returned the completed survey.

After the survey in Alberta was completed, some minor changes to the wording of particular sections were made to enhance the quality of the information obtained.

However, no fundamental modifications were made to any of the questions or the factors contained in the survey.

In addition, the survey feedback from Alberta showed that the initial survey population of industrial design consultants in Alberta would not yield enough responses for the survey to be successful and meet its objectives. Therefore, in July 1999, a listing of all registered Canadian engineering design firms was purchased from the Association of Consulting Engineers of Canada. The population of companies based in British Columbia and involved in the industrial sector was determined through the profiles provided, and these firms were subsequently contacted to obtain a contact name. The modified survey was mailed to each of the companies on this list at the beginning of July 1999. Similar to the procedure followed for Alberta-based companies, follow-up calls were made to non-responsive companies in British Columbia at the beginning of August 1999.

4.3 Survey Response Rate

The survey was completed at the end of September 1999. 78 surveys were mailed out and 26% (18) sent back the completed questionnaires. A 30% response rate is typically considered acceptable for surveys administrated by mail (according to Statistics Canada). The response rate obtained in this survey is typical of mail-out surveys used in construction research, which normally have response rates of 20% to 30% (Ahmad and Minkarah, 1988). The reasons that led to such a low response rate came from several aspects. First, the questionnaire is too complex and it requires the company to devote sufficient time and knowledge to complete it. The survey was conducted in the summer time, which is the busy construction season, so it is understandable that the survey cannot get more companies' participation. Second, some companies provided inaccurate information at first. Some companies indicated that they worked in the industrial sector when initially contacted, however, upon receiving the survey, responded that they did not fit into this target population. Due to this reason, the actual sample population was reduced to 68 instead of 78. The survey response rate is shown in Table 4-1.

Table 4-1: Survey Response Rate

	Total	AB	BC
Surveys Mailed	78	52	26
Completed Surveys	18	14	4
Not Industrial Sector	10	9	1
Response Rate	**26%**	**33%**	**16%**

- companies not in the industrial sector are not included in calculation of response

4.4 Data Processing Procedure

The survey consists of three sections. Section 1 was developed to collect general information from the respondents. This type of information will be used in future research on context variables. The respondent results of this section are shown in Appendix 2.

The next two sections are focused on the details of a specific project selected by the respondent. Section 2 lists the fourteen input factors and sub-facts that impact design performance. Similarly, Section 3 lists three output factors and sub-factors that affect design performance. For these two sections, the respondents were asked to provide both a numerical value (or in some cases a rating on a scale of 1 to 10) and a linguistic description (e.g., small, average, large) for each factor.

For modeling purposes, the information in these two sections was compiled using the following technique. First, the overall 18 responses were divided into two sets of data, that is, one training data set and one testing data set. Second, a graph that represents the frequency of the responses for each linguistic term in the survey was developed for each factor using the training data set. The graphs in Figure 4-1 to 4-4 illustrate this method. Based on these graphs, the general shape of each membership function can be visually determined. This step provides the base stone of generating membership functions.

If a factor has an upper limit to its values (i.e. x-axis), then the frequencies of numerical responses for each linguistic term for this factor were plotted on one graph (see Figure 4-1). If a factor does not have an upper limit to its values, the x-axis value may vary a lot for different linguistic terms. In this case, a separate graph was developed for each linguistic term. Figure 4-2 to Figure 4-4 illustrate this concept. If a linguistic term did not receive any responses, then there will be no values on its graph (Figure 4-4 is an example).

Number of Responses = 15

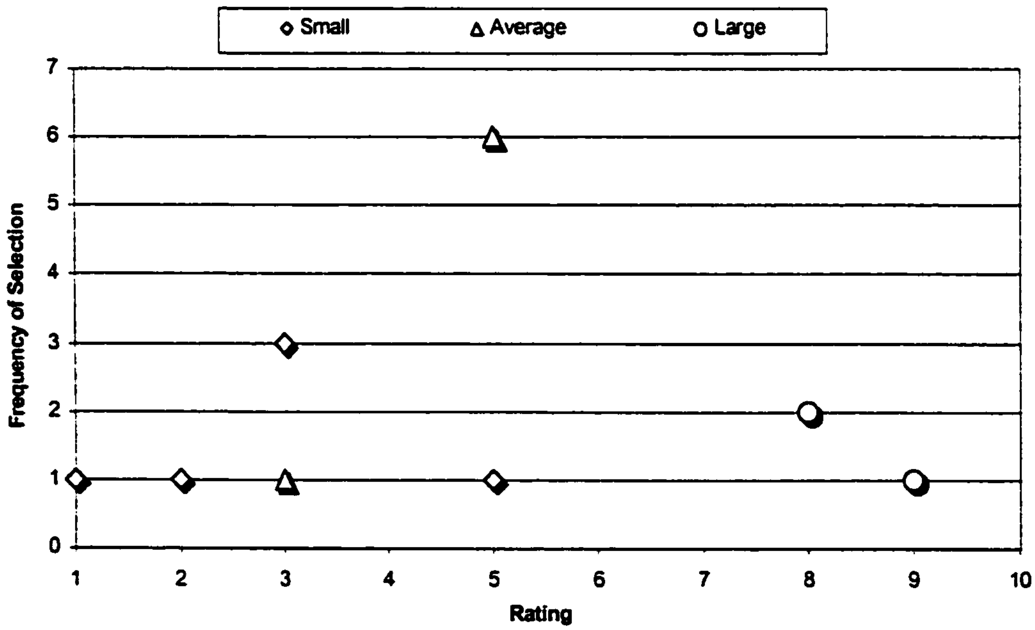


Figure 4-1: Rating of the Overall Size of Industrial Division of Firm

Number of Responses = 12

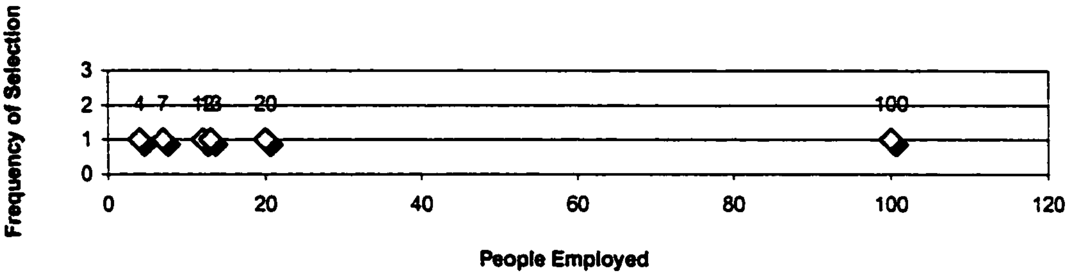


Figure 4-2: Number of People Employed by Firm, Variable 'Small '

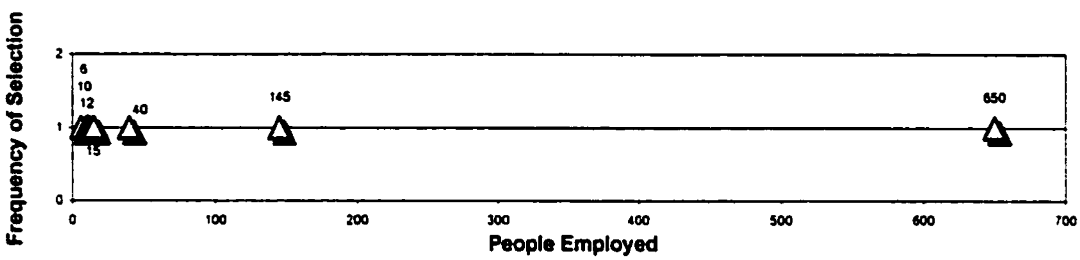


Figure 4-3: Number of People Employed by Firm, Variable 'Average'

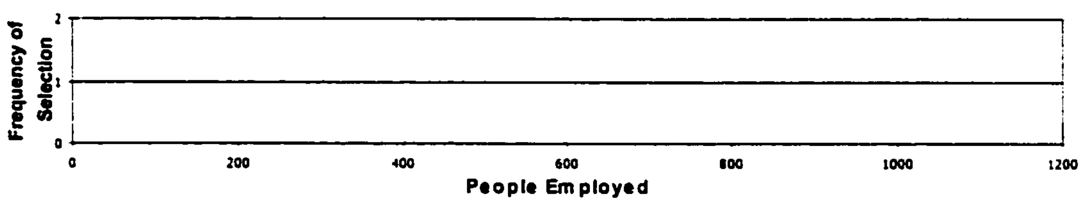


Figure 4-4: Number of People Employed by Firm, Variable 'Large'

This data processing procedure was performed for all the factors in each of the two different training data sets. The complete results are shown in Appendix 3.

The characteristics of input Factor 9 (complexity of project conditions) as shown in Chapter 3, required a slightly different analysis. Even though each sub-factor, representing a type of problem such as insufficient working area, restricted access to site, in-situ conditions, has two variables (number of occurrences and magnitude), only the overall effect of each sub-factor on the factor was considered as follows:

1) For the linguistic terms describing each sub-factor, Table 4-2 shows the overall effect based on all possible combinations of the two variables.

Table 4-2: Linguistic Combinations for Sub-factors related to Input Factor 9

Combinations (in any order)	**Overall effect**
Small-small	Small
Small-average	Small
Small-large	Average
Average-average	Average
Average-large	Large
Large-large	Large

- The above rules are based on intuition and approximation.
- If one descriptor is missing, then the other that is given is used.
- If neither of the two variables was rated, then the response was ignored.
- This method does not take into account the nature of the sub-factors (i.e., the type of the problem).

2) For the numerical data describing the variables in each sub-factor, the overall effect of each problem (i.e. sub-factor) is calculated as follows:

$$Overall\_effect = \frac{(Number\_of\_Occurrences)*(Magnitude)}{10} \qquad (4\text{-}1)$$

If a range of numbers was given for each variable, then a fuzzy multiplication operation is performed to obtain the overall effect, as follows:

$$[a,b]\otimes[c,d]=(\min[ac, ad, bc, bd], \max[ac, ad, bc, bd]) \quad (4\text{-}2)$$

4.5 Summary

This chapter presented the survey technique used to collect actual data for modeling purposes. The data collected was separated into a training data set and a testing data set. The training set was used to develop the fuzzy membership functions and to generate the If-Then rules for the fuzzy expert system. The testing training data set was used to test the accuracy of both. The following three chapters describe these procedures.

CHAPTER 5 GENERATING MEMBERSHIP FUNCTIONS

5.1 Problem Description

Generating membership functions is the first step involved in applying fuzzy set theory; yet this important step is often one of the most subjective. The reasons for its lack of definition arise from several aspects. First, methods of generating membership functions are often dependent on extensive data, which may not be available or are difficult to obtain. Second, some methods have limitations that make them inapplicable to certain problems. Often, membership functions are generated from experts or experience, making them context specific. Consequently, generating membership functions becomes one of the bottlenecks of applying fuzzy set theory.

Until now, research on design performance evaluation has been limited. The problem is also quite complex. The model for design performance evaluation described in Chapter 3 has 14 higher-level input factors, 3 higher-level output factors, and numerous sub-factors for each higher-level factor. To model this problem using fuzzy set theory, membership functions must be generated for each higher-level factor and its sub-factors.

A review of the literature (refer to Chapter 2) shows that no existing technique can be applied to this problem; that is, either the existing techniques cannot deal with multi-variables, say more than 4 (Hong and Lee, 1996) or else the technique requires a substantial data set to statistically plot membership functions (Li and Yen, 1995). Consequently, one of the goals of this research is to develop a technique for generating membership functions that are based on objective data, so that they can be calibrated to suit different contexts. Defining membership functions on the basis of objective data is a first step towards developing membership functions that are widely applicable in a given context. The survey described in Chapter 4 was specifically designed to collect the data required to implement and test the proposed method of generating membership functions. The context chosen is industrial engineering design. Each factor in the model can be described in terms of three linguistic terms, each of which is represented by a membership function. Each membership function represents the degree to which numerical levels of the factor reflect the linguistic term describing the factor. For example, for input factor 1, there are 3 membership functions, small, average, and large, used to describe the size of the design firm. The size of the firm can also be evaluated numerically on a scale from 1 to 10; each numerical value has a different degree of membership in each of the three membership functions.

This chapter describes the new method developed for generating membership functions, and how the data collected from the survey was used to develop and test this method.

5.2 Data Segregation

Chapter 4 describes the procedure used to collect and compile the data used for modeling. For generating membership functions, the 18 survey responses were separated into two sets: one is used as the training data set, and the other is the testing data set. In order to obtain more reliable results from the proposed technique, the procedure for generating membership functions was performed and tested twice, using two different training and testing data sets.

The steps for separating the complete data set into training and testing data sets are as follows:

- For the first trial, 13 survey responses were selected for training to build the membership functions and the remaining 5 responses were used to test the membership functions.
- For the second trial, 11 survey responses were chosen as the training data set, and the remaining 7 were used for testing purposes.

The criteria used for separating the data set are as follows:

- The two testing data sets cannot have any common survey responses. This means that none of the seven responses used for testing in the second trial were in the first trial.
- The testing samples should be typical among all of the responses in terms of type of projects and location. Therefore, the testing data should neither be from extreme outliers nor from non-typical projects.
- According to the above criteria, the samples in both testing data sets are shown in Table 5-1.

Table 5-1: Data Segregation for the Two Trials

	Project type	Quantity
1st trial testing data set	Chemical processing (extraction plant)	1 (ALBERTA)
	Mining	1 (ALBERTA)
	Oil/gas refinery/compressor	1 (ALBERTA)
	Water (waste water treatment plant)	2 (one in BRITISH COLUMBIA, one in ALBERTA)
2nd trial testing data set	Water treatment	1 (ALBERTA)
	Mining	1 (ALBERTA)
	Chemical Processing	1 (ALBERTA)
	Oil/gas pipeline	1 (ARGENTINA)
	Office and warehouse	1 (ALBERTA)
	Bulk material handling and transportation	1 (ALBERTA)
	Water supply	1 (BRITISH COLUMBIA)

5.3 Assumptions Underlying the Method of Generating Membership Functions

Using the training data set, the responses to each question were plotted separately on a graph, which represents the frequency of numerical responses for each linguistic descriptor. The method for generating membership functions was developed. The basic assumptions of this methodology are as follows:

1) According to Lorterapong and Moselhi (1996), "theoretically, fuzzy numbers can take various shapes. In modeling real-life problems, however, linear approximations such as the trapezoidal and triangular fuzzy numbers are frequently used." To simplify the problem, these two common membership function shapes (triangular and trapezoidal) were chosen to model each factor and sub-factor in the model. Another good feature of these two shapes is that they best represent the trend of the data collected from the survey. The triangular

shape can deal with peaks, and the trapezoidal one can handle ranges of numbers. The trapezoidal fuzzy number can be represented by a quadruple (a, b, c, d), while the triangular fuzzy number is a special case of the trapezoidal shape, with b=c (refer to Figure 5-1).

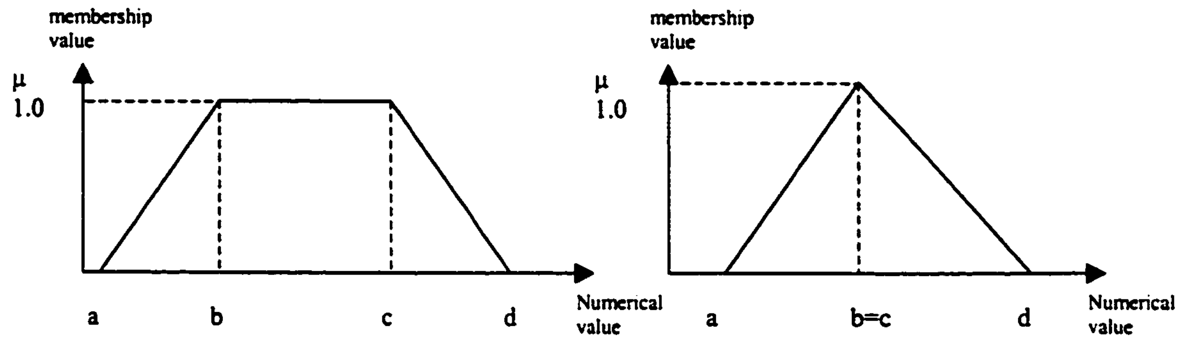


Figure 5-1: Trapezoidal and Triangular Membership Function

2) According to Li (1995), a stable frequency resulting from fuzzy statistical tests can serve as the degree of membership in the objective sense. In this case, the frequency of each response is used to reflect the membership value of the numerical value for the given linguistic term.
3) Since each factor in the model has 3 linguistic descriptors, then each may have 3 membership functions (small, average, large; or low, average, high; or poor, average, good). Basically, we have 3 groups: the first one includes small, low, or poor; the second includes average; and the third includes large, high, or good. The author assumes that the first and the third groups always have the trapezoidal membership functions, so that on the membership function any element in the first group less than a specific value (for example, point x in Figure 5-2) will be considered to have the membership value of 1. Further, any elements in the third

group with a larger value than a certain number (for example, point y in Figure 5-2), which has a membership value of 1, will be regarded as having the membership value of 1 as well. These shapes are illustrated in the Figure 5-2.

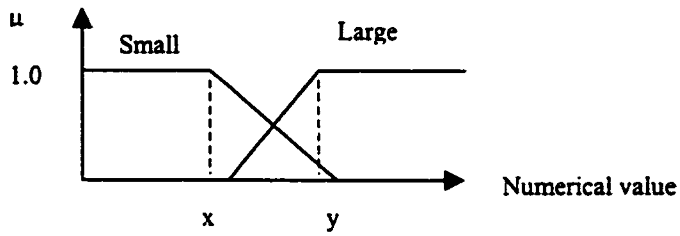


Figure 5-2: Trapezoidal Membership Functions for Smallest and Largest Linguistic Terms

There are two exceptions: when the smallest value (equal to the x-axis left limit) in first group has the highest frequency, and when the largest number (equal to the x-axis right limit, for example, point y in Figure 5-3) in the last group has the highest frequency. In these two cases only, the membership function of the first and third groups may be triangular. These shapes are illustrated in Figure 5-3:

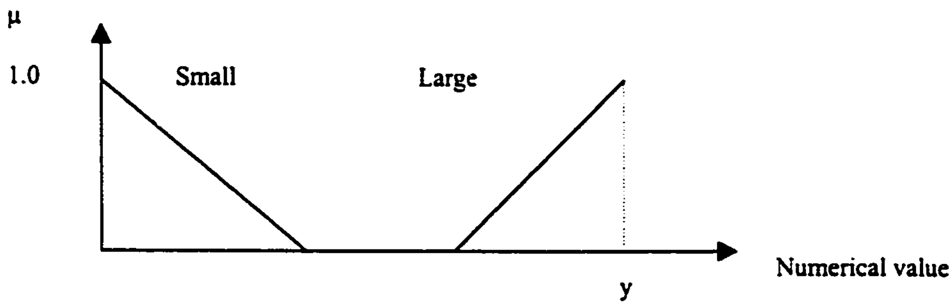


Figure 5-3: Special Membership Functions for Smallest and Largest Linguistic Terms

However, the second group, which deals only with averages, has the possibility of possessing both a triangular shape and a trapezoidal one.

The generic membership function of these two shapes can be expressed as:

$$\mu(x) = \begin{cases} 0, x < a \\ \dfrac{x-a}{b-a}, a <= x < b \\ 1, b <= x < c \\ \dfrac{x-d}{c-d}, c <= x <= d \\ 0, x > d \end{cases} \tag{5-1}$$

For the triangular shape, b=c. a, b, c, d are parameters for the membership functions which are illustrated in Figure 5-1. μ represents the membership value for a given numerical value x.

4) Choosing the shape of membership functions is only required for the second group of linguistic descriptors, that is, the group representing the linguistic term average.

4.1) Situations for choosing trapezoidal membership functions.

a) If there is more than one distinct peak (i.e., there are two or more values with the highest frequency), then choose the trapezoidal shape of membership functions.

b) If there is no peak at all (i.e., all the values have the same frequency), then also choose the trapezoidal shape.

4.2) Situations for choosing triangular membership functions.

a) If there is one distinct peak in the data set (i.e., there is one single data point with the highest frequency), then choose the triangular shape of membership functions.

b) If there is only one single value in the graph, then also assume that the shape of the membership function is triangular.

5) Some of the survey data contains overlapping values; in other words, a given numerical value may belong to more than one linguistic term. Overlapping occurs because different respondents may have different opinions on the level that a given value

represents. When the sample size is small, the answers may not be distinguishable. Dealing with this involves combining the overlapped membership functions into one group, that is, a new group called Small-Average or Average-Large may be created. This means that for this sample size, the problem cannot be described by three categories; perhaps only two are enough.

Taking Input Factor 1, for example, Figure 5-4 shows that the data for Small and Average cannot be distinguished. These two groups are therefore combined into one membership function called Small-Average. Thus, there are only two membership functions for this factor, Small-Average and Large, such that the two functions do not overlap extensively (refer to Figure 5-5).

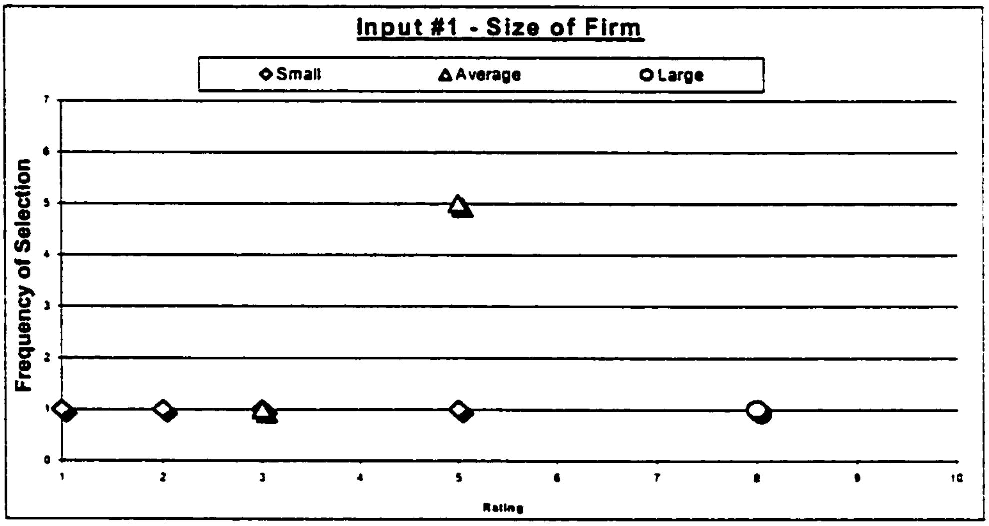


Figure 5-4: Original Frequency Graph for Input Factor 1 (Trial 1)

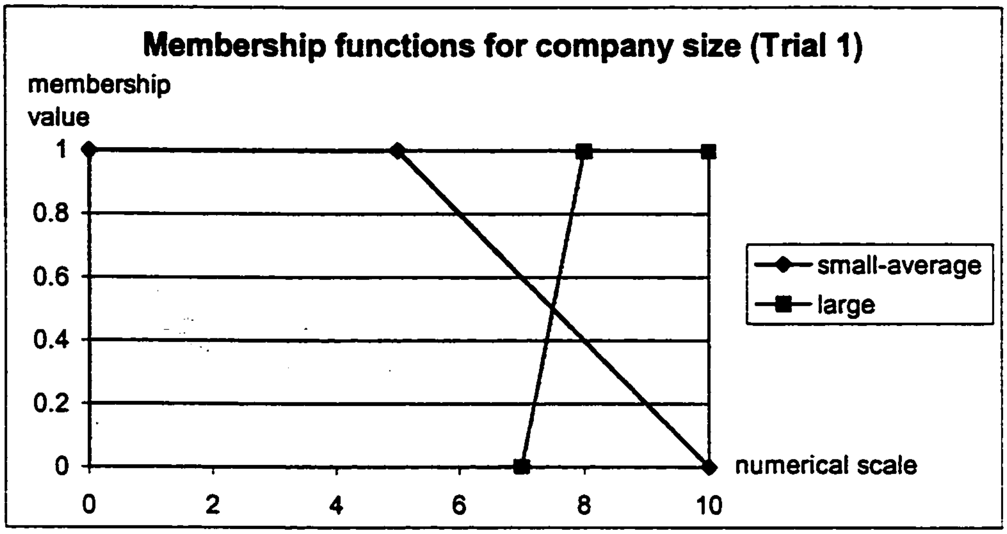


Figure 5-5: Membership Functions for Input Factor 1 in Trial 1

5.4 Description of Method for Generating Membership Functions

The following rules are used to determine the parameters of each membership function using the survey data. The peak number refers to the value with the highest frequency of responses.

<u>Case 1</u>: For the first group of linguistic terms (Small, Poor, or Low), let the peak = x, then μ_x =1.0, μ_0 to μ_x =1.0 (refer to Figure 5-6).

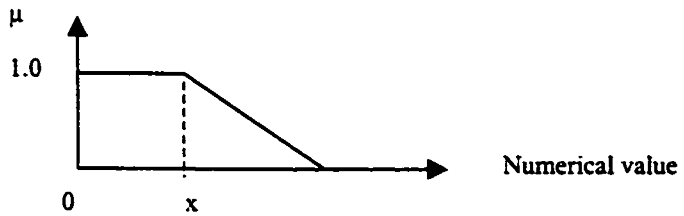


Figure 5-6: Membership Function for Small, Poor or Low

Rule 1: If the peak value x is given and also there is a value y given, where y > x, and y is the largest value on the x-axis, then let $\mu_y = F_y / F_x$, where F_y is the frequency of y and F_x is the frequency of x (refer to Figure 5-7).

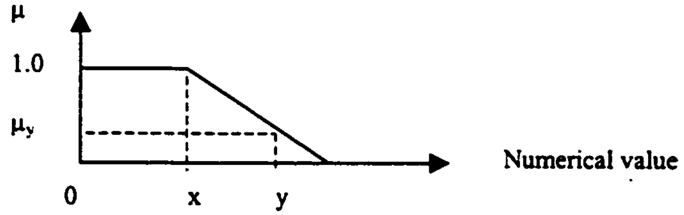


Figure 5-7: Membership Function for Small, Poor or Low using Rule 1

Rule 2: If the peak (x) is the largest or only value given and the adjacent membership function is available, let $y_{small} = y_{average}$, where $\mu_{ysmall} = \mu_{yaverage} = 0.5$. In this case, $z_{small} = 2\, y_{average} - x$ (shown in Figure 5-8).

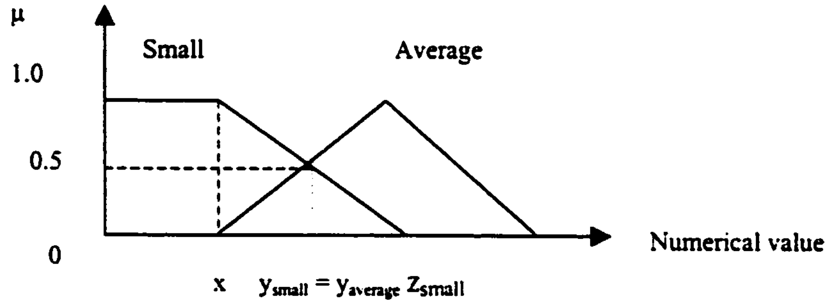


Figure 5-8: Membership Function for Small, Poor or Low using Rule 2

Rule 3: If the peak value is given, but there is no reference to the membership function of Average or the adjacent function is unavailable, (for example, the mid point of the average membership function is less than the peak value of small membership function), then let parameter d of the Small membership function equals 2 times the peak value x, that is 2*x (refer to Figure 5-9). In this case, only the peak value is required to determine the shape of the membership function.

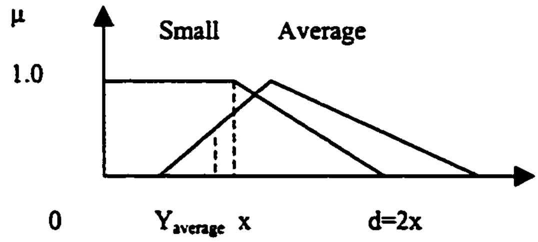


Figure 5-9: Membership Function for Small, Poor or Low using Rule 3

Rule 4: If there is no data at all for this linguistic term, but the relevant information of Average is available, then the following steps should be followed: let $y_{small} = y_{average}$, where $\mu_{ysmall} = \mu_{yaverage} = 0.5$. Then $x = \frac{2}{3} * y_{average}$ and z=2x. The four parameters for this membership function are then decided. This rule is illustrated in Figure 5-10.

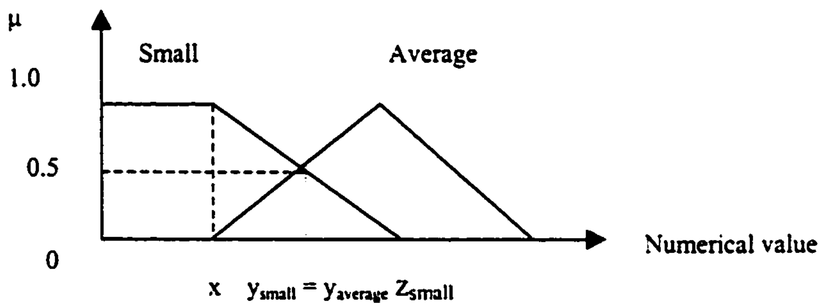


Figure 5-10: Membership Function for Small, Poor, or Low using Rule 4

Case 2: For the second group of linguistic term (Average), no matter whether the shape of the membership functions in this group is triangular or trapezoidal, the two legs are the keys to deciding the shape. Take a triangular shape membership function for example (Figure 5-11), let the peak = x, and $\mu_x = 1.0$.

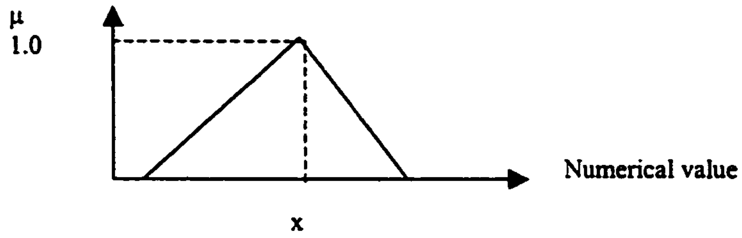


Figure 5-11: Membership Function for Average

Rule 5: If there is a lower limit y given, where y<x, and y is the lowest value, then let $\mu_y = F_y / F_x$, where F_y is the frequency of y and F_x is the frequency of x. If there is a higher limit z given, where z>x, and z is the highest value, then let $\mu_z = F_z / F_x$, where F_z is the frequency of z and F_x is the frequency of x (refer to Figure 5-12).

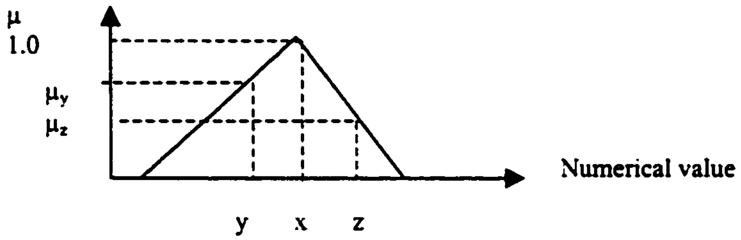


Figure 5-12: Membership Function for Average using Rule 5

Rule 6: If the peak is the smallest (or only value given) and there is a membership function for Small available, then let $y_{average} = y_{small}$, where $\mu_{ysmall} = \mu_{yaverage} = 0.5$ (refer to Figure 5-13).

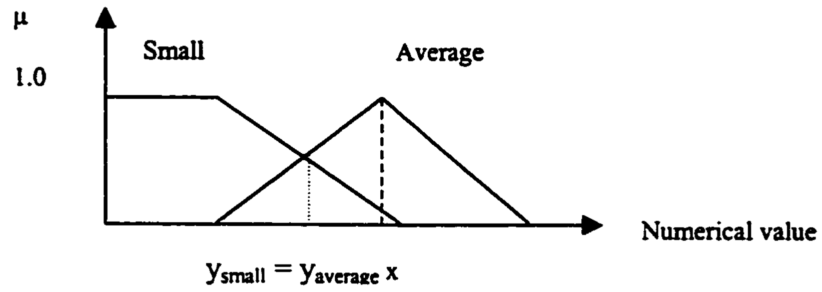


Figure 5-13: Membership Function for Average using Rule 6

Rule 7: If the peak is the largest or only value given, and there is a membership function for Large available, then let $Z_{average} = Z_{large}$, where $\mu_{zlarge} = \mu_{zaverage} = 0.5$ (refer to Figure 5-14).

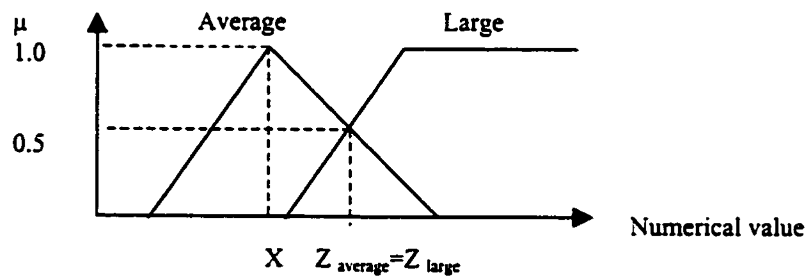


Figure 5-14: Membership Function for Average using Rule 7

Rule 8: If there is no information available to decide the other leg of the triangular membership function, then assume that the membership function for average is symmetric so that one can get the other leg from either available membership function.

Rule 9: If there are no data points for Average, then as long as the data from either the first group or the third group is available, the membership function can be generated using the above combined rules (Rules 6, 7 and 8).

Case 3: For the third group of linguistic terms (Large, Good, or High), let the peak = x, and μ_x =1.0, μ_x to $\mu_{x\text{-axis limit}}$ =1.0 (shown in Figure 5-15).

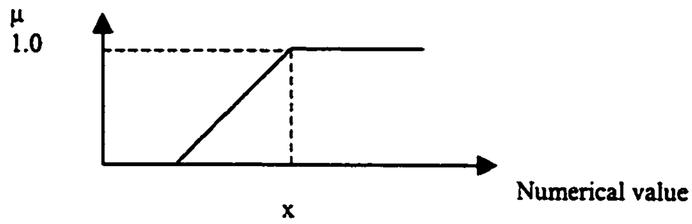


Figure 5-15: Membership Function for Large, Good or High

Rule 10: If the peak value x is given and there is a value y given, where y < x, and y is the lowest value on the x-axis, then let $\mu_y = F_y / F_x$, where F_y is the frequency of y and F_x is the frequency of x (refer to Figure 5-16).

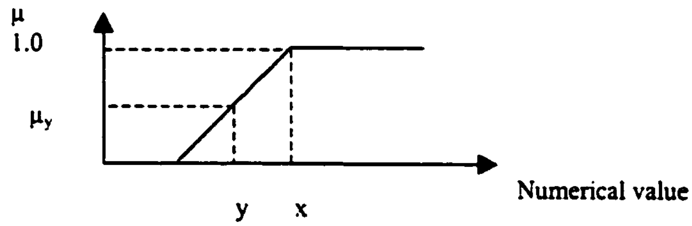


Figure 5-16: Membership Function for Large, Good or High using Rule 10

Rule 11: If the peak is the lowest or only value given, let $y_{large} = y_{average}$, where $\mu_{ylarge} = \mu_{yaverage} = 0.5$ (refer to Figure 5-17).

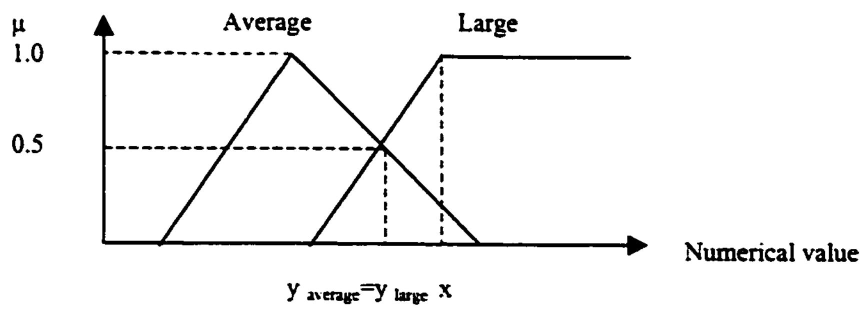


Figure 5-17: Membership Function for Large, Good or High using Rule 11.

Rule 12: If the peak value is given, but there is no reference to the membership function of Average, then let $\mu_y = 0.0$, $\mu_x = 1.0$, $y = x/2$. In this case, only the peak value is required to determine the shape of the membership function (refer to Figure 5-18).

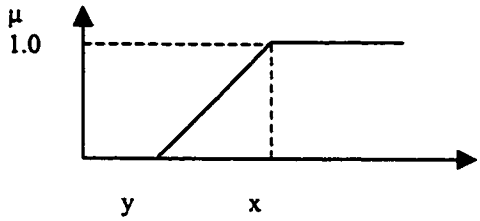


Figure 5-18: Membership Function for Large, Good or High using Rule 12

Rule 13: If there is no data at all for this group, but the relevant information of Average is available, then let $y_{large} = y_{average}$, where $\mu_{ylarge} = \mu_{yaverage} = 0.5$. $x_{large} = \frac{2}{3} y_{large}$ and $z_{large} = 2\ x_{large}$.

Another issue for generating membership functions for this group is the X-axis limit. If the X-axis has a natural limit (e.g., 10 for a numerical scale, 100 for percentage), then this is not an issue. If not, then a limit has to be decided for implementing the function in Matlab. In most cases, a sufficiently large number, such as 10 times the largest value given in the survey, is appropriate.

5.5 Examples of Membership Function Generation

According to the procedures described in the previous section, the membership functions for all of the factors were developed for each of the two training data sets. The complete results are shown in Appendix 4. The following examples illustrate how this method was implemented. Input Factor 1 is used as an example in both trials.

Example 1 (in trial 1: 13 training data, 5 testing data):

First, observe the graph. If there are too many overlaps, then some membership functions have to be combined. Second, apply the methodology to obtain the parameters of the membership functions.

The frequency graph for this factor (Figure 5-4) shows that Small and Average cannot be significantly distinguished. To eliminate too much overlap, these two membership functions are combined. The original and new frequency tables are shown in Table 5-2 and Table 5-3, respectively. The new frequency figure is then generated and shown in Figure 5-19.

Table 5-2: Original Frequency Table for Factor 1

Input #1	Frequency of Selection		
Scale	Small	Average	Large
1	1		
2	1		
3	1	1	
4			
5	1	5	
6			
7			
8			1
9			
10			

Table 5-3: Revised Frequency Table for Factor 1

Input #1	Frequency of Selection	
Scale	Small-Average	Large
1	1	
2	1	
3	2	
4		
5	6	
6		
7		
8		1
9		
10		

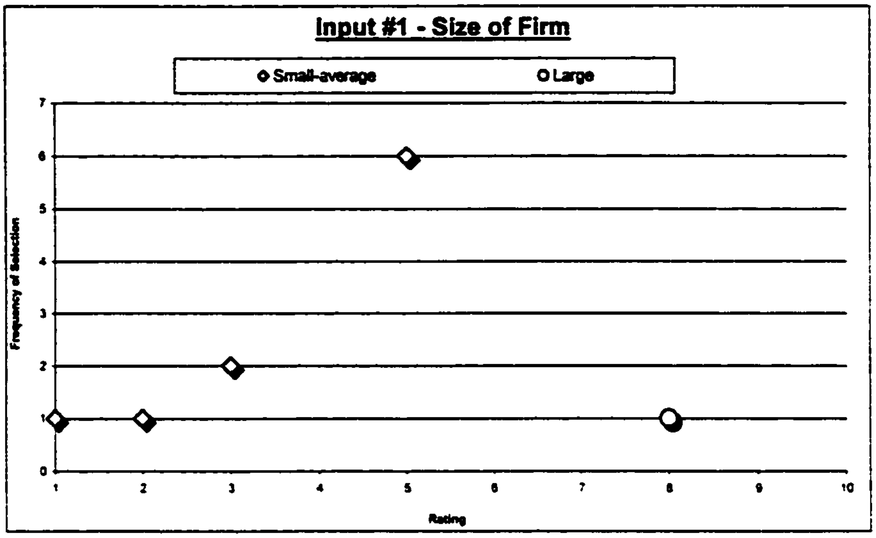


Figure 5-19: New Frequency Graph for Input Factor 1

Based on this new frequency figure, the membership functions for Small-Average and Large are developed, as shown in Figure 5-5.

The four parameters for the Small-Average membership function are (0, 0, a, b). Among them, parameter "a" equals 5, which obtains the highest frequency in the Small-Average section of the graph. There is no other data that is greater than 5 for the Small-Average function and also no reference to the adjacent membership function. This is because the membership function for Large also lacks the information needed to determine the left leg. Then Rule 3 applies. So another parameter b is equal to 2a, which in this case equals 10.

The four parameters for the large membership function are (a, b, c, d, which d is equal to c). Parameter b is equal to 8, which has the membership value of 1 because of its highest frequency in Large. However, no smaller value is given, which is smaller than 8, so rule 10 can not be used. Since the membership function for Small-Average has been developed in the previous step, it gives the large membership function a reference where Rule 11 applies, and $\mu_{ylarge} = \mu_{yaverage} = 0.5$. As a result, y equals 7.5. Then parameter "a" equals 7. This factor has an x-axis limit, which is 10, so parameter c is set to 10. Thus, the membership functions for this factor are as shown in Figure 5-5.

Example 2 (in trial 2: 11 training data, 7 testing data):

The procedures for generating membership functions are applied to the second set of training data (11 training cases, 7 testing cases). The three membership functions do not overlap (refer to Figure 5-20). Each of them can be developed using the proposed technique.

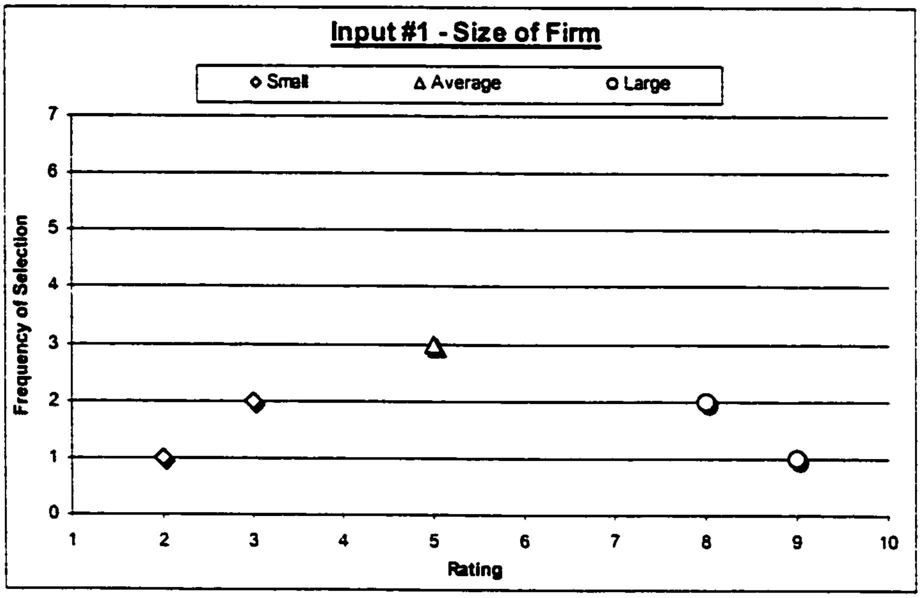


Figure 5-20: Membership Functions for Input Factor 1 in trial 2

For the Small membership function, the four parameters are 0, 0, c, d. 3 has the highest frequency in the graph for Small, so c is equal to 3. Since 3 is the largest value given and the membership function for Average has not been decided yet, only Rule 3 applies. According to Rule 3, parameter "d" equals 2c, which is 6.

For the Average membership function, 5 is the only value given. Assumption 4.2 a, indicates that if there is one distinct peak in the data set, then the shape of this membership function should be triangular. So the shape of this membership function is triangular. For triangular-shaped membership functions, the three parameters are a, b, and c. 5 is the only value given in the data, so b is equal to 5 and only Rule 6 can be used. The membership function for Small is available now, so let $y_{average} = y_{small}$, where $\mu_{ysmall} = \mu_{yaverage} = 0.5$. As a result, "a" is set to 4. As for the other leg of the Average membership function, one must refer to the membership function of Large. The Large membership function has not been decided yet, so the next step is to generate the membership function for Large.

The four parameters for the Large membership function are a, b, c, and d. The graph (Figure 5-20) shows that 8 has the highest frequency devoted to Large, so b is equal to 8. This factor has an x-axis limit, so c is set to this limit, which is 10. "a" is decided by Rule 12, which is 8/2=4.

Now according to Rule 7, let $Z_{average} = Z_{large}$, where $\mu_{zlarge} = \mu_{zaverage} = 0.5$, so parameter "c" in the Average membership function is 7.

The three membership functions for Input Factor 1 are thus developed and are shown in Figure 5-21.

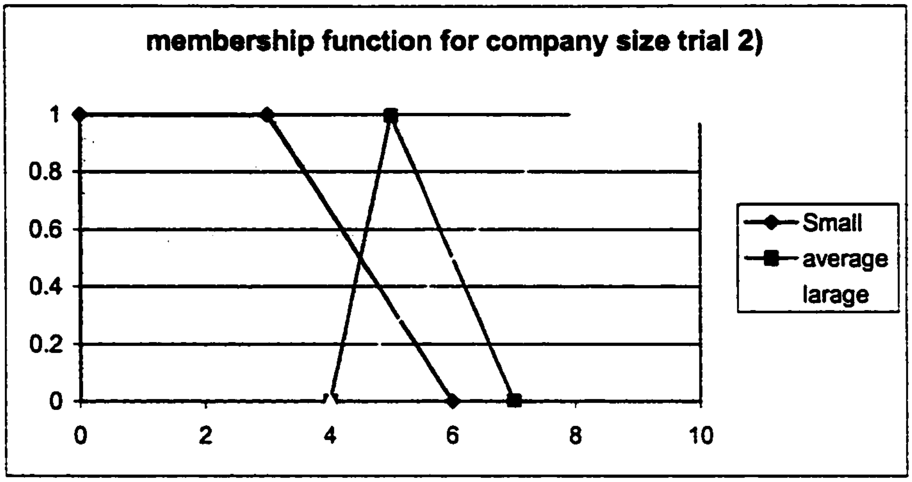


Figure 5-21: Frequency Graph for Input Factor 1 (trial 2)

The above two examples illustrate the technique developed to generate membership functions based on the frequency of responses. The proposed technique is very sensitive to the data being analyzed. Which case is better in terms of the accuracy of membership functions can not be decided until the testing procedure is executed. The testing procedure is presented in the next section.

5.6 Validation of Membership Functions

Validation of the membership functions was performed using the testing data set. Validation was performed twice, once each for trial 1 and trial 2. All the membership functions were developed from the training data. For each trial, the remaining testing data were used to test the membership functions. A small Visual Basic program (shown in Appendix 5) was developed to obtain the membership values from each membership function. The total testing results for the two trials are found in Appendix 6. Input Factor 1 was used as an example to illustrate the validation procedure.

In trial 1, the membership function for Input Factor 1 was developed, and the five testing data groups were respectively fitted to this membership function. The testing results are shown in Table 5-4.

Table 5-4: Membership Function Testing Results for Factor 1 in Trial 1

	Parameters	a	b	c	d
Membership function	Small-Average	0	0	5	10
	Large	7	8	10	10
			Membership Value		
	Real number	Linguistic term	Small-Average		Large
Group 1	3	Small	1		0
Group 2	5	Average	1		0
Group 3	8	Large	0.4		1
Group 4	3	Small	1		0
Group 5	9	Large	0.2		1

In trial 2, 7 data groups were consecutively fitted to the membership fʃnctions for Input Factor 1 using the functions developed for this factor using the training data set for trial 2. The results are shown in Table 5-5.

Table 5-5: Membership Function Testing Results for Factor 1 in Trial 2

	Parameters	a	b	c	d
Membership Function	Small	0	0	3	6
	Average	4	5	5	7
	Large	4	8	10	10
			Membership Vale		
	Real number	Linguistic term	Small	Average	large
Group 1	(N/A)				
Group 2	5	Average	0.33	1	0.25
Group 3	5	small	0.33	1	0.25
Group 4	1	small	1	0	0
Group 5	3	small	1	0	0
Group 6	3	average	1	0	0
Group 7	5	average	0.33	1	0.25

The criteria used for deciding a match between the testing sample and the membership function are if the matched linguistic term attains the highest membership value, and if this membership value is equal to or greater than 0.5. A successful membership function has to have at least 60% of its testing groups match. This percentage rate is chosen to be greater than a random rate. If a factor has 3 fuzzy sets, then the random success rate of this factor is 33% (i.e., 1/3). Similarly, a factor with only 2 fuzzy sets should have a 50% random success rate (i.e., ½). For the model to provide better results than a random choice, its success rate was set to 60%, which is almost double 33% and greater than 50%.

In trial 1, all of the testing groups for Input Factor 1 succeeded and the total accuracy of the membership function is 100%. Therefore, this membership function in trial 1 is acceptable (see Table 5-6).

Table 5-6: Matching Results for Input Factor 1 In Trial 1

	Actual number	Linguistic term	Membership Value		Highest value	Match Y/N	Total accuracy (%)
			Small-Average	Large			
Group 1	3	Small	1	0	1	Y	100
Group 2	5	Average	1	0	1	Y	
Group 3	8	Large	0.4	1	1	Y	
Group 4	3	Small	1	0	1	Y	
Group 5	9	Large	0.2	1	1	Y	

In Table 5-7, 2 predicted linguistic terms failed to match the actual ones, resulting in a total accuracy of 4/6*100=67%. This result is still higher than 60%, so the membership function in trial 2 is still acceptable.

Table 5-7: Matching Results for Input Factor 1 In Trial 2

	Actual number	Linguistic term	Membership Value			Highest value	Match Y/N	Total accuracy (%)
			Small	Average	Large			
Group 1	(N/A)							67
Group 2	5	Average	0.33	1	0.25	1	Y	
Group 3	5	Small	0.33	1	0.25	1	N	
Group 4	1	Small	1	0	0	1	Y	
Group 5	3	Small	1	0	0	1	Y	
Group 6	3	Average	1	0	0	1	N	
Group 7	5	Average	0.33	1	0.25	1	Y	

After generating and testing all of the membership functions, a number of problems were encountered and addressed.

First, in some cases, the respondents gave two linguistic terms. For example, they may have circled both small and average for a factor with three membership functions (small, average, large). What is considered a match is the answer matching either of the two linguistic terms while still meeting the general criteria for a match (membership value ≥ 0.5 and highest membership value)

Second, if the respondents circled two linguistic terms for a factor with only two membership functions, then there is a match if the answer tends towards lower (or upper) values and still meets all the general criteria. For example, if the factor has two membership functions, say small and average-large, and the response is small-average, then it matches the linguistic term "small" if this answer has the highest membership value and it is greater than or equal to 0.5. If the factor has two membership functions, say small-average and large, and the respondents gave an answer of "average-large", then there is a match on "large" if the answer meets the general criteria, which means this answer obtains the highest membership value and also this value is greater than 0.5.

Third, sometimes there is a tie in the membership values because of an overlap in the membership functions. Then there is a match as long as the values meet all the general criteria. Due to the limited data, this can not be avoided.

Fourth, if close membership values were found, and the difference between these two is less than or equal to 0.05, then these two values become approximately equal. Thus, either one is matched, and the matching rate is potentially increased.

The final testing results are shown in Appendix 7. In both cases, 17 models of membership functions out of 89 failed. The overall accuracy of this method is 81%, which is an acceptable accuracy rate. However, there are 5 models of membership functions that failed in both trials. The main reason that led to the failure of these membership functions is lack of data. If a large enough sample size can be obtained, these membership functions can be improved and may yield better results. Another way to increase the accuracy of the membership functions is to reduce the number of fuzzy sets, for example, instead of 3 linguistic categories, choosing 2 linguistic categories. Thus, the accuracy of the model can be improved, however, the distinction of the categories will be reduced accordingly.

These failed membership functions were not fine-tuned in this research. A simplification procedure is performed before implementing the functions in the fuzzy expert rules. After simplification, some membership functions are not used in the rules. This simplification process is described in Chapter 6.

5.6 Summary

In this chapter, a technique for generating membership functions based on objective numerical data was presented. The data collected from the survey was segregated into training and testing data sets. The membership functions were generated using the proposed technique and tested twice. The results of the testing are satisfactory, indicating the feasibility of the proposed technique. Much of the inaccuracy lies in the lack of sufficient data with which to generate and test the membership functions. Consequently, some factors are modeled with overlapping and /or combined membership functions. This lack of distinction in the membership functions can be improved as more data is obtained.

Nevertheless, the proposed technique is significant in that it illustrates how membership functions can be developed on the basis of objective (i.e., numerical) data. If sufficient data can be collected for a given context, then these membership functions can reflect a widely–held concept of a subjective (i.e., linguistic) term. If the context variables that affect the shapes and ranges of the membership functions can be identified, then these membership functions can be calibrated to suit different contexts. The technique presented in this chapter provides a foundation for a method of generating and calibrating membership functions that are widely applicable and have the same meaning to different users in a given context.

The proposed model for design performance prediction and evaluation is very complex. Implementation in the form of a fuzzy expert system is difficult due to the number of possible combinations of factors and therefore rules. Correlation analysis is used to reduce the number of factors in the model. Simplification of the model using correlation analysis is described in the next chapter.

CHAPTER 6 MODEL SIMPLIFICATION – USING CORRELATION ANALYSIS

6.1 Problem Description

As shown in Chapter 3, the original model is complex: it consists of 14 higher-level input factors and 3 higher-level output factors. Among them, 16 higher-level factors have sub-factors related to them. These sub-models consist of various numbers of sub-factors ranging from 2 to 12. Given that each factor and sub-factor can take on any one of three linguistic values, this complex structure leads to an unmanageably large number of possible combinations of different levels of each factor and sub-factor. This complex structure leads to the following disadvantages that greatly limit research efficiency and accuracy:

- It significantly increases the number of possible combinations of expert rules, increasing the specification and calculation load by powers of the number of factors (e.g., 3^5 for 5 factors)
- No existing software or method can be effectively used to specify all possible rules.
- Even though the model is comprehensive, it does not mean that all of the information it provides is potentially useful and significant to the output. Its complexity may make it difficult for the decision -maker to pinpoint the main factors influencing positive or negative design performance.

To yield a more realistic and manageable model, reduction of the number of sub-factors and factors was required by eliminating insignificant ones. Correlation analysis was performed for each sub-model on its input and output to reduce the number of variables in each sub-model.

6.2 Correlation Method

6.2.1 Why correlation analysis?

Correlation analysis was chosen to simplify the model for the following reasons:

- One of the basic assumptions of the expert rules is that all of the input factors have independent and equal influences on the output. This feature makes correlation analysis suitable for the simplification purpose.
- The goal of simplifying the model is to retain factors that have significant effects on the output and eliminate any unrelated factors. Many methods for data reduction can be used, such as principle component analysis, stepwise linear regression analysis, etc. However, due to the lack of data the stepwise linear regression method can not be used. On the other hand, the results from principle component analysis are not what the author expected. The author already generated membership functions for each factor in Chapter 5, but this method combines certain factors into one general factor. Thus, all the generated membership functions will become void. As a result, principle component analysis is also not suitable. Correlation analysis, then, becomes available and feasible in this case. The correlation coefficient will directly tell the author whether the input factors contribute to the output or not. Those factors that are highly correlated to the output will be kept and those that are not will be eliminated. Thus the model is simplified.
- Correlation analysis not only presents the ability to identify input to output relationships, but also shows the direction of the relationship. That is, the input is either positively or negatively correlated to the output.

6.2.2 Introduction to correlation analysis

Correlation measures the linear relationship between two quantitative variables. In other words, correlation analysis will tell us whether the input values and the output values have a linear relationship or not. The commonly used measure for correlation analysis is called the Pearson Correlation Coefficient, denoted by γ. It is defined as (Norusis, 1993)

$$\gamma = \frac{\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})}{(N-1)S_x S_y} \tag{6-1}$$

where N is the number of cases and S_x and S_y are the standard deviations of the two variables. The absolute value of γ indicates the strength of the linear relationship. The correlation coefficient ranges in value from -1 to +1. A value of 0 indicates no linear relationship between the two variables, while a value of +1 indicates a perfect positive relationship (as the values of one variable increase, the values of the other also increase). A value of -1 indicates a perfect negative relationship (as the values of one variable increase, the values of the other decrease).

6.2.3 Correlation analysis application

Correlation analysis was used in three categories of relationships.

- One category consists of the sub-input factors to the higher-level input factor. This category has 13 models altogether, since one factor (input factor 14) does not have sub factors.
- One category includes all higher-level input factors to each sub-output factor. Here we have 13 models as well.
- The third category is sub output factors to higher level output factors. There are 3 models in this category, since we have only 3 higher-level output factors.

Special considerations when applying this method include:

- The numerical data in the actual survey was used for the correlation analysis.
- A total of 18 actual projects were used for this analysis.
- Since the direction of the relationship can not be determined in advance, a two-tailed significance test was used.
- Due to the small sample data size (only 18 in total), a significance level of 0.10 was applied, since the correlation analysis is used as a rough guide to the relationship between factors.

6.3 Results of Correlation Analysis from SPSS

The tool used for correlation analysis is called SPSS (Statistical Package for Social Sciences) for Microsoft Windows 7.5 (SPSS Inc., 1996). SPSS is a very comprehensive, flexible, and powerful statistical analysis and data management system.

An example (Input factor 1) will be used to illustrate how the model was simplified using correlation analysis step by step.

- The original model

Sub input model 1:

Input 1.1: # of employees
Input 1.2 total volume of work
Input 1.3 # of projects on hand

→ Input 1: overall size of the company

- Input all the values for each factor into SPSS (refer to Table 6-1).

Table 6-1: Data Input Format in SPSS

1	3	2	0.3	4
2	5	900	10	10
3	8	60	10	5
4	3	12	0.11	4
5	9	1000	.	100
6	.	12	.	7
7	2	7	7.5	7
8	5	15	0.25	60
9	5	100	1	100
10	2	13	0.3	15
11	2	4	0.1	2
12	3	20	0.2	5
13	5	145	1.75	5
14	3	6	.	5
15	8	650	135	6
16	5	10	1	5
17	1	4	.	25
18	5	40	1.5	10

- Run correlation analysis function, then the output was obtained and shown in Table 6-2.

Table 6-2: Output format from SPSS for correlation results

Correlations					
		INPUT1	INPUT11	INPUT12	INPUT13
Pearson Correlation	INPUT1	1	0.645004	0.565173	0.384072
	INPUT11	**0.645004**	1	0.579487	0.364056
	INPUT12	**0.565173**	0.579487	1	-0.13199
	INPUT13	0.384072	0.364056	-0.13199	1
Sig. (2-tailed)	INPUT1	.	0.005178	0.035196	0.128003
	INPUT11	**0.005178**	.	0.029867	0.137492
	INPUT12	**0.035196**	0.029867	.	0.652871
	INPUT13	0.128003	0.137492	0.652871	.
N	INPUT1	17	17	14	17
	INPUT11	17	18	14	18
	INPUT12	14	14	14	14
	INPUT13	17	18	14	18
**	Correlation is significant at the 0.01 level (2-tailed).				
*	Correlation is significant at the 0.05 level (2-tailed).				

- Taking 0.1 for the significance level, some factors will be eliminated. Table 6-2 shows that the correlation significance test results of input 1.1 and input 1.2 are both lower than 0.1. It shows that these two sub factors have very strong relationships to input factor 1, while input 1.3 will fail this test (<90% confidence level) if the sample size large enough. So input 1.1 and input 1.2 were kept for further development. Input 3.1 was eliminated.
- The simplified model.

  The simplified model is shown as follows:

Sub input model 1:

Input 1.1 # of employees
Input 1.2 total volume of work

Input 1: overall size of the company

6.4 Simplified Model

The results of this analysis of the three categories of factors are shown in Appendix 8 (factors with significant relationship to the output have been highlighted).

From these results, some sub models were eliminated completely because there exists no strong relationship between any of the inputs to the output. These sub models are for input factor 7, input factor 11, input factor 13, higher-level input factors to sub output factor 3.3 model and sub output factor 3.4 model. The output sub factors did not show very strong relationships to each higher-level output factors. So all these sub factors can not be used to describe the design performance. As a result, all 3 of these models were eliminated as well.

After the model was simplified, the framework for the fuzzy expert rules of each model could be more easily built. Since we assume all the input factors independently and equally affect the output, the "and" operator should be used in these expert (If-Then) rules.

The simplified models are shown as follows:

Sub input model 1:

Input 1.1 # of employees
Input 1.2 total volume of work → Input 1: overall size of the company

Sub input model 2:

Input 2.2: # of projects available in the market → Input 2: competition level

Sub input model 3:

Input 3.8: # of design personnel changes → Input 3: design firm profile

Sub input model 4:

Input 4.2: design process duration → Input 4: size of design contract

Sub input model 5:

Input 5.1: manhour per week for each designer → Input 5: continuity of design

Sub input model 6:

Input 6.1: clarity of design definition
Input 6.2: clarity of design alternative definition
→ Input 6: scope definition level

Sub input model 8:

Input 8.4: number of review authorities involved → Input 8: design process complexity

Sub input model 9:

Input 9.3: effect in-situ soil conditions → Input 9: project condition complexity

Sub input model 10:

Input 10.1: duration of the owner to make a decision → Input 10: quality of owner's profile

Sub input model 12:

Input 12.1: number of construction packages
Input 12.2: percentage of non-standard forms and conditions
→ Input 12: complexity of tendering process for construction

Higher level input factors to sub output factor 1.1:

Input 6: scope definition level → Sub output 1.1:Actual manhours over budgeted ones

Higher level input factors to sub output factor 1.2:

Input 7: project function complexity
Input 9: project condition complexity
→ Sub output 1.2:percentage of design manhours due to owner changes

Higher level input factors to sub output factor 1.3:

Input 11: quality of vendor's profile

→ Sub output 1.3: percentage of design manhours due to rework

Higher level input factors to sub output factor 1.4:

Input 6: scope definition level
Input 8: design process complexity
Input 14: economic conditions

→ Sub output 1.4: actual design cost over budgeted

Higher level input factors to sub output factor 1.5:

Input 12: complexity of tendering process for construction

→ Sub output 1.5: design cost over the cost due to the owner changes

Higher level input factors to sub output factor 1.6:

Input 5: continuity of design
Input 7: project function complexity
Input 11: quality of vendor's profile
Input 12: complexity of tendering process for construction

→ Sub output 1.6: design cost over total construction cost

Higher level input factors to sub output factor 2.1:

Input 1: overall size of the company
Input 11: quality of vendor's profile

→ Sub output 2.1: actual design duration over scheduled

Higher level input factors to sub output factor 2.2:

Input 3: design firm profile
Input 10: quality of owner's profile

→ Sub output 2.2: percentage of the miss-deadline design document release

Higher level input factors to sub output factor 3.1:

Input 4: size of design contract
Input 13: Complexity of construction process

→ Sub output 3.1: number of approved changes

Higher level input factors to sub output factor 3.2:

Input 4: size of design contract
Input 12: complexity of tendering process for construction

Sub output 3.2: total cost of approved changes

Higher level input factors to sub output factor 3.5:

Input 1: overall size of the company

Sub output 3.5: number of problems due to design errors during the construction.

6.5 Summary

Based on this simplified model structure, the expert rules describing the relationship between input and output factors for each model are developed. The simplification process yields a smaller number of possible rules, leading to a more manageable rulebase. The rule specification and testing process is described in Chapter 7.

CHAPTER 7 GENERATING IF-THEN RULES FOR FUZZY EXPERT SYSTEM

7.1 Introduction

7.1.1 Introduction to Fuzzy Expert System

To understand the If-Then rules in a fuzzy expert system, one has to start from the fuzzy expert system. A fuzzy expert system is basically an expert system that uses a group of membership functions and If-Then rules for reasoning. The If-Then rules in a fuzzy expert system commonly look similar to the following:

If x is small and y is large, then z is average

where x and y are input variables and z is the output. Small is a membership function defined on x, large is a membership function defined on y and average is another membership function defined on z. The rule's premise gives the degree to which the rule applies, while the rule's consequent assigns a membership function to each of the output variables. The collection of these If-Then rules in a fuzzy expert system is called the rulebase, which is the core of the fuzzy expert system

7.1.2 Introduction to Fuzzy Inference Process

No matter what type of fuzzy inference system, the processing procedures are virtually the same. They all include the following steps:

- Fuzzification: applied membership functions defined on a certain input variable to an actual value to determine the degree of truth for the rule premise.
- Fuzzy operator: if the antecedent contains more than one input variable, the calculated truth value for each input variable is then combined, using a certain fuzzy

operator to obtain one single number that represents the result of this rule's premise. The commonly used fuzzy operators are "AND" and "OR." For the operator "AND", there are two common operation methods. One is "MIN" operator, which takes the minimum of the membership values. The other is the "PRODUCT", which is the product of all the membership values.

- Implication: the single value obtained from each rule premise is applied to the consequent of each rule. Two common implication methods are "MIN" and "PRODUCT". They either truncate (MIN) or scale (PRODUCT) the membership function of each consequent of each rule, using the single value given by the antecedent.
- Aggregation: a process in which all the fuzzy outputs of each rule are combined into a single fuzzy set, which happens only once for each output variable. The commonly used aggregation methods are "MAX" and "PROBOR". "MAX" method involves taking the maximum value from each output. The algorithm for "PROBOR" is the algebraic sum of the entire output from each fired rule.
- Defuzzification (optional): used to convert the fuzzy output obtained from the aggregation step into a single, crisp value. There are many methods for defuzzification, among which the "CENTROID" and "MAXIMUM" methods are commonly used. The crisp value calculated from "CENTROID" method is the center of gravity of the membership function for the output. "MAXIMUM" method involves LOM (the largest crisp value of the maximum) method, MOM (mean of the maximum value), SOM (the smallest crisp value of maximum) method.

7.2 Description of the Method of Generating If-Then Rules

The great limitation when choosing a method to generate If-Then rules results from the limited data collected. Without sufficient data, none of the methods described in Chapter 2 could be properly used. Based on all the available information, the author proposes a new technique for generating the If-Then rules for the fuzzy expert system, which is tested using real values.

The key features of this technique are as follows:.

- The second set of membership functions was chosen to build the fuzzy expert system, since it yielded the best results. This data set contains 11 training data and 7 testing data. The fuzzy rulebase was also constructed using the same 11 training data and was tested with the 7 testing data.
- The results from correlation analysis were used to determine the logical relationship of the rule premise and rule consequent.
- Actual values maintain the highest priority in the building of the fuzzy expert system. If there is a conflict with an actual value when implementing the method, the actual value is kept. The assumption of this method is that all input factors are independent and have an equal effect on the output.
- This technique carefully considers the completeness and consistency of the rulebase. If all the possible rules can be collected, different weighting of the rules can be used to distinguish the more significant rules and the other rules.

The detailed procedures for this technique are as follows:

- The first step is to count how many possible rules there are to maintain the completeness of the rulebase. This depends on the number of input variables and also on the number of membership functions that each variable has. For example, if a model has 2 input variables and one variable has 3 membership functions and the other has 2, the complete rulebase should have 3*2=6 rules altogether. Another example, is a model has 3 input variables and each variable has 3 membership functions, then the complete rulebase will have 3^3 =27 rules altogether (i.e., twenty-seven possible combinations).

- All the actual rules that occur in each model are noted, and the frequency of each rule in that model is recorded.
- For each model, the frequency of the rule indicates its relative likelihood of occurrence. If one rule obtains the highest frequency and its output makes sense, then this rule is kept in the rulebase. If a rule has the highest frequency in the actual rule system, but its output is not correct based on the correlation results, then this rule cannot be kept. For example, if the correlation analysis says that input 1 and input 2 are positively correlated to the output, and if the actual rule says:

  "If input 1 is small and input 2 is small, then output is large."

  Then in that case, this rule should be ignored. If it says:

  "If input 1 is small and input 2 is small, then output is small."

  Then this rule should be kept in the rulebase.
- In order to maintain the consistency of the rulebase, the same premise can not have two or more different conclusions. The inconsistent rules can also be eliminated based on the frequency of the rules and based on the results of the correlation analysis.

- Redefine the rules based on the number of membership functions, after combining overlapping functions.
- All the actual and reasonable rules are kept in the rulebase. The rest of the rules to cover all other combinations of input are derived using the patterns in the existing rules and using the results of the correlation analysis.
- The technique used for deriving rules is based on the actual frequency of the responses and the results of the correlation analysis. Assuming that all the input factors are independent and have an equal effect on the output factor, all combinations in any order will yield the same result. For example, suppose that the actual rule says: if input 1 is small and input 2 is average, then output is small; then, in that case, the derived rule will be: if input 1 is average and input 2 is small, then output is small. Another consideration for generating the derived rules is to use the results of the correlation analysis to determine the pattern. For example, if the actual rule says:

  If input 1 is <u>small</u> and input 2 is large, then output is large. And all the inputs are positively related to the output.

  Then the derived rules will be:

  If input 1 is <u>average</u> and input 2 is large, then output is large.

  If input 1 is <u>large</u> and input 2 is large, then output is large.

  This procedure is repeated until all the remaining rules are derived. Thus, a complete and consistent rulebase can be built for each model.
- For missing data in the rules, the following steps are taken. If only one variable has input or output missing, the missing value is replaced with all possible

linguistic terms. Keep the variables that already exist in the actual rules and increase the frequency of those rules accordingly by 1. If more than two variables have missing data, then ignore this data record.

- Repeat the entire procedure for each model.

7.3 Application of the Method

The above technique has been applied to the fuzzy expert system for design performance evaluation. Appendix 9 shows the completed rulebase for each model in the problem. An example was chosen to illustrate this technique, which can be described as follows:

- Take Input Factor 1, for example. The total number of input variables is 2. Input 1 has 2 membership functions, which are small and average-large. Input 2 also has two membership functions, which are small and average-large. So the total number of possible rules is 2*2=4. So the framework of the model is:

  If sub input 1.1 is (small, average- large), and sub input 1.2 is (small, average-large), then output (input 1) is (small, average, large).

- Count all the actual rules. The results are shown in Table 7-1:

Table 7-1: Rules Derived from Actual Data for Input Factor 1

Rule number	Sub input 1.1 (positive correlation)	Sub input 1.2 (positive correlation)	Input 1	Frequency	Check consistency
1	Small	Small	Small	3	Keep
2	Average	Average	Average	2	Keep
3	Average	Small	Average	1	Keep
4	Small	Small	Average	1	Eliminate
5	Large	Large	Large	1	Keep
6	Average	Average	Large	1	Eliminate

- Check the consistency of the rules. Rule 1 and rule 4 conflict. According to correlation analysis and frequency, keep rule 1. Apply the same procedure to rule 2 and rule 6, eliminate rule 6.
- Based on the number of the membership functions, some rules are combined to reflect the combined membership functions, such as Small-Average. After combination, the following rules are kept (refer to Table 7-2):

Table 7-2: Actual Remaining Rules

Rule number	Sub input 1.1 (positive correlation)	Sub input 1.2 (positive correlation)	Input 1
1	Small	Small	Small
2	Average-large	Average-large	Large

- Based on the total rule number of possible rules (combinations), that is 4 in this case, 2 more rules were derived. According to actual rule 3, the two rules can be derived as follows (refer to Table7-3):

Table 7-3: Derived Rules

Rule number	Sub input 1.1 (positive correlation)	Sub input 1.2 (positive correlation)	Input 1
3	Small	Average-large	Average
4	Average-large	Small	Average

- The rulebase for model 1 is now complete.

7.4 Building the RuleBase In Matlab

The completed rulebase is implemented in Matlab, (Mathworks Inc., 1999). Its Fuzzy Logic Toolbox is used to build the rulebase.

The steps can be shown in a sequence of graphs.

1. Open fuzzy toolbox in Matlab and input all the information about the model, such as number of inputs and outputs, and name of the inputs and outputs (Figure 7-1).
2. Input membership function specification for each variable (Figure 7-2).
3. Input all the rules using the Rule Editor (Figure 7-3).
4. Examining the performance of the model (Figure 7-4).

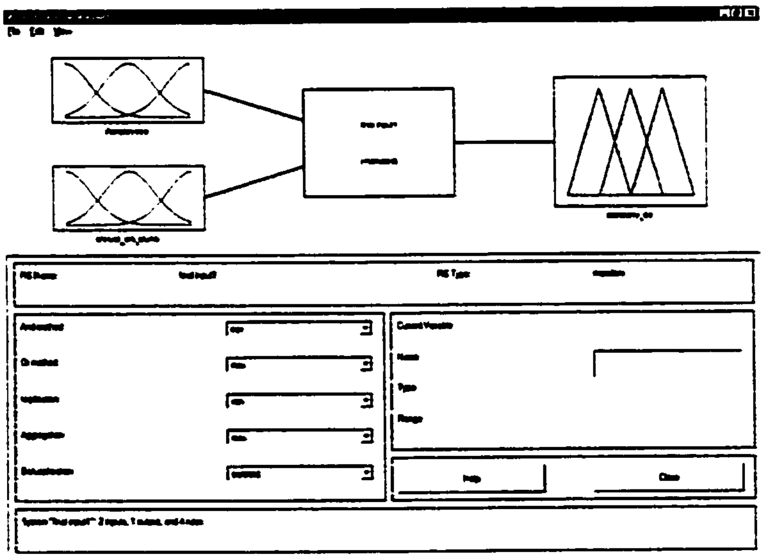


Figure 7-1: Input and Output Variables in Matlab

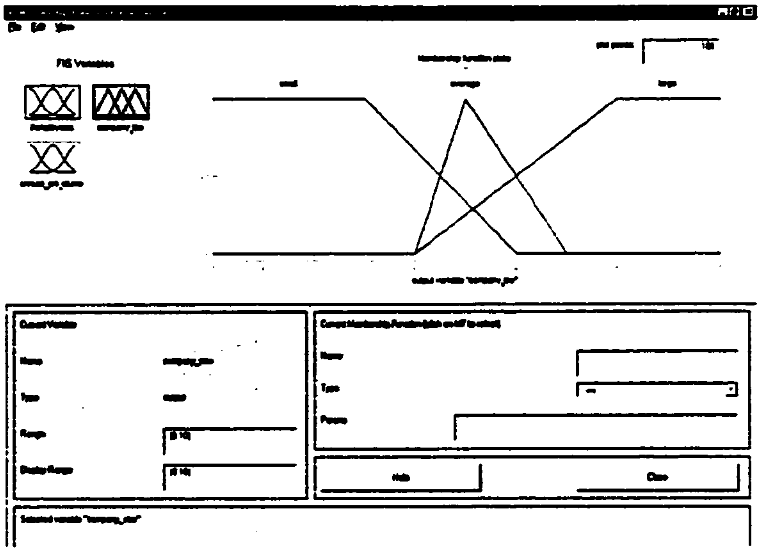


Figure 7-2: Membership Functions for Each Variable in Matlab

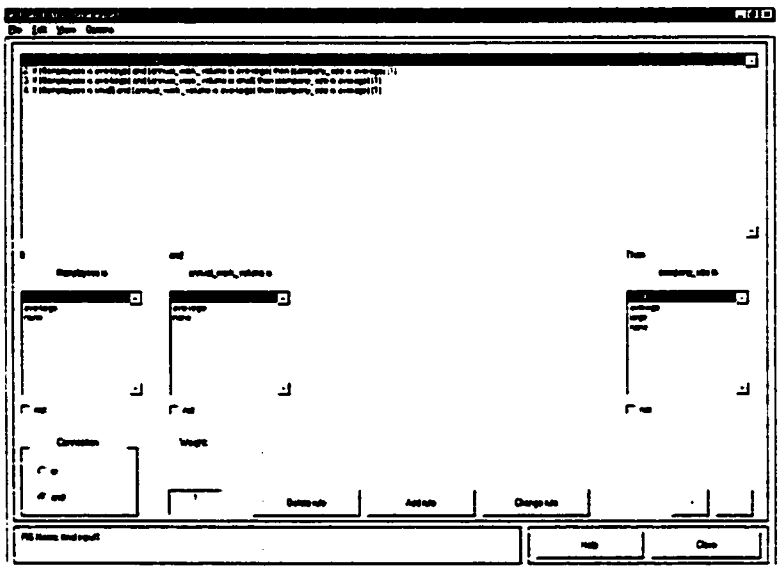


Figure 7-3: Rulebase in Rule Editor of Matlab

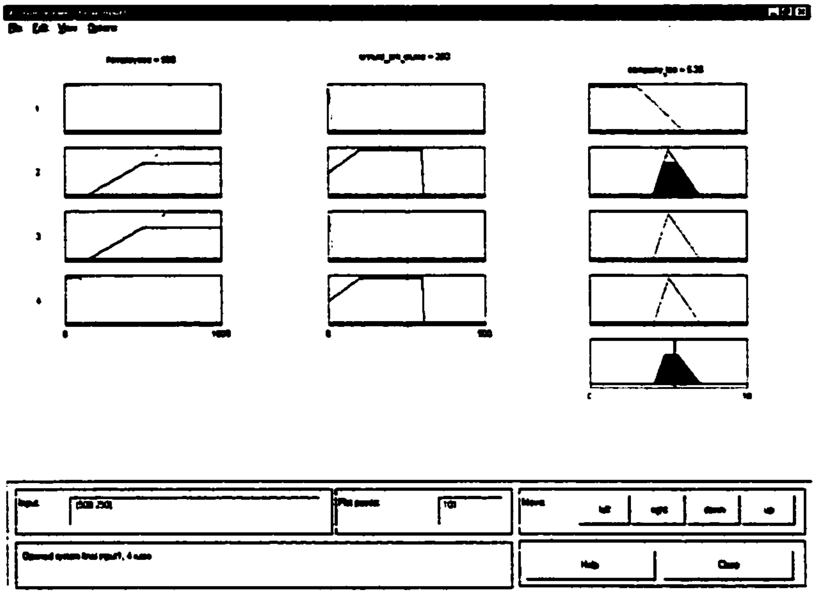


Figure 7-4: A Complete Model in Matlab

7.5 Model Validation

Model validation involves applying the remaining 7 responses not used for training to test the accuracy of the rulebase. The base case testing was done using the "and" operator with the "min" operation, the min-max method for rule implication-aggregation, and the centroid method for defuzzification. For each model, the 7 survey data were input to the model respectively. The output of each test was a crisp number after defuzzification and is shown on a defuzzified graph. Whether the model is successful or not depends on two criteria: one is how close the defuzzified crisp number is to the actual value that the respondents gave, and the other is whether it is a linguistic match with the linguistic term given.

Numerically, the match is defined on the basis of the calculated error, which is represented by:

Error = [|(predicted value-actual value)|/actual value]*100 (7-1)

If this error is less than 33%, then the rule is considered successful.

For a model to be successful, at least 50% of its testing cases should succeed, based on the numerical match criteria.

Linguistically, a match is defined in terms of whether the linguistic area that the defuzzified crisp value belongs to is the same as the actual linguistic term. If it is, then this is considered a match.

In some cases, the predicted linguistic term is average-large, while the actual answer is small-average. Then the author would consider it a match at "average." This is mainly because some of the membership functions have a large overlap with the two membership functions. This overlap is due to the lack of sufficient data to distinguish the membership functions.

The overall results of the base case testing are shown in Table 7-4. Only 38% of the total number of models were numerically successful. The linguistic testing produced good results, with a success rate of 71% of the models. The current fuzzy expert system model does not achieve a high success rate for numerical prediction: there is a large bias between the actual value and the predicted crisp value. Because of the roughness of the membership functions, the fuzzy expert system can not conclude an accurate enough crisp number. In the design context, however, linguistic terms are more commonly used than crisp numbers to describe the dynamics of design project performance. Consequently, the performance of the model is acceptable on the basis of linguistic term prediction.

Table 7-4: Overall Testing Results for Base Case

	Numerically Matched	Linguistic Term Matched
Total number of models	21	21
Number of Matching Models	8	15
Success rate (%)	38	71

7.6 Model Sensitivity Analysis

Sensitivity analysis is used to determine which variables affect the output of the model to the greatest degree. The results of the sensitivity analysis provide a way to increase the model's accuracy. In this research, the sensitivity analysis was conducted by changing the defuzzification method, the implication method, and the operator method.

7.6.1 LOM Defuzzification Method

The LOM (largest of the maximum) defuzzification method takes the largest crisp value that attains the largest membership value. Matlab takes the maximum of these x-axis values. Changing the defuzzification method to this one and repeating the above procedure, led to the following results: 57% of the models matched numerically and 81% matched linguistically. The complete results are shown in Appendix 10.

7.6.2 SOM Defuzzification Method

The SOM defuzzification method takes the minimum crisp value that attains the highest membership value. 29% of the models matched numerically and 71% matched linguistically.

7.6.3 MOM Defuzzification Method

Similar to LOM and SOM defuzzification method, MOM simply takes the average crisp value that attains the highest membership value. The results of this method turns out to be 43% numerically matched and 71% linguistically matched.

7.6.4 Product-Probor Implication-Aggregation Method

The implication-aggregation method used for the base case testing is the "min-max" method. The implication-aggregation method is changed to the Product-Probor method, which does not change the accuracy of the model.

7.6.5 Product for "and" operation

The "min" is usually used in "and" inference rules. The "min" operator is changed to "product" to examine the impact of different operators on the output. The results do not change from the base case.

7.6.6 Conclusions from Sensitivity Analysis

The comparison results for the sensitivity analysis are shown in Table 7-5. From this table, LOM defuzzification is the best method for increasing the accuracy of the model. Other defuzzification methods changed the model accuracy to different degrees. Changing the implication method or the operator does not improve the model's accuracy. Based on this analysis, the method of defuzzification is the most significant factor in improving the accuracy of the model. When designing the rulebase system, one has to pay close attention to this parameter. Table 7-5 also shows that no matter what method is used, the results for numerical data prediction are not very satisfactory. On the other hand, all the linguistic term testing results are good, especially for the LOM defuzzification method. How numerical testing results can be improved will be left for future research. However, linguistic term testing is satisfactory.

Table 7-5: Sensitivity Analysis Results Comparison

	Methods Applied					
	Base case	LOM	SOM	MOM	Prod-Probor	"and"- product
Accuracy (%)	38	57	29	43	38	38
Linguistic Match (%)	71	81	71	71	71	71

For those unmatched linguistic terms, a further investigation can be done by determining how far off the predicted linguistic terms are from the actual. To do this, a matrix was developed to show the distribution of the errors (refer to Table 7-6). Table 7-7 shows an example of Factor 1 in the base case.

Table 7-6: Error Distribution Matrix

		Actual Linguistic Term		
		Small	**average**	**large**
Predicted Linguistic Term	Small	Match	1 term off	2 term off
	Average	1 term off	Match	1 term off
	Large	2 term off	1 term off	Match

All the numbers on the diagonal line indicate the number of matches. If there are 3 linguistic terms, and, for example, the actual term is "Average" and the predicted term is "Small", then the predicted term is 1 term off from the actual term. Two terms off means the actual term is "Large", while the predicted one is small or vice versa. This relationship is shown in Table 7-7. The complete results for all the factors and for both cases are shown in Appendix 11.

Table 7-7: Error Distribution Matrix for Input Factor 1

		Actual Linguistic Term		
		Small	**Average**	**large**
Predicted Linguistic Term	Small	**2**	1	
	Average		1	
	Large			

Table 7-8 shows that in the base case, most (65%) of the linguistic terms are a 100% match, while 21% are 1 term matched, and only 14% are two-term matched. Among those 1-term matching models, 78% of the predicted linguistic terms are larger than the actual terms. If the percentage of complete matches is less than 50%, then the model is considered a failure. Altogether, 5 models failed in this case, which are shown as shaded in Table 7-8.

Table 7-8: Detailed Results for Base Case Testing

	# of responses	# match	% match	# 1 term	% 1 term match	larger #	# 2 term	% 2 term match	larger #
Model 1	4	3	75	1	25	0	0	0	
Model 2	4	1	25	3	75	1	0	0	
Model 3	7	4	57	1	14		2	29	2
Model 4	7	4	57	1	14	1	2	29	2
Model 5	7	6	86	1	14		0	0	
Model 6	7	5	71	2	29	2	0	0	
Model 7	7	5	71	2	29	2	0	0	
Model 8	5	4	80	1	20	1	0	0	
Model 9	7	3	43	3	43	2	1	14	
Model 10	7	6	86	1	14	1	0	0	
Model 11	7	7	100	0	0		0	0	
Model 12	5	5	100	0	0		0	0	
Model 13	7	5	71	0	0		2	29	2
Model 14	4	3	75	1	25	1	0	0	
Model 15	6	2	33	2	33	2	2	33	
Model 16	7	1	14	4	57	4	2	29	2
Model 17	5	3	60	1	20	1	1	20	1
Model 18	6	2	33	3	50	3	1	17	1
Model 19	7	6	86	0	0		1	14	1
Model 20	6	5	83	0	0		1	17	1
Model 21	5	3	60	0	0		2	40	2
Total	**127**	**83**	**65**	**27**	**21**	**21**	**17**	**13**	**14**
					% larger	**78**		**% larger**	**82**
	Note:	Fail (<50)	5						

Table 7-9 shows the results for the LOM method of defuzzification. A higher percentage, namely 69%, obtained 100% matching terms. 1-term matches were 20% and two-term matches were 11%. Only 3 models failed the linguistic testing in this case (refer to the shaded models in Table 7-9).

Table 7-9: Detailed Testing Results for LOM Defuzzification Method

	# of responses	# match	% match	# 1 term	% 1 term match	# larger	# 2 term	% 2 term match	# larger
Model 1	4	3	75	1	25		0	0	
Model 2	4	1	25	2	50	1	1	25	
Model 3	7	4	57	1	14		2	29	2
Model 4	7	5	71	1	14	1	1	14	
Model 5	7	5	71	2	29	2	0	0	
Model 6	7	4	57	2	29	2	1	14	
Model 7	7	5	71	2	29	2	0	0	
Model 8	5	4	80	1	20	1	0	0	
Model 9	7	3	43	2	29	1	2	29	1
Model 10	7	6	86	1	14	1	0	0	
Model 11	7	7	100	0	0		0	0	
Model 12	5	5	100	0	0		0	0	
Model 13	7	6	86	0	0		1	14	1
Model 14	4	3	75	0	0		1	25	
Model 15	6	3	50	2	33	2	1	17	
Model 16	7	3	43	4	57	4	0	0	
Model 17	5	3	60	1	20	1	1	20	1
Model 18	6	5	83	1	17	1	0	0	
Model 19	7	4	57	1	14		2	29	1
Model 20	6	4	67	1	17		1	17	1
Model 21	5	4	80	0	0		1	20	
Total	**127**	**87**	**69**	**25**	**20**	**19**	**15**	**12**	**7**
					% larger	**76**		**% larger**	**47**
	note:	Fail (<50)	3						

The following conclusions can be made based on the results of the sensitivity analysis.

- Overall, the LOM defuzzification method provides the most accurate results. The comparison of this method with the base case method is shown in Figure 7-5.
- Changing defuzzification methods most significantly changes the accuracy of the models. So this parameter should receive more attention.
- The failed models do not deviate very far according to the 1-term and 2-term analysis.

- The predicted linguistic terms tend to be larger than the actual if they are not perfectly matched.
- The failed models can be attributed to several factors. First, not having enough satisfactory data to generate very accurate membership functions contributed to the deviation. Second, the actual values may have a bias in themselves.

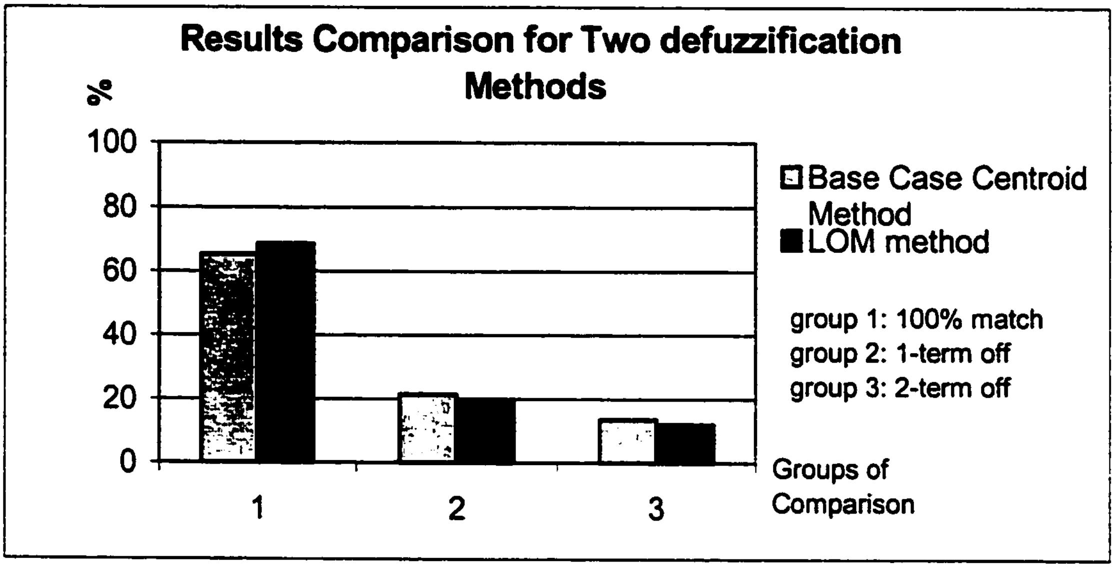


Figure 7-5: Two Defuzzification Methods Results Comparison

7.7 Uses of Fuzzy Expert System

The uses of the constructed fuzzy expert system can be summarized as follows:

- Given any detailed information (represented by the sub-input factors in each model) for each input factor model, it will predict a crisp value and a linguistic term to describe each input factor. For example, once the information for the number of employees in the design firm and the total volume of work of the company is available, then the fuzzy expert model can predict the overall size of the company. (1 means small, 10 means large). This information provides the project manager with a good description of the project environment. Furthermore, it is an essential step to predict design performance if pre-construction information is not available.

- Given all the pre-construction information (i.e., values for the higher-level input factors), the performance of the design project can be predicted before the actual construction is complete, using higher-level input factors to sub-output factor models. This is very useful for project management. It will alert all the construction parties so that they will be aware of any potential problems that may arise, such as cost overrun, schedule overrun, or poor quality of the design.
- This system can also be used to help the project manager to decide which factors are the most significant inputs to the system based on the most significant performance criteria (i.e., output measure).

7.8 Summary

In this chapter, the complete steps for building the fuzzy expert system were presented. The constructed model was tested using actual data. Sensitivity analysis provided insight into methods to improve the accuracy of the model. 71% of the models succeeded in predicting linguistic terms using the centroid method of defuzzification. 81% of the models succeeded using the LOM method. The accuracy of the models in predicting numerical values is low. Much work remains to be done to improve the accuracy of the fuzzy expert system. The following chapter will discuss what future work needs to be done and how it might be achieved.

CHAPTER 8 CONCLUSIONS AND FUTURE RESEARCH

8.1 Conclusions

The objective of this research was to develop a model for use in evaluating and predicting a design firm's performance using fuzzy set theory and fuzzy logic techniques. This goal was achieved by identifying the potential factors that would affect design performance, generating membership functions for these factors, building the rulebase for a fuzzy expert system, and then constructing and testing the fuzzy model in a computerized environment.

The first stage of the research was to identify a thorough list of factors that would affect design performance. After extensive literature review as well as discussions with experts both in the industry and the academic area, the factors were identified. For the purpose of modeling, all these factors were classified into 3 categories: context variables, input variables, and output variables. The preliminary model was constructed based on the structure of these factors.

The second stage dealt with data collection; a mailout survey was used to collect actual data from design companies. Eighteen responses were collected based on actual design projects. The data was stored and processed for modeling purposes.

The third stage dealt with generating membership functions based on objective values. A new technique was developed that made use of the limited data set. The generated membership functions were tested twice to prove the feasibility and accuracy of this new technique. In both cases, 81% out of 89 models of membership functions succeeded.

The fourth stage of the study focused on the model's simplification. Due to the complexity of the proposed model, it was difficult to generate rules and to implement them in computerized form. Correlation analysis was used to reduce the number of factors in the model. This method not only simplified the model, but also provided the direction of the variable relationship for generating the If-Then rules.

The fifth stage of the study dealt with the development of the fuzzy expert system. The fuzzy expert system consists of membership functions, a fuzzy rulebase, a fuzzy inference system, and a defuzzification module. At this stage the study focused on developing fuzzy If-Then rules, since the rest of the components had already been developed. A new technique was developed and used to generate If-Then rules, based on the frequency of actual data and the correlation analysis. Once the rules were generated, the fuzzy expert model was implemented in Matlab. The model was tested with the actual data collected from the survey. The model produced good results, especially in linguistic terms. 71% of the sub-models were successful on the basis of predicted linguistic terms. In the design context, linguistic terms are more commonly used than crisp numbers to describe the dynamics of design project performance. Consequently, the performance of the model is acceptable.

The last stage of this study involved conducting a sensitivity analysis of the fuzzy expert system, by changing several of its parameters. The results indicate that the model is sensitive to the defuzzification method. Changing the operator or the implication-aggregation method does not change the accuracy of the model. Moreover, changing the defuzzification method from the most common centroid method to the LOM method slightly increased the predication success rate. By using the LOM method, 81% of the sub-models were successful on the basis of predicted linguistic terms.

This study shows how fuzzy set theory can be used to model design performance evaluation and prediction, which involves much linguistic and subjective evaluation. The fuzzy expert system developed shows a good ability to relate multiple inputs to one or more outputs. The use of natural language for reasoning in the model is a realistic and desirable feature for decision making in project management.

8.2 Contributions

This study has made several contributions to both academic research and industrial applications. Academically, it has demonstrated the appropriateness of the application of fuzzy set theory in design performance prediction and evaluation. Secondly, the study proposed a new technique for generating membership functions based on objective data. The fact that these membership functions are based on objective data means that they can be calibrated to suit different contexts. This is a first step towards developing membership functions that have the same meaning to different users. This technique for generating membership functions is useful for building membership functions when there is not enough information or experience available, except for a few limited sources of data. Third, the study developed a new technique for generating If-Then rules based on a limited data set. These techniques can be improved significantly if a large enough data set can be obtained. Based on the available information, these techniques produced very promising results, and laid the foundation for future research in this area. Finally, this research demonstrates that statistical analysis can be used as an intermediate step in the development of rules for a fuzzy expert system. It also shows the flexibility of fuzzy set theory when used with other techniques.

Industrial contributions include the development of a model that provides insight into the factors that affect design performance and evaluation. The fuzzy expert system developed provides a tool for design performance prediction and evaluation, both of which are difficult to quantify and measure. This fuzzy expert system may be useful to project management personnel in evaluating design projects.

8.3 Limitations of the Research and Recommendations for Future Research

There are some limitations of this research that need to be addressed. These limitations lie in the following areas:

- Model design:

  The present research did not examine the impact of the context variables on the model, and how these factors affect design performance.

- Data collection:

  Because the data collection technique consisted of a mailout survey, not enough data were collected. Many of the ensuing limitations are due to limited data.

- Membership functions:

  Even though membership function testing produced good results, there are still some membership functions that failed, largely due to lack of data. This may lead to some failures of the fuzzy If-Then rules. Also, because of the limited data, some membership functions exist with great overlaps between them. These factors lowered the accuracy rate of the membership functions' performance.

- Fuzzy expert system:

  The fuzzy expert system does not achieve a high success rate for numerical prediction; there is a large bias between the actual value and the predicted crisp value. Because of the roughness of the membership functions, the fuzzy expert system can not conclude an accurate enough crisp number; however, prediction of linguistic terms is reasonably accurate.

Despite the limitations of the current model, this research provides very promising results for future development. The following are recommendations for future research:

1) Before carrying out data collection, thorough research on the proposed method of data collection needs to be done to increase its success rate. The method of data collection should be matched with the type, quantity, and quality of data required. The proposed models should be refined by redefining the factors and examining case studies to make sure the proposed models are realistic and useful. The linguistic descriptors for each factor may also need to be redefined to be more appropriate.
2) Case studies should be used for future data collection to eliminate unclear and missing information in the survey, thus increasing the quality of the data.
3) To achieve better results from the survey, cross validation can also be used, that is, repeat the same question in two different ways to validate the accuracy of the responses.
4) More research needs to be done on how context variables affect design performance and how to use this information when building the model.

5) For membership function generation, other shapes of membership functions should be explored. Since current membership functions have many overlaps, perhaps all of the membership functions should be redefined as triangular shapes to increase the accuracy rate.

6) In addition to correlation analysis, non-linear regression analysis can be used to explore the effect of partial correlation between input factors and for differential weighting of the rules.

7) A better way to generate If-Then rules is to collect a large enough data set, then use the frequency of combinations of input data with output data to generate the rules. Thus, the accuracy of actual rules may be improved significantly.

8) To build a complete relationship rulebase, combinations of "and" and "or" operators in one rule can be explored to incorporate all input factors that are correlated to an output factor, and are correlated with varying degrees to each other.

9) For linguistic term matching, there is also some work that can be done to improve the match criteria. In addition to a visual observation to decide the range of a predicted linguistic term location, a degree to which that predicted term belongs to this area should also be considered. This could be done manually from the membership functions of the output factor.

10) To improve the accuracy of the model, a different defuzzification method may need to be developed. Instead of LOM or centroid, the deffuzification method should try to defuzzify the area where the maximum linguistic term is located. This may produce a more accurate crisp number.

11) Whether changing the shape of the membership functions will increase the accuracy of the models or not can be examined in sensitivity analysis. The use of triangular membership functions rather than trapezoidal functions may improve the accuracy of the model's predictions.
12) Other methods of mapping inputs to outputs besides If-Then rules should be examined, such as neural network and decision trees.

This research has explored several aspects of fuzzy logic and fuzzy expert system development. Despite their limitations, many of the proposed techniques are useful in laying the groundwork necessary to further develop the proposed concepts. The advancement of these concepts will contribute to our knowledge of fuzzy set theory and to its practical application.

REFERENCES

Ahmad, I., and Minkarah, I., *Questionnaire Survey on Bidding in Construction*, Journal of Management in Engineering, 1988 4(3), pp. 229-243.

Araki, S., Nomura H., Hayashi, I., and Wakami, N., *A Self-Generating Method of Fuzzy Inference Rules*, Fuzzy Engineering toward Human Friendly Systems, Proceedings of IFES 1991, pp. 1047-1059.

Ayyub, B. M., and Haldar, A., *Project Scheduling using Fuzzy Set Concepts*, Journal of Construction Engineering and Management, vol. 110, 1983, pp. 189-204.

APEGGA, *Consultant Fees for General Engineering Assignments – A Guideline*, the Association of Professional Engineers, Geologists and Geophysicists of Alberta, September 1990a.

APEGGA, *Consultant Fees for Building Projects – A Guideline*, the Association of Professional Engineers, Geologists and Geophysicists of Alberta, September 1990b.

APEGGA and the Alberta Association of Architects, *Schedule of Designated Services for Recommended Conditions of Engagement and Schedule of Professional Fees for Building Projects*, the Association of Professional Engineers, Geologists and Geophysicists of Alberta, April 1998a.

APEGGA and the Alberta Association of Architects, *Recommended Conditions of Engagement and Schedule of Professional Fees for Building Projects*, the Association of Professional Engineers, Geologists and Geophysicists of Alberta, May 1998b.

Armentrout D. R., *Engineering Productivity Management and Performance Measurement*, Journal of Management in Engineering, vol. 2, 1986, pp. 141-147.

Baldwin, A. N., Austin, S. A., Hassan, T. M., and Thorpe, A., *Modelling Information Flow during the Conceptual and Schematic Stages of Building Design*, Construction Management and Economics, 17, 1999, pp. 155-167.

Bernd, M. and Beer, M, *Safety Assessment Using Fuzzy Theory*, Computing in Civil Engineering, conference proceeding, 1998, pp. 756-759.

Bezdek, J., C., *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum Press, New York, 1981.

Bobrowicz, O., Choulet, C., Haurat, A., Sandoz, F. and Tebaa, M., *A Method To Build Membership Functions – Application to numerical/symbolic interface building*, 3rd International Conference on Information Processing and Management of Uncertainty in Knowledge – based Systems, IPMU '90, Paris, France, July 1990, Proceedings 521, pp.136-142.

Bogardi, I, Bardossy, A, and Duckstein, L, *Risk Management for Groundwater Contamination: Fuzzy Set Approach*, Optimizing the Resources for Water Management, conference proceeding, 1990, pp. 442-448.

Chalabi, A. F., Salazar, G. F. and Beaudin, B. J., *Defining and Evaluating Input Variables Impacting Design Effectiveness Research Phase I*, A Report to the Construction Industry Institute, the University of Texas at Austin, January 1986.

Chalabi, A. F., Beaudin, B. J., and Salazar, G., *Input Variables Impacting Design Effectiveness*, A Report to the Construction Industry Institute, the University of Texas at Austin, April 1987.

Chao, C. and Cheng, F., *Fuzzy Pattern Recognition Model for Diagnosing Cracks in RC Structures*, Journal of Computing in Civil Engineering, Vol. 12, No. 2, April 1998, pp. 111-119.

Chao, L., C., and Skibniewski, M. J., *Fuzzy Logic for Evaluating Alternative Construction Technology*, Journal of Construction Engineering and Management, vol. 124, 1998, pp. 297-304.

Chen, J. E., and Otto, K., N., *Constructing Membership Functions using Interpolation and Measurement Theory*, Fuzzy Sets and Systems 73, 1995, pp. 313-327.

Chen, M., S., and Wang, S. W., *Fuzzy Clustering Analysis for Optimizing Fuzzy Membership Functions*, Fuzzy Sets and Systems 103, 1999, pp. 239-254.

Chiu, S. L., *Fuzzy Model Identification Based on Cluster Estimation*, Journal of Intelligent and Fuzzy Systems, vol. 2, 1994, pp. 267-278.

CII, *Evaluation of Design Effectiveness*, Report 8-1, the Construction Industry Institute, the University of Texas at Austin, July 1986.

CII, *Input Variables Impacting Design Effectiveness*, Report 8-2, the Construction Industry Institute, the University of Texas at Austin, July 1987.

Civanlar, M.R. and Trussell, H. J., *Constructing Membership Functions using Statistical Data*, Fuzzy Sets and Systems 18, 1986, pp. 1-13.

Dahab, M., F., Fürhacker, M, and Zibuschka, F., *Pathogen Risk Assessment of Land Applied Wastewater and Biosolids: A Fuzzy Set Approach*, Water Resources and the Urban Environment-98, 1998, conference proceeding, pp. 173-178.

Dunn, J., *A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well Separated Clusters*, J. Cibernet., 3, 1974, pp. 32-57.

Eldin, N. N., *Management of Engineering / Design Phase*, Journal of Construction Engineering and Management, vol. 117, 1991, pp. 163-175.

Fayek, A., *Competitive Bidding Strategy Model and Software System for Bid Preparation*, Journal of Construction Engineering and Management, vol. 124, 1998, pp. 1-10.

Gue, K. and Fayek, A., *Report on the University of Alberta's Study of the Factors Affecting Design Performance*, University of Alberta, Edmonton, August 1999.

Herbsman, Z., and Ellis, R., *Research of Factors Influencing Construction Productivity*, Construction Management and Economics, 8, 1990, pp. 49-61.

Hong, T. P., and Chen, J. B., *Finding Relevant Attributes and Membership Functions*, Fuzzy Sets and Systems 103, 1999, pp. 389-404.

Hong, T. P., and Lee, C. Y., *Induction of Fuzzy Rules and Membership Functions from Training Examples*, Fuzzy Sets and Systems 84, 1996, pp. 33-47.

Krishnapuram, R., *Generation of Membership Functions via Possibilistic Clustering*, IEEE Proceedings, 1994, pp. 902-908.

Lee, Y., W., Dahab, M., and Bogardi, I., *Nitrate-Risk Assessment Using Fuzzy-Set Approach*, Journal of Environmental Engineering, Vol. 121, No. 3, March 1995, pp. 245-256

Li, R. P. and Mukaidono, M., *Gaussian Clustering Method based on Maximum-Fuzzy-Entropy Interpretation*, Fuzzy Sets and Systems 102, 1999, pp. 253-258.

Li, H. X., and Yen, V. C., *Fuzzy Sets and Fuzzy Decision-Making*, CRC Press, Florida, 1995, pp. 93-122.

Lorterapong, P., and Moselhi, O., *Project-Network Analysis using Fuzzy Sets Theory*, Journal of Construction Engineering and Management, vol. 122, 1996, pp. 308-318.

Mason, A. K., and Kahn, D. J., *Estimating Costs with Fuzzy Logic*, 1997 AACE International Transactions, EST. 03.1-EST. 0.3.6.

Mathworks, *Matlab -- the Language of Technical Computing, (version 5.3.0 10183 R11)*, the Mathworks, Inc., 1999.

McGeorge, J. F., *Design Productivity: A Quality Problem*, Journal of Management in Engineering, vol. 4, 1988, pp. 350-362.

Norusis, M. J., *SPSS for Windows Base System User's Guide Release 6.0*, SPSS Inc., 1993.

Pedrycz, W. and Gomide, F., *An Introduction to Fuzzy Sets: Analysis and Design*, A Bradford Book, the MIT Press, Cambridge, Massachusetts, London, England, 1998.

Portas J., and AbouRizk, S., *Neural Network Model for Estimating Construction Productivity*, Journal of Construction Engineering and Management, vol. 123, 1997, pp. 399-410.

Ross, T., Trumm, J., and Donald, S., *Decision-Making Paradigm for Risk Assessment of Hazardous Wastes Using Fuzzy Sets*, Computing in Civil Engineering, conference proceeding, 1994, pp. 200-207.

Russell, A. D., and Fayek, A., *Automated Corrective Action Selection Assistant*, Journal of Construction Engineering and Management, vol. 120, 1994, pp. 11-33.

Scarlett, B. R., *Evaluation of Design Effectiveness*, Master Thesis, the University of Texas at Austin, 1986.

Sintas, A. F., Cadenas, J. M., and Martin, F., *Membership Functions in the Fuzzy C-means Algorithm*, Fuzzy Sets and Systems 101, 1999, pp. 49-58.

Stull, J. O., and Tucker, R. L., *Objectives Matrix Values for Evaluation of Design Effectiveness*, A Report to the Construction Industry Institute, the University of Texas at Austin, November, 1986.

Thiel, C, Singh, J, and Boissonnade, A, *Risk Assessment of Sea Bottom Scouring Using Fuzzy Set Theory*, Civil Engineering in the Arctic Offshore, F. Lawrence Bennett, et al, eds., 1985, conference proceeding, pp. 430-438.

Thomas, H. R., and Yiakoumis, I, *Factor Model of Construction Productivity*, Journal of Construction Engineering and Management, vol. 113, 1987, pp. 623-639.

Thomas, H. R., Korte, Q. C., Sanvido, V. E., and Parfitt, M. K., *Conceptual Model for Measuring Productivity of Design and Engineering*, Journal of Architectural Engineering, vol. 5, 1999, pp. 1-8.

Yager, R. R. and Filev, D., *Template-Based Fuzzy Systems Modeling*, Journal of Intelligent and Fuzzy Systems, vol. 2, 1994, pp. 39-54.

Yager, R. R. and Filev, D., *Generation of Fuzzy Rules by Mountain Clustering*, Journal of Intelligent and Fuzzy Systems, vol. 2, 1994, pp. 209-219.

Yin, M. Yuan W., and Stenstrom, M., K, *Fuzzy-Supported Estimator for HPO-AS Process*, Journal of Environmental Engineering, Vol. 125, No. 2, February 1999, pp. 137-145.

Zhang, Z. and Tumay, M., *Statistical to Fuzzy Approach Toward CPT Soil Classification*, Journal of Geotechnical and Geoenvironmental Engineering, Vol. 125, No. 3, March 1999, pp. 179-186.

Appendix 1: Sample Questionnaire

General Information

In order to ensure confidentiality, your company information will not be linked in any way to the project information in subsequent sections. You may respond anonymously if you wish.

Name of Company:______________________________

Location of Company:______________________________

Name of Respondent:______________________________

Title of Respondent:______________________________

Date of Response:___/___/______

Would you like your company name acknowledged as a participant in this survey in the report to be released to all participating companies? Your company will not be identified with any of the data you provide.

Yes ☐

No ☐

Project Details

For the following sections, please report on a single recent industrial sector design project that was completed by your firm in either 1998 or 1999.

Name of Project:______________________________

Location of Project:______________________________

Name of Owner:______________________________Public ☐ Private ☐

Start Date of Design Contract:_____/___/___

Duration of Design Contract:______________________

Total $ Value of Design Contract:______________________

Design Firm's Role: Prime Consultant ☐ Sub-consultant ☐

Start Date of Construction Contract:_______/___/________

Duration of Construction Contract:______________________

Total $ Value of Construction Contract:______________________

Project Information

Please select the most appropriate answer from each group of options for the project that was described in the previous section. You may select more than one category, if applicable.

1. What is the type of project?
   - Oil and Gas - Pipeline ☐
   - Oil and Gas – Refinery, Compressor Station ☐
   - Chemical Processing or Extraction Plant ☐
   - Mining ☐
   - Pulp and Paper Mill ☐
   - Power Plant ☐
   - Water Treatment Plant ☐
   - Other (specify):__________ ☐

2. What is the type of design contract?
   - Lump Sum ☐
   - Unit Price ☐
   - Cost Plus (Cost Reimbursable) ☐
   - Guaranteed Maximum Price ☐
   - Negotiated ☐
   - Other (specify):__________ ☐

3. What is the scope of the design contract?
   - Design Only ☐
   - Design and Manage (construction or project management) ☐
   - Design and Build ☐
   - Other (specify):__________ ☐

4. What is the type of construction contract?
   - Lump Sum ☐
   - Unit Price ☐
   - Cost Plus (Cost Reimbursable) ☐
   - Guaranteed Maximum Price ☐
   - Negotiated ☐
   - Other (specify):__________ ☐

5. What is the scope of the construction contract?

Construct Only ☐
Design and Build ☐
Management (construction or project management) ☐
Other (specify):_______________ ☐

6. What was the method of tendering for the construction contract?
   Open ☐
   Prequalified ☐
   Other (specify):_______________ ☐

7. Please rate the following project priorities in their order of importance for the project (from 1 to 8).

Factor	Relative Importance (1 = most important)
Cost	
Schedule	
Quality	
Safety	
Aesthetics	
Environmental Impact	
Constructability	
Potential For Future Development Or Expansion	

Variables That Impacted the Project

The following questions require both a numeric response and a linguistic rating of the magnitude of the response. If you are not sure of the numeric value, you may specify a range of values. You may also circle more than one linguistic rating if you believe the magnitude falls somewhere between the specified values (e.g. small and average).

The information gathered from this survey will be used to model the influence of each factor on design performance. ***In order to achieve accurate relationships, the survey must be filled out as completely as possible. If you are not sure of an exact value for any particular question, please estimate an approximate value.***

The term "average" implies that the condition is standard in the experience of the respondent.

1. Rate the overall size of the industrial division of your design firm, relative to other industrial design firms.

Rating: 1 2 3 4 5 6 7 8 9 10
(small) (large)

Industry wide, do you consider this (circle): small average large

1.1. How many people does the industrial division employ?

Number:___________

Industry wide, do you consider this (circle): small average large

1.2. What is the average annual volume of work of the industrial division in the previous 3 years?

Amount ($):________________________

Industry wide, do you consider this (circle): small average large

1.3. How many design projects does the industrial division currently hold?

Number:___________

Industry wide, do you consider this (circle): small average large

2. Rate the current level of competition in the industrial design market.

Rating: 1 2 3 4 5 6 7 8 9 10
(low) (high)

Do you consider this (circle): low average high

2.1. How many similar industrial design firms are currently in the market?

Number:___________

Do you consider this (circle): small average large

2.2. How many industrial design projects are available in the current market?

Number:__________

Do you consider this (circle): small average large

3. Rate the overall quality of your firm's profile, relative to other industrial design firms.

Rating: 1 2 3 4 5 6 7 8 9 10
(poor) (good)

Industry wide, do you consider this (circle): poor average good

The following questions refer to the specific project that was previously chosen.

3.1. Rate the extent to which the project design is within the normal scope of your firm.

Rating: 1 2 3 4 5 6 7 8 9 10
(small) (large)

Industry wide, do you consider this (circle): small average large

3.2. How many designers were involved on the project?

Number:__________

Industry wide, do you consider this (circle): small average large

3.3. What was the ratio of senior to junior designers on the project?

Ratio:_____/_____

Industry wide, do you consider this (circle): small average large

3.4. Rate the level of skill of the design team.

Rating: 1 2 3 4 5 6 7 8 9 10
(poor) (good)

Industry wide, do you consider this (circle): poor average good

3.5. What is the average experience (in years) of design team members?

Number of Years:__________

Industry wide, do you consider this (circle): small average large

3.6. Rate the motivational and leadership skills of the design team supervisor.

Rating: 1 2 3 4 5 6 7 8 9 10
(poor) (good)

Industry wide, do you consider this (circle): poor average good

3.7. How many years of experience does the supervisor have?

Number of Years:_____________

Industry wide, do you consider this (circle): small average large

3.8. How many designer personnel changes occurred during the project?

Number of Changes:__________

Industry wide, do you consider this (circle): small average large

3.9. Rate the familiarity of the design team with the CAD or other software used on the project.

Rating: 1 2 3 4 5 6 7 8 9 10
(low) (high)

Industry wide, do you consider this (circle): low average high

4. Rate the size of the design contract.

Rating: 1 2 3 4 5 6 7 8 9 10
(small) (large)

Industry wide, do you consider this (circle): small average large

4.1. What was the total cost of design?

Amount ($):________________

Industry wide, do you consider this (circle): small average large

4.2. How long was the total duration of the design process?

Length of Time:_____________

Industry wide, do you consider this (circle): short average long

4.3. How many manhours were expended on the design?

Manhours:_________________

Industry wide, do you consider this (circle): small average large

5. Rate the continuity of the manhour commitment for the design project.

Rating: 1 2 3 4 5 6 7 8 9 10
(small) (large)

Industry wide, do you consider this (circle): small average large

5.1. On average, how many manhours per week did each individual designer expend on the project?

Manhours:__________________

Industry wide, do you consider this (circle): small average large

5.2. On average, how many manhours per week did the entire design team expend on the project?

Manhours:__________________

Industry wide, do you consider this (circle): small average large

6. Rate the level of scope definition for the project prior to detailed design.

Rating: 1 2 3 4 5 6 7 8 9 10
(low) (high)

Industry wide, do you consider this (circle): low average high

Please rate the extent to which each of the following requirements were met on the project.

6.1. The definition of project type, description of facility, project priorities and objectives were made clear.

Rating: 1 2 3 4 5 6 7 8 9 10
(low) (high)

Industry wide, do you consider this (circle): low average high

6.2. A description of alternatives being considered and their potential impact on scope were made clear.

Rating: 1 2 3 4 5 6 7 8 9 10
(low) (high)

Industry wide, do you consider this (circle): low average high

6.3. What was the percent of basic design data available prior to detailed design?

Percent (%):__________________

Industry wide, do you consider this (circle): small average large

6.4. Rate the amount of similar design information available from previous projects.

Rating: 1 2 3 4 5 6 7 8 9 10
(small) (large)

Industry wide, do you consider this (circle): small average large

For all of Question 7, "average" indicates that your answer is typical for the project type.

7. Rate the complexity of function of the project.

Rating: 1 2 3 4 5 6 7 8 9 10
(low) (high)

Industry wide, do you consider this (circle): low average high

7.1. What was the percent of repetition of design features on the project?

Percent (%):________________

Industry wide, do you consider this (circle): small average large

7.2. What was the percent of unique or new design features on the project?

Percent (%):________________

Industry wide, do you consider this (circle): small average large

7.3. What percent of the project was in upgrades or changes to an existing facility?

Percent (%):________________

Industry wide, do you consider this (circle): small average large

7.4. As a percent of structural, mechanical and/or electrical systems, to what degree were there specialized structural, mechanical and/or electrical system requirements on the project?

Percent (%):________________

Industry wide, do you consider this (circle): small average large

7.5. As a percent of the building envelope design (if applicable), to what degree were there special considerations on the project for the building envelope?

Percent (%):________________

Industry wide, do you consider this (circle): small average large

8. Rate the complexity of the design process for the project.
"Average" conditions indicate single owner, single basic review authority, approvals at completion of each stage, design within normal scope of design firm.

Rating: 1 2 3 4 5 6 7 8 9 10
(low) (high)

Industry wide, do you consider this (circle): low average high

8.1. How many design contracts did the entire project involve?

Number:__________

Industry wide, do you consider this (circle): small average large

8.2. In how many locations was the project engineered?

Number:__________

Industry wide, do you consider this (circle): small average large

8.3. How many owners or stakeholders were involved in the project?

Number:__________

Industry wide, do you consider this (circle): small average large

8.4. How many review authorities were involved in the project?

Number:__________

Industry wide, do you consider this (circle): small average large

8.5. What was the average length of time for review and approvals on the project?

Number of Days:____________

Industry wide, do you consider this (circle): short average long

8.6. How many environmental assessment reviews were required for the project?

Number:__________

Industry wide, do you consider this (circle): small average large

9. Rate the complexity of the project conditions.
"Average" conditions indicate a relatively uncomplicated site, compatible land use, stable soils, good access, services available, similar neighbours.

Rating: 1 (low) 2 3 4 5 6 7 8 9 10 (high)

Industry wide, do you consider this (circle): low average high

For each of the following variables, please provide the number of times the problem affected design throughout the project and a rating of the magnitude of the problem.

9.1. Insufficient working area.

Number:____________

Industry wide, do you consider this (circle): small average large

Rating of magnitude of problem: 1 (small) 2 3 4 5 6 7 8 9 10 (large)

Industry wide, do you consider this (circle): small average large

9.2. Restricted access to site.

Number:______________

Industry wide, do you consider this (circle): small average large

Rating of magnitude of problem: 1 2 3 4 5 6 7 8 9 10
(small) (large)

Industry wide, do you consider this (circle): small average large

9.3. In-situ soil conditions.

Number:______________

Industry wide, do you consider this (circle): small average large

Rating of magnitude of problem: 1 2 3 4 5 6 7 8 9 10
(small) (large)

Industry wide, do you consider this (circle): small average large

9.4. Air temperature.

Number:______________

Industry wide, do you consider this (circle): small average large

Rating of magnitude of problem: 1 2 3 4 5 6 7 8 9 10
(small) (large)

Industry wide, do you consider this (circle): small average large

9.5. Amount of Precipitation.

Number:______________

Industry wide, do you consider this (circle): small average large

Rating of magnitude of problem: 1 2 3 4 5 6 7 8 9 10
(small) (large)

Industry wide, do you consider this (circle): small average large

9.6. Lack of services available to site.

Number:______________

Industry wide, do you consider this (circle): small average large

Rating of magnitude of problem: 1 2 3 4 5 6 7 8 9 10
(small) (large)

Industry wide, do you consider this (circle): small average large

9.7. Compatibility with land use zoning.

Number:______________

Industry wide, do you consider this (circle): small average large

Rating of magnitude of problem: 1 (small) 2 3 4 5 6 7 8 9 10 (large)

Industry wide, do you consider this (circle): small average large

9.8. Disposal of contaminated materials.

Number:______________

Industry wide, do you consider this (circle): small average large

Rating of magnitude of problem: 1 (small) 2 3 4 5 6 7 8 9 10 (large)

Industry wide, do you consider this (circle): small average large

9.9. Other (specify):___________________________

Number:______________

Industry wide, do you consider this (circle): small average large

Rating of magnitude of problem: 1 (small) 2 3 4 5 6 7 8 9 10 (large)

Industry wide, do you consider this (circle): small average large

9.10. Other (specify):___________________________

Number:______________

Industry wide, do you consider this (circle): small average large

Rating of magnitude of problem: 1 (small) 2 3 4 5 6 7 8 9 10 (large)

Industry wide, do you consider this (circle): small average large

9.11. Other (specify):___________________________

Number:______________

Industry wide, do you consider this (circle): small average large

Rating of magnitude of problem: 1 (small) 2 3 4 5 6 7 8 9 10 (large)

Industry wide, do you consider this (circle): small average large

10. Rate the quality of the owner's profile and participation during the project.

Rating: 1 2 3 4 5 6 7 8 9 10
(poor) (good)

Industry wide, do you consider this (circle): poor average good

10.1. On average, how long did the owner take to make a decision on the project?

Number of Days:______________

Industry wide, do you consider this (circle): short average long

10.2. How many times did the owner interfere or change mind?

Number of Times:____________

Industry wide, do you consider this (circle): small average large

10.3. During the project, how many times was there a change in the owner's personnel?

Number of Times:____________

Industry wide, do you consider this (circle): small average large

10.4. How many years of experience did the owner's representative have?

Number of Years:____________

Industry wide, do you consider this (circle): small average large

10.5. Rate the owner's attitude toward risk.

Rating: 1 2 3 4 5 6 7 8 9 10
(risk averse) (risk prone)

Industry wide, do you consider this (circle): risk averse average risk prone

11. Rate the quality of the primary vendors' profiles on the project.

Rating: 1 2 3 4 5 6 7 8 9 10
(poor) (good)

Industry wide, do you consider this (circle): poor average good

11.1. How long did it take to receive certified information for instrumentation, equipment drawings, and specifications for the project?

Number of Days:______________

Industry wide, do you consider this (circle): short average long

11.2. How complete, as a percent complete, was the certified information provided for instrumentation, equipment drawings, and specifications for the project?

Percent Complete (%):________

Industry wide, do you consider this (circle): small average large

11.3. How many errors were contained in the certified information provided for instrumentation, equipment drawings, and specifications for the project?

Number of Errors:____________

Industry wide, do you consider this (circle): small average large

12. Rate the complexity of the tendering process used for construction of the project. *"Average" indicates open or pre-qualified tender, basic documents, standard agreement forms, stipulated sum contract.*

Rating: 1 2 3 4 5 6 7 8 9 10
(low) (high)

Industry wide, do you consider this (circle): low average high

12.1. How many construction work packages did the project involve?

Number:__________

Industry wide, do you consider this (circle): small average large

12.2. What percentage of forms and conditions were non-standard for the project?

Percent (%):________

Industry wide, do you consider this (circle): small average large

13. Rate the complexity of the construction process that followed the design project. *"Average" indicates single prime contractor, normal construction site, periodic site visits, pre-qualified tenders.*

Rating: 1 2 3 4 5 6 7 8 9 10
(low) (high)

Industry wide, do you consider this (circle): low average high

13.1. What percentage of the design was complete prior to construction?

Percent (%):________

Industry wide, do you consider this (circle): small average large

13.2. How many phases of construction were there?

Number:__________

Industry wide, do you consider this (circle): small average large

13.3. How many prime contractors were involved in construction?

Number:__________

Industry wide, do you consider this (circle): small average large

13.4. How many subcontractors were involved in construction?

Number:__________

Industry wide, do you consider this (circle): small average large

13.5. What was the frequency of site visits required by the design consultant?

Number per Week:__________ **OR** Number per Month:__________

Industry wide, do you consider this (circle): small average large

13.6. What was the percentage of work that was in renovations or additions?

Percent (%):________

Industry wide, do you consider this (circle): small average large

13.7. How many days of early occupation of site by the owner were required?

Number of Days:__________

Industry wide, do you consider this (circle): small average large

14. Rate the economic (market) conditions surrounding the design project.

Rating: 1 2 3 4 5 6 7 8 9 10
(unfavourable) (favourable)

Do you consider this (circle): unfavourable average favourable

Outcomes of the Project

1. Rate the level of performance against the cost of design for the design project.

   Rating: 1 2 3 4 5 6 7 8 9 10
   (poor) (good)

   Industry wide, do you consider this (circle): poor average good

   1.1. What was the percentage increase (+) or decrease (-) in actual design manhours vs. budgeted design manhours for the project?

   Percent (%):________ Increase ☐ Decrease ☐

   Industry wide, do you consider this (circle): small average large

   1.2. What percentage of design manhours was due to owner changes during the project?

   Percent (%):________

   Industry wide, do you consider this (circle): small average large

   1.3. What percentage of design manhours was due to rework during the project?

   Percent (%):________

   Industry wide, do you consider this (circle): small average large

   1.4. What was the percentage increase (+) or decrease (-) in actual design cost (including changes) vs. budgeted design cost for the project?

   Percent (%):________ Increase ☐ Decrease ☐

   Industry wide, do you consider this (circle): small average large

   1.5. What percentage of design cost was due to owner changes during the project?

   Percent (%):________

   Industry wide, do you consider this (circle): small average large

   1.6. What was the design cost (including changes) as a percentage of the total cost of construction of the project?

   Percent (%):________

   Industry wide, do you consider this (circle): small average large

2. Rate the level of performance against the schedule for the design project.

   Rating: 1 2 3 4 5 6 7 8 9 10
   (poor) (good)

   Industry wide, do you consider this (circle): poor average good

2.1. What was the percentage increase (+) or decrease (-) in actual design duration vs. scheduled design duration for the project?

Percent (%):________ Increase ☐ Decrease ☐

Industry wide, do you consider this (circle): small average large

2.2. What percentage of design document release deadlines were missed during the project?

Percent (%):________

Industry wide, do you consider this (circle): small average large

3. Rate the level of accuracy of design documents issued throughout the project.

Rating: 1 (low) 2 3 4 5 6 7 8 9 10 (high)

Industry wide, do you consider this (circle): low average high

3.1. How many approved changes were made during construction?

Number:___________

Industry wide, do you consider this (circle): small average large

3.2. What was the total cost of approved changes made during construction?

Cost ($):___________

Industry wide, do you consider this (circle): small average large

3.3. What percentage of the total value of the construction contract was due to changes?

Percent (%):________

Industry wide, do you consider this (circle): small average large

3.4. What was the total number of design rework manhours during construction?

Manhours:_________

Industry wide, do you consider this (circle): small average large

3.5. What was the total number of problems that occurred during construction due to errors, incompleteness, or lack of clarity in the design documents?

Number:___________

Industry wide, do you consider this (circle): small average large

Thank you for your participation!

Appendix 2: Respondent Results from the Survey

The questionnaire developed for the survey began with a section on general company information that gathered general data from the respondents. The next two sections in the questionnaire focused on the details of a specific project selected by the respondent and information on the selected project. The section following listed the fourteen input factors that impact design performance. Respondents were asked to rate and describe the impact of each factor during the project. The questionnaire ended with a similar section requiring respondents to rate and describe the impact of the three output indicators of design performance. The results of the survey are presented in this section.

1.1. General Information

Title of Respondent

All respondents to the survey occupied a position in their respective companies that allowed them to have a detailed and complete knowledge of the project on which they reported. The majority were in managerial positions and oversaw multiple projects. Thirty-three percent (33%) of respondents were project managers and eleven percent (11%) worked in the estimating department. The percent of total respondents with each job description is shown in Figure 2.

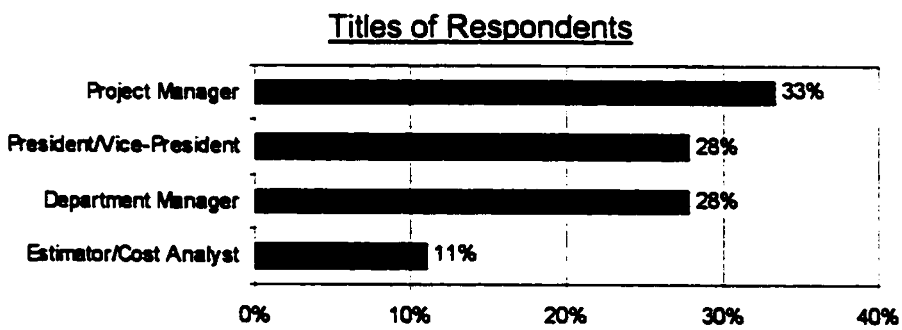


Figure 2. Titles of Respondents to Survey

1.2. Project Details

Location of Project

The projects that were chosen by the responding companies were mainly located in Alberta or British Columbia. Citing confidentiality, some companies withheld the location of their project. A graph comparing project locations is presented in Figure 3.

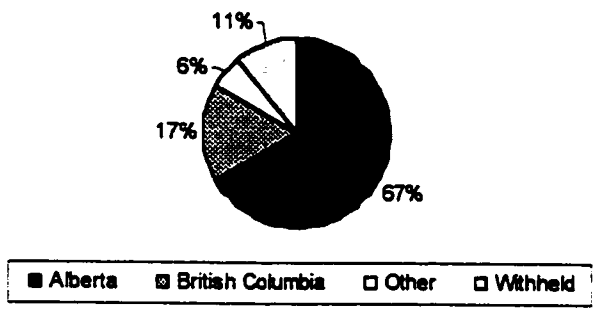


Figure 3. Location of Project Selected

Identity of Owner

Sixty-seven percent (67%) of the projects surveyed were completed in the private sector. Figure 4 compares the proportion of projects completed in the public and private sectors.

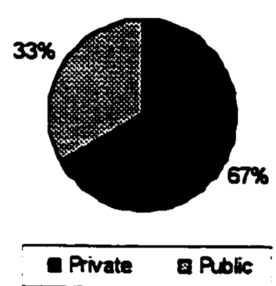


Figure 4. Identity of Project Owner

Duration of Design Contract

Of the projects that were reported, most had design contracts that had durations of less than one year. However, one-third (33%) had contracts that extended longer than one year. The shortest design project lasted three months and the longest was thirty-six months. The range of project durations is displayed in Figure 5.

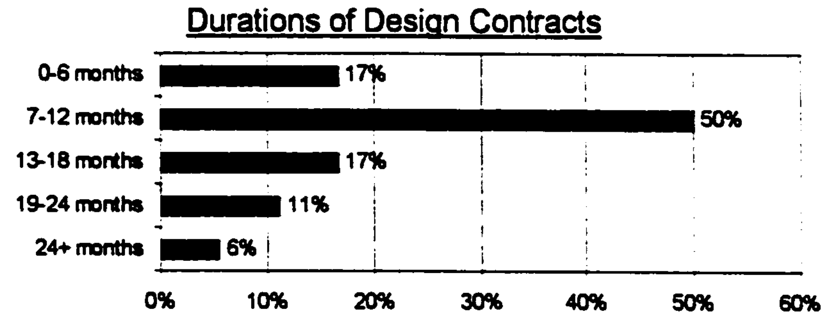


Figure 5. Durations of Design Contracts

Value of Design Contract

The total value of the design contracts ranged from $60,000 to $240,000,000. This excellent coverage of a wide variety of sizes of industrial design contracts ensured that the survey produced results that were representative of a large part of the industry. Most of the projects surveyed had design contract values of less than $1,000,000. The design contract values of the projects are shown in Figure 6.

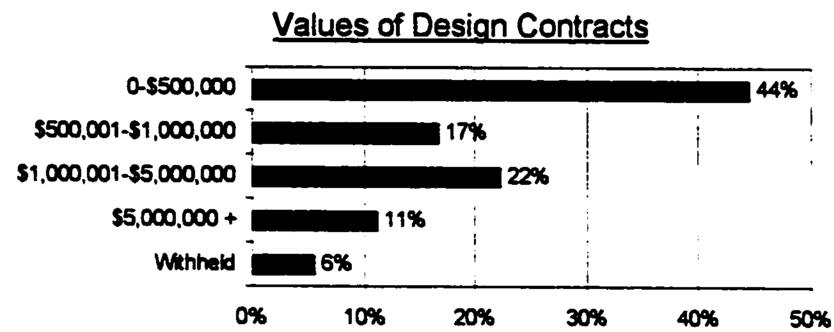


Figure 6. Values of Design Contracts

Role of Designer

Most of the responding design firms were the prime consultant on the project they selected. In fact, only twenty-two percent (22%) of respondents were sub-consultants on their project. The proportion of respondents that were prime and sub-consultants is shown in Figure 7.

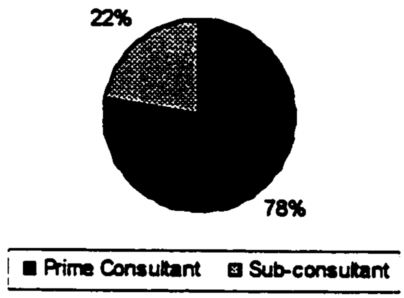


Figure 7. Role of Designer

Duration of Construction Contract

Generally, the construction contracts of the projects selected had durations of less than one year. However, the range of durations spanned from three months to two years. For firms that reported a project that was design only, this section may not be applicable or may be unknown by the respondent. The durations of the projects are listed in Figure 8.

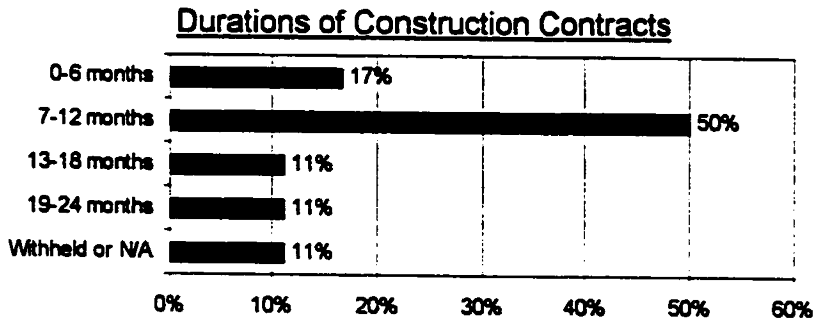


Figure 8. Durations of Design Contracts

Value of Construction Contract

Forty-four percent (44%) of the surveyed construction contracts had values in the one to ten million-dollar range. In addition, several construction contracts exceeded ten million dollars and a few had values less than one million dollars. Once again, if the scope of the project selected was design only, this section may not be applicable or may be unknown by the respondent. The range of construction contract values is shown in Figure 9.

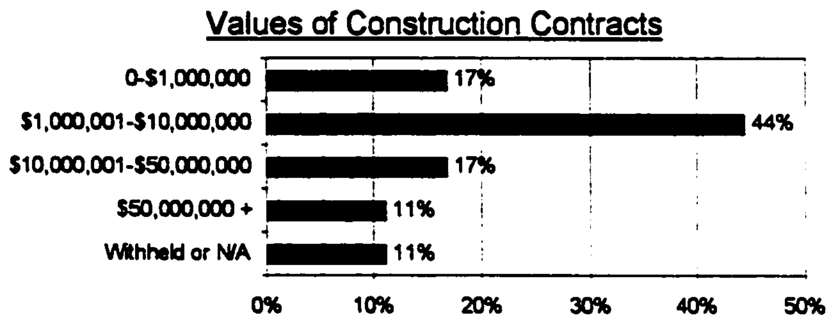


Figure 9. Values of Construction Contracts

1.3. Project Information

Type of Project

A list of seven different areas of industrial activity was included in the survey. Some projects fell into these predefined categories, but many respondents filled in a project type that was not already specified on the questionnaire. The questionnaire was only intended to survey design projects in the industrial sector; therefore, each new category provided by a responding company was reviewed to ensure the project fit within the target population. If it did fall into the industrial sector, the new category was then added to the list of project types. A complete list of types of projects that were chosen is shown in Table 2.

Table 2. Types of Projects Chosen by Companies Surveyed

Project Type	Number
Chemical Processing or Extraction Plant	2
Mining	2
Oil and Gas - Pipeline	1
Oil and Gas - Refinery or Compressor Station	2
Water/Waste Water Treatment Plant	3
Other: Bulk materials handling/transportation	1
Other: Highway	1
Other: Landfill	1
Other: Manufacturing	1
Other: Offices and Warehouse	1
Other: Port/Marine Terminal	1
Other: Vehicle Maintenance Facility, Shop and Office	1
Other: Water Supply Line, Wells, Pumps, Reservoir	1

Type of Design Contract

The types of design contracts used for the projects selected by the respondents included virtually all the common types used in the construction industry today. The most common among them was a cost plus arrangement, which was the contract type for thirty-nine percent (39%) of the projects reported. The next most common type was a lump sum design contract, used on twenty-eight percent (28%) of the projects. The various types of design contract types are shown in Figure 10.

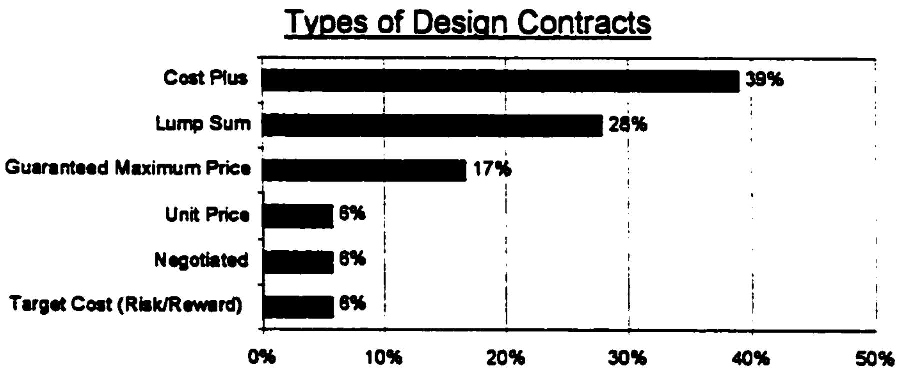


Figure 10. Types of Design Contracts

Scope of Design Contract

Respondents to the survey mainly chose projects that had a design only scope or a scope of design contract that included design and management. Several of the design contracts included field services or procurement services as well. A comparison of the different scopes of design contracts is shown in Figure 11.

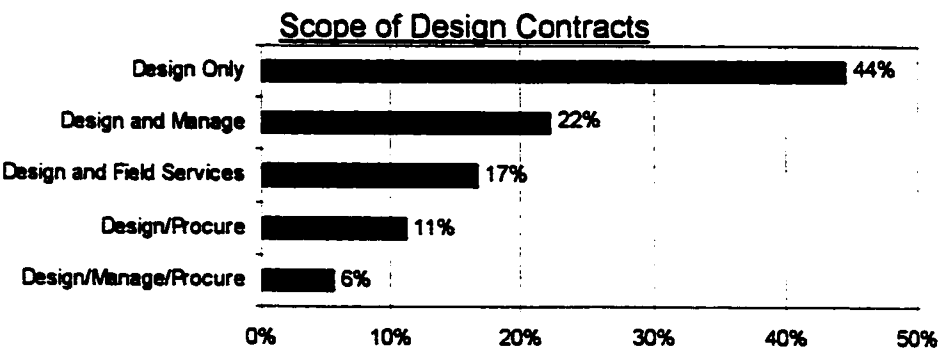


Figure 11. Scope of Design Contracts

Type of Construction Contract

Thirty-nine percent (39%) of the construction contracts were lump sum contracts, however unit price, guaranteed maximum price and cost plus contracts were also each selected on over ten percent of the projects. Some design companies that had design only contracts found this section not applicable to their selected project. A complete list of construction contracts is shown in Figure 12.

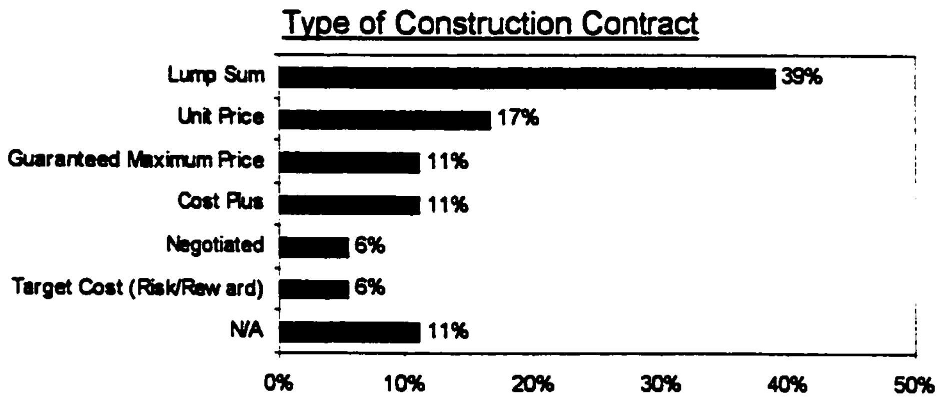


Figure 12. Types of Construction Contracts

Scope of Construction Contract

The majority (61%) of survey responses indicated a construct only contract for construction. Other contract types, in smaller numbers, were also reported. All of the construction contract scopes are listed in Figure 13.

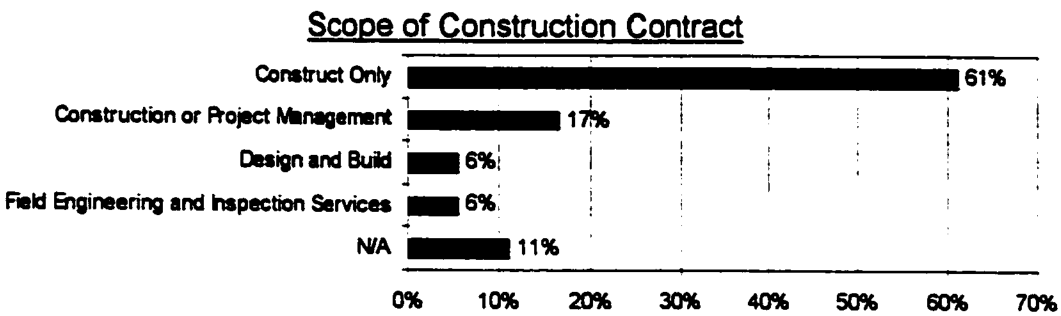


Figure 13. Scope of Construction Contracts

Method of Tendering for Construction

Of the responses given, fifty percent (50%) of construction contracts were open competitive tenders. The next most common method of tendering was prequalification of bidders (28%). The breakdown of the methods of tendering is shown in Figure 14.

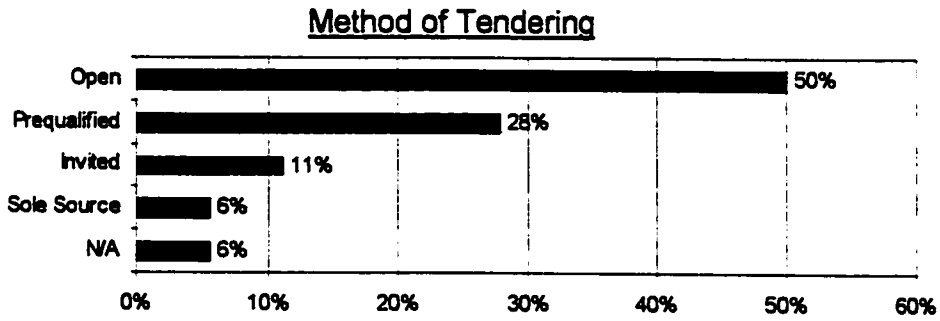


Figure 14. Methods of Tendering for the Construction Contracts

Project Priorities

In this section of the survey, respondents were asked to rank the relative importance of a list of eight priorities with one (1) being the highest ranking and eight (8) the ranking of the lowest priority. Table 3 gives each priority an overall ranking by averaging the sum of the rankings of each priority over the total number of responses. In Table 4, a method of ranking weighted priorities by calculating an importance factor uses the formula:

$$I = \Sigma\,(9-w)\cdot(n/N)\cdot(100/8)$$

where:
I = the importance factor
w = a constant expressing the weighting given to each response. The weighting ranges from 1 to 8, where 1 is the highest importance and 8 is the least importance.
n = the frequency of response, for each weighting
N = the total number of responses = the total sample size for the question
$(100/8)$ = a factor to convert the weighting into a percentage

Both of the two methods of determining the importance of the project priorities result in identical overall rankings of each individual priority, with cost being the most important.

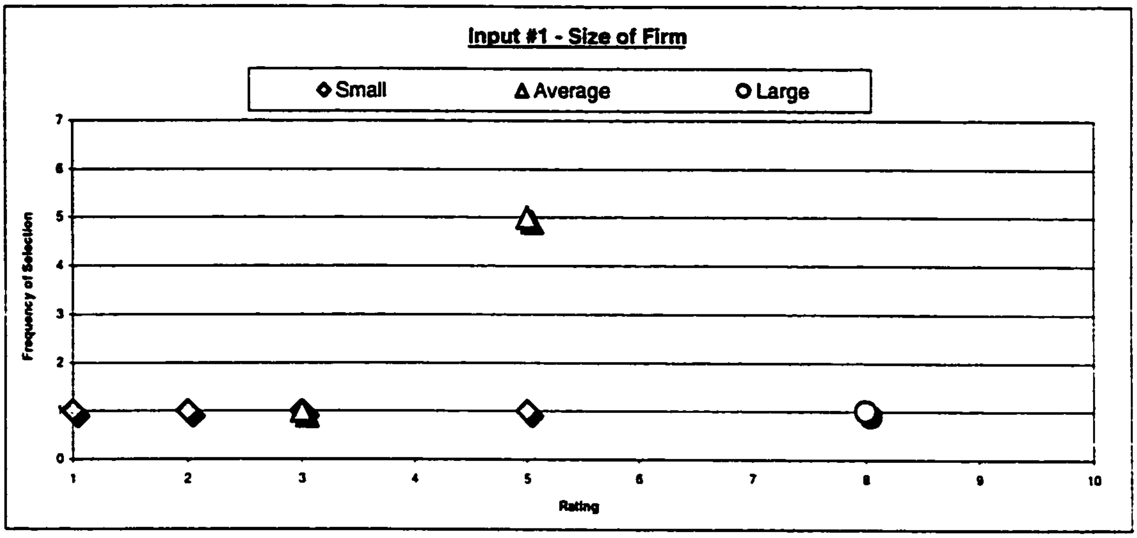
Table 3. Ranking of Project Priorities

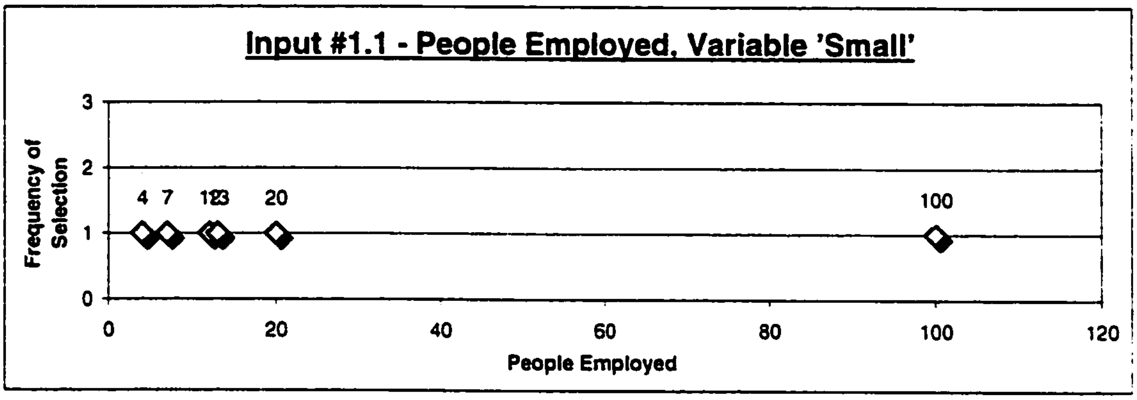
Priority	Total Score	# Of Responses	Average Importance [illegible]	Overall Ranking
Cost	32	16	2.00	1
Quality	52	16	3.25	2
Safety	56	16	3.50	3
Schedule	61	16	3.81	4
Constructability	67	16	4.19	5
Environmental Impact	78	16	4.88	6
Aesthetics	108	16	6.75	7
Future Development	113	16	7.06	8

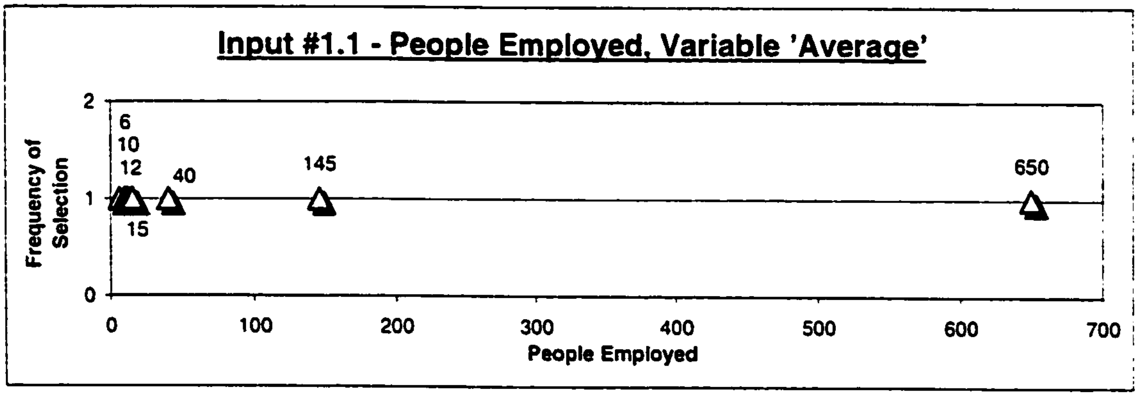
Table 4. Importance Factor for Each Project Priority

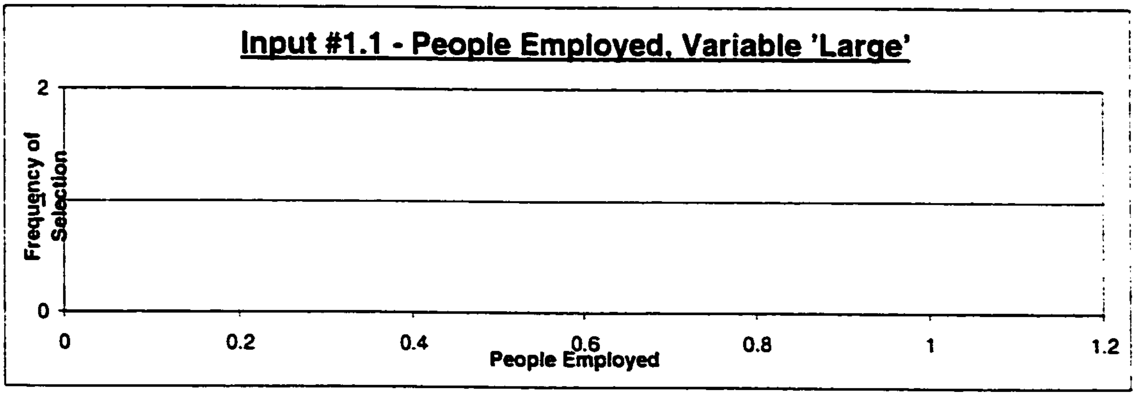
Priority	Number of Responses	Frequency of Selection of Each Ranking								Importance Index	Overall Ranking
		1	2	3	4	5	6	7	8		
Cost	16	8	4	2	1	0	1	0	0	87.5%	1
Quality	16	0	4	6	5	0	1	0	0	71.9%	2
Safety	16	4	1	2	5	1	2	1	0	68.8%	3
Schedule	16	2	4	2	1	3	2	2	0	64.8%	4
Constructability	16	1	2	3	3	2	4	1	0	60.2%	5
Environmental Impact	16	2	1	0	1	6	3	2	1	51.6%	6
Aesthetics	16	0	0	1	0	3	1	4	7	28.1%	7
Future Development	16	0	0	0	0	2	2	5	7	24.2%	8

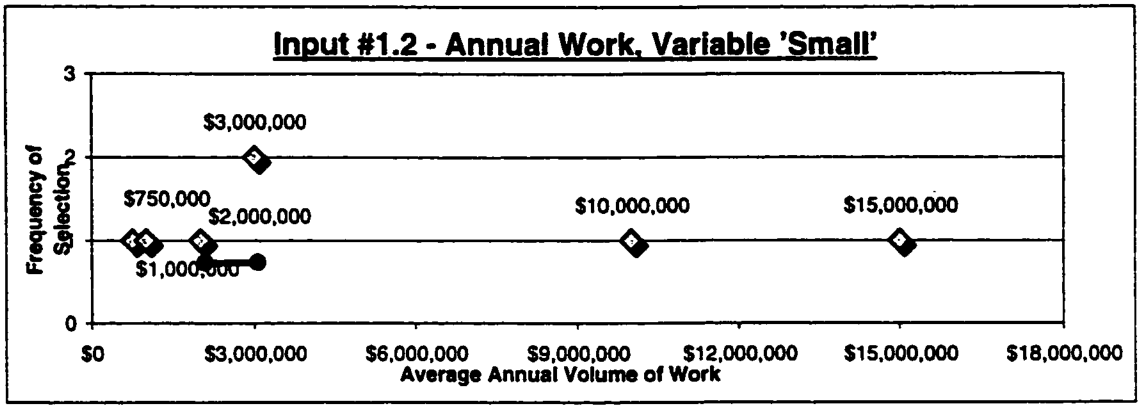
Appendix 3: Data Processing Results (Trial 1)

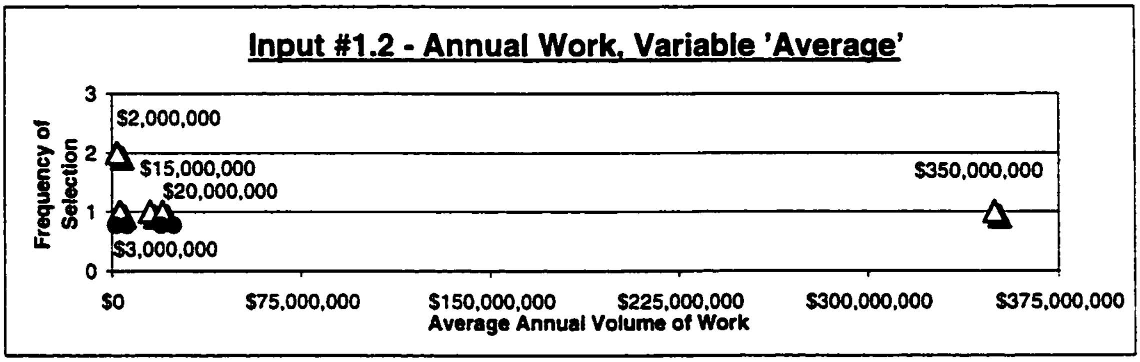


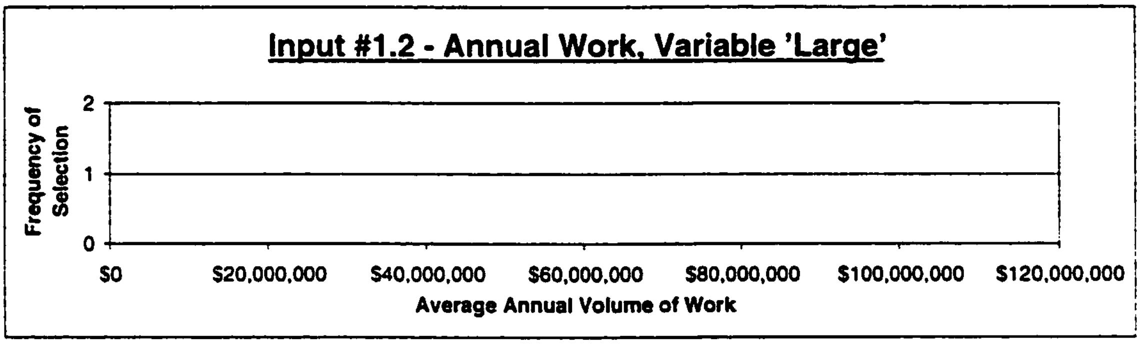


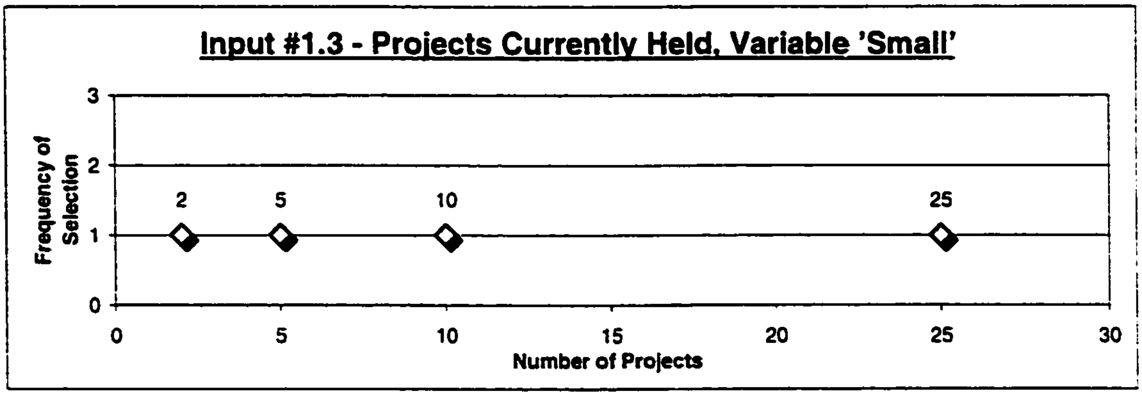


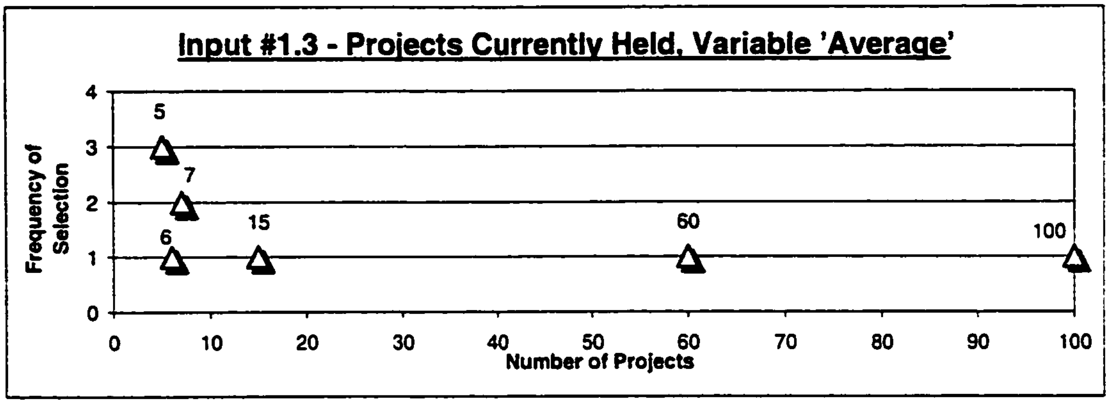


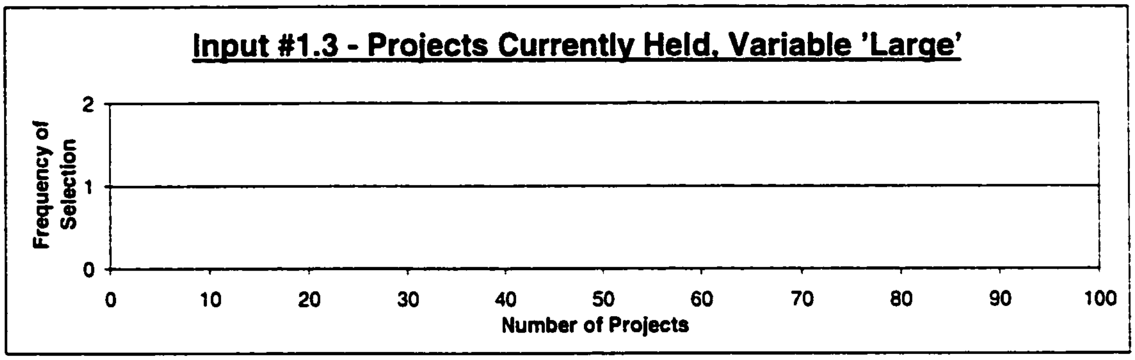

Input #1.2 - Annual Work, Variable 'Small'
Frequency of Selection
3
2
0
$3,000,000
$750,000
$2,000,000
$10,000,000
$15,000,000
$1,000,000
$0
$3,000,000
$6,000,000
$9,000,000
$12,000,000
$15,000,000
$18,000,000
Average Annual Volume of Work

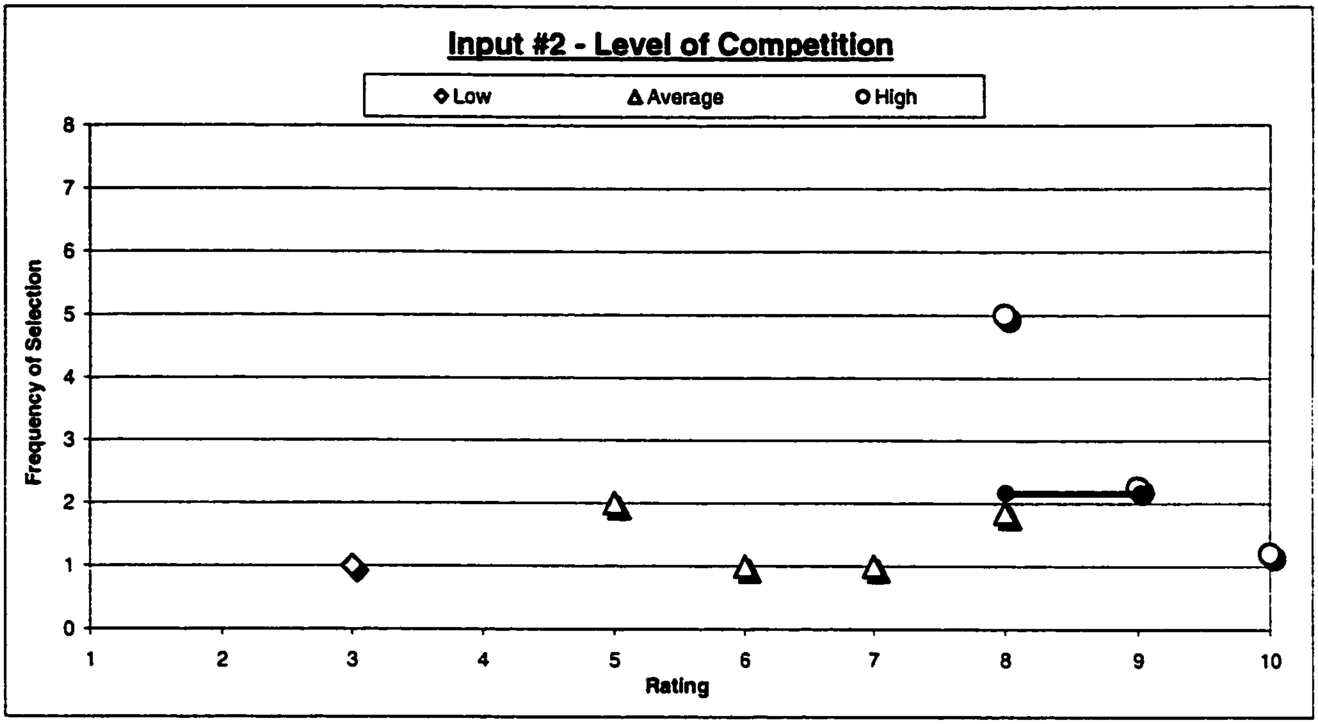

Input #1.2 - Annual Work, Variable 'Average'
Frequency of Selection
3
2
1
0
$2,000,000
$15,000,000
$20,000,000
$3,000,000
$350,000,000
$0
$75,000,000
$150,000,000
$225,000,000
$300,000,000
$375,000,000
Average Annual Volume of Work

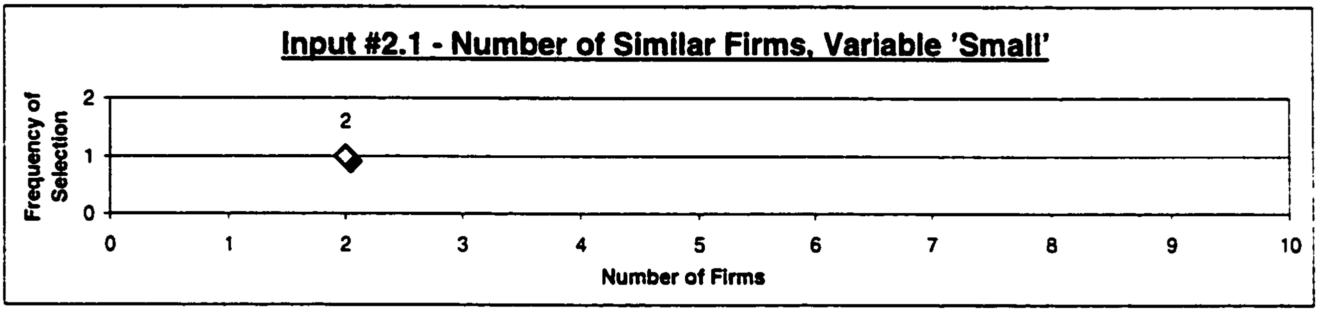

Input #1.2 - Annual Work, Variable 'Large'
Frequency of Selection
2
1
0
$0
$20,000,000
$40,000,000
$60,000,000
$80,000,000
$100,000,000
$120,000,000
Average Annual Volume of Work

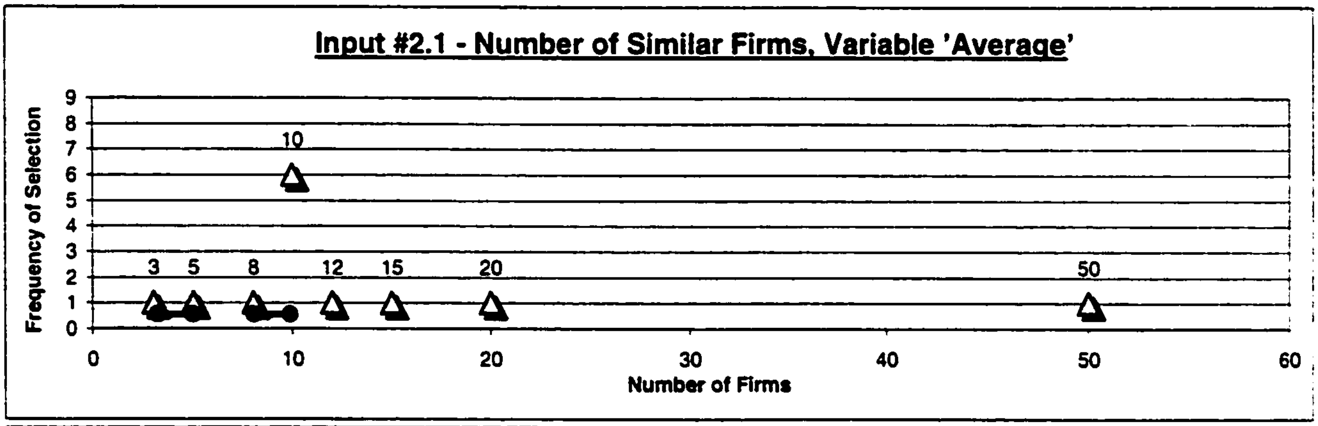

Input #1.3 - Projects Currently Held, Variable 'Small'
Frequency of Selection
3
2
1
0
2
5
10
25
0
5
10
15
20
25
30
Number of Projects

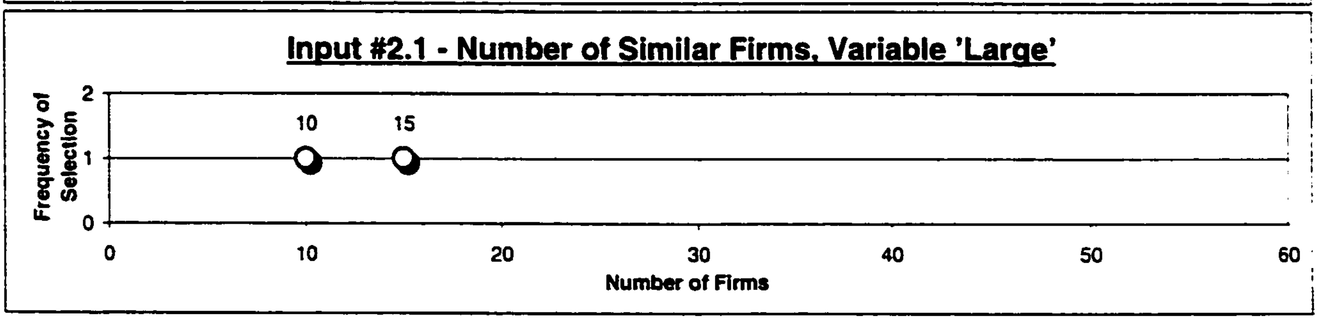

Input #1.3 - Projects Currently Held, Variable 'Average'
Frequency of Selection
Number of Projects
5
7
6
15
60
100
0
1
2
3
4
0
10
20
30
40
50
60
70
80
90
100

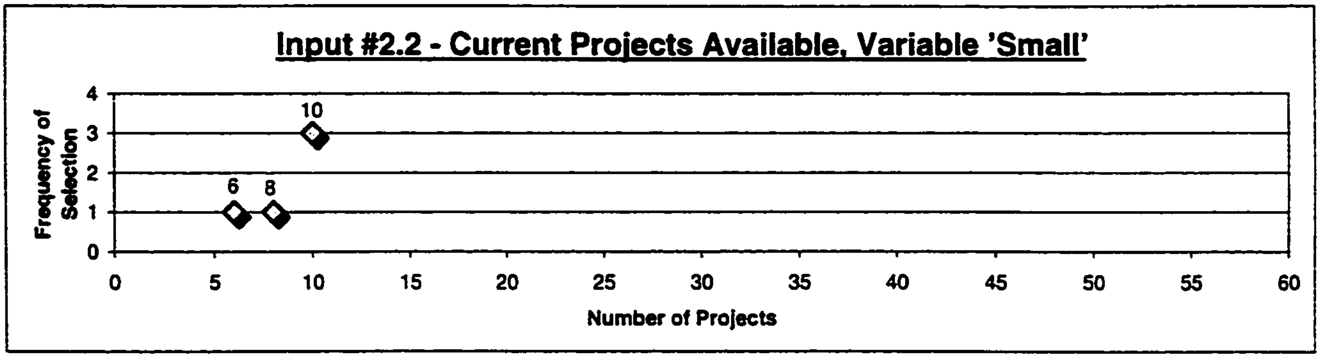

Input #1.3 - Projects Currently Held, Variable 'Large'
Frequency of Selection
Number of Projects
0
1
2
0
10
20
30
40
50
60
70
80
90
100

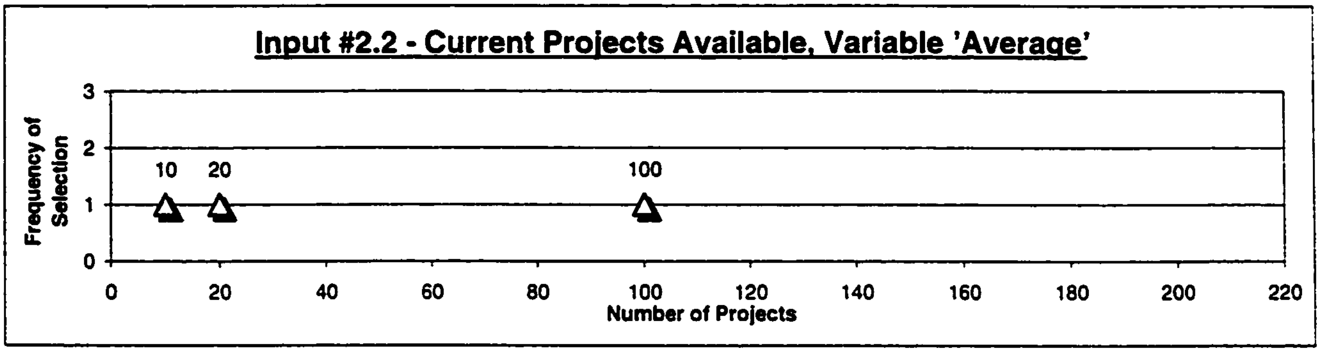

Input #2 - Level of Competition
Low
Average
High
Frequency of Selection
Rating
0
1
2
3
4
5
6
7
8
1
2
3
4
5
6
7
8
9
10

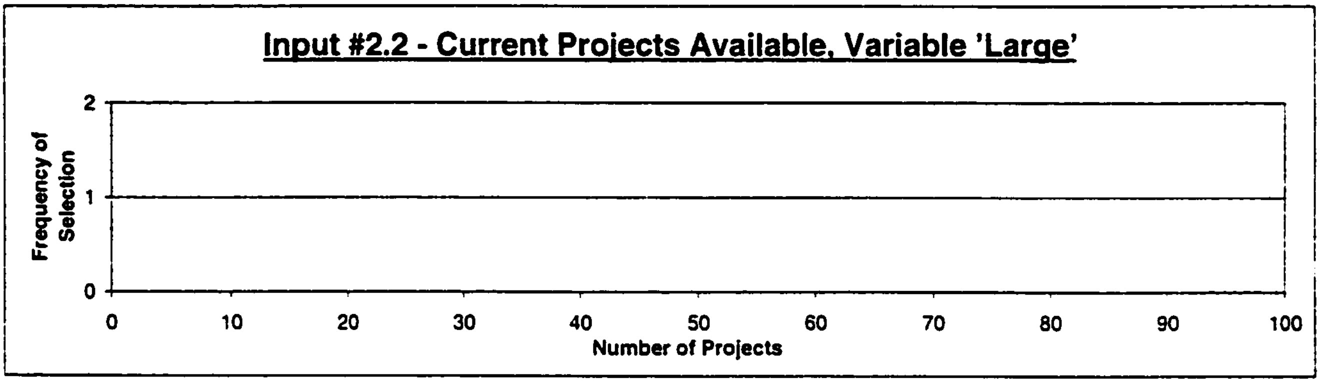

Input #2.1 - Number of Similar Firms, Variable 'Small'
Frequency of Selection
Number of Firms
2
0
1
2
0
1
2
3
4
5
6
7
8
9
10

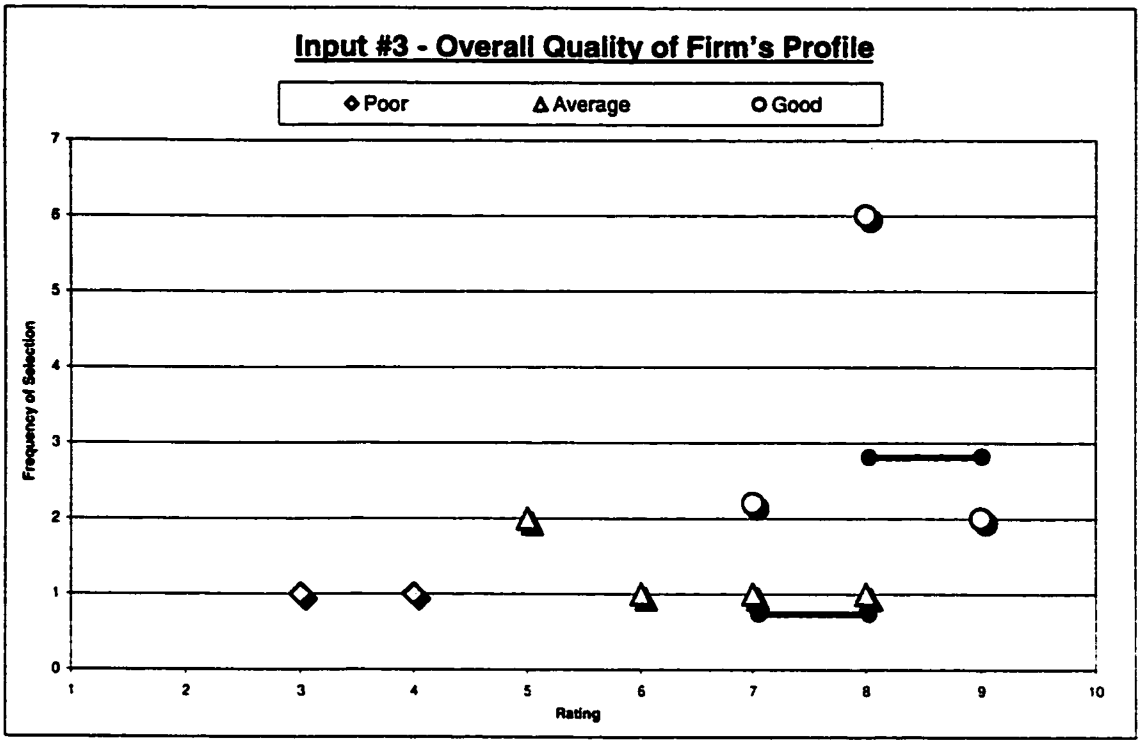

Input #2.1 - Number of Similar Firms, Variable 'Average'
Frequency of Selection
Number of Firms
10
3
5
8
12
15
20
50
0
1
2
3
4
5
6
7
8
9
0
10
20
30
40
50
60

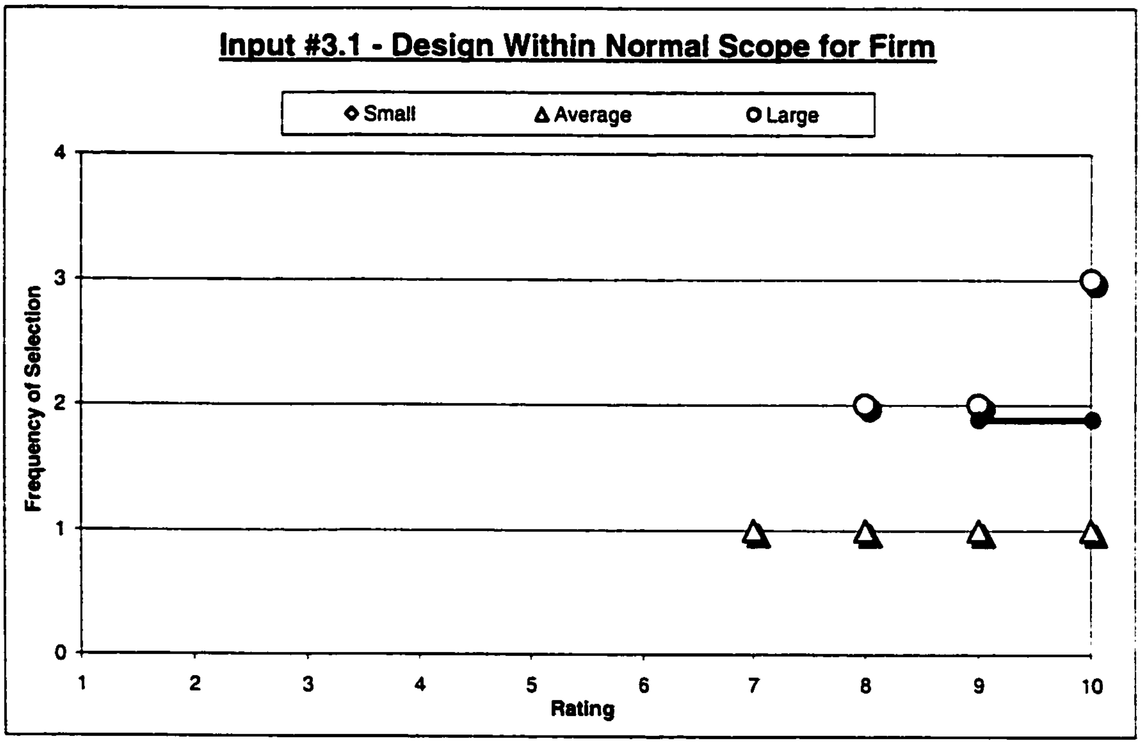

Input #2.1 - Number of Similar Firms, Variable 'Large'
Frequency of Selection
Number of Firms
10
15
0
1
2
0
10
20
30
40
50
60

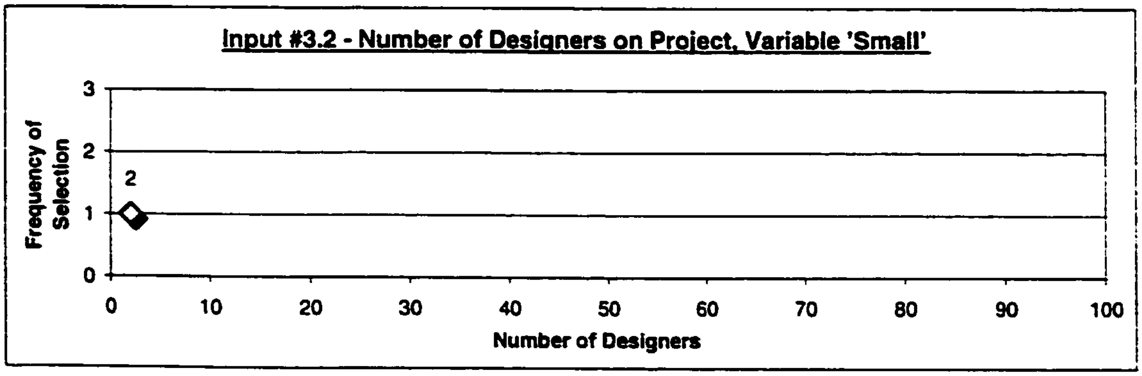

Input #2.2 - Current Projects Available, Variable 'Small'
Frequency of Selection
Number of Projects
10
6
8
0
1
2
3
4
0
5
10
15
20
25
30
35
40
45
50
55
60

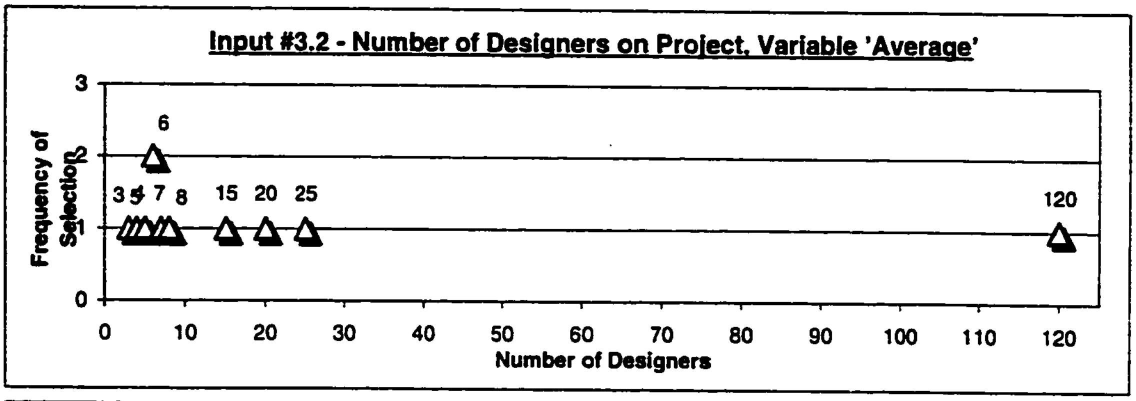

Input #2.2 - Current Projects Available, Variable 'Average'
Frequency of Selection
Number of Projects
10
20
100
0
1
2
3
0
20
40
60
80
100
120
140
160
180
200
220

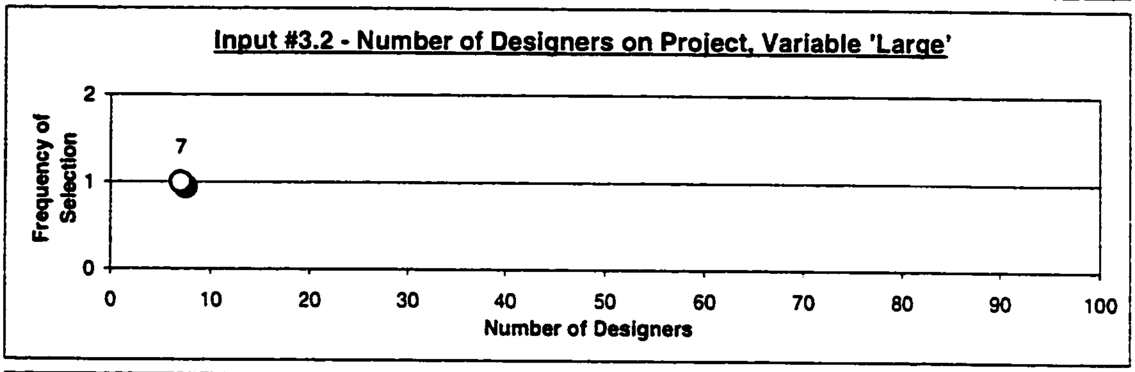

Input #2.2 - Current Projects Available, Variable 'Large'
Frequency of Selection
Number of Projects
0
1
2
0
10
20
30
40
50
60
70
80
90
100

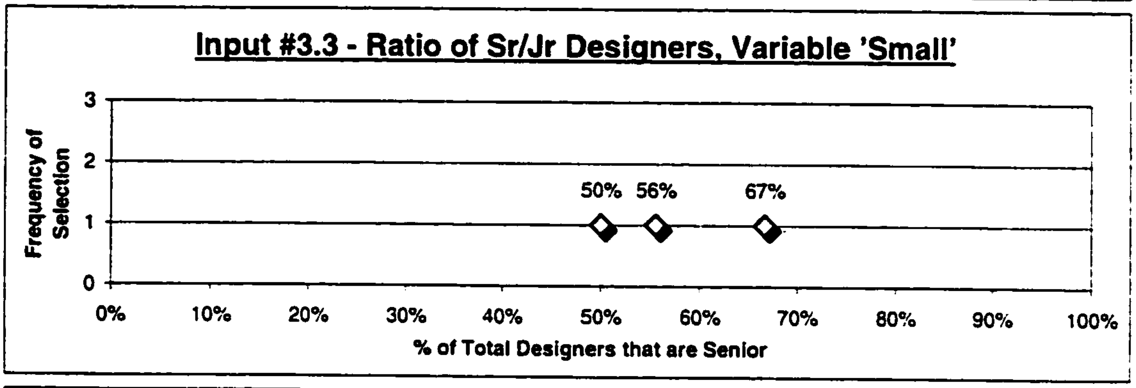

Input #3 - Overall Quality of Firm's Profile
Poor
Average
Good
Frequency of Selection
Rating

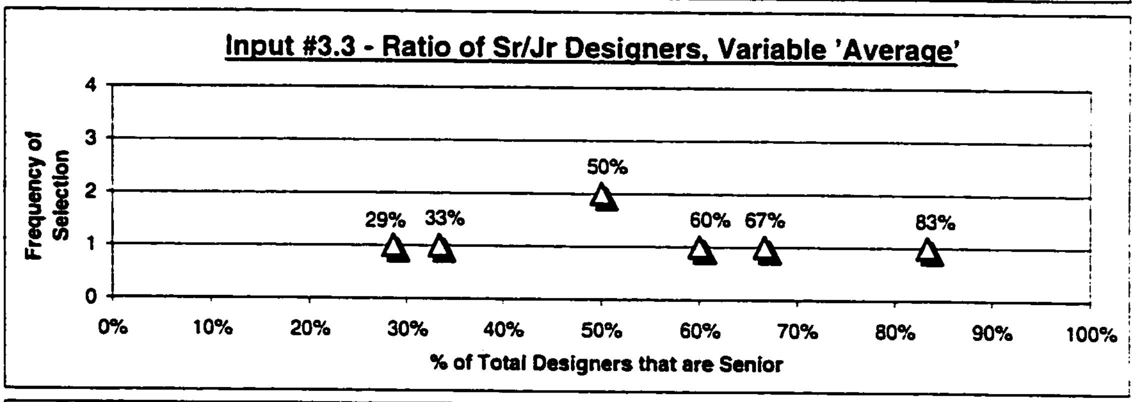

Input #3.1 - Design Within Normal Scope for Firm
Small
Average
Large
Frequency of Selection
Rating

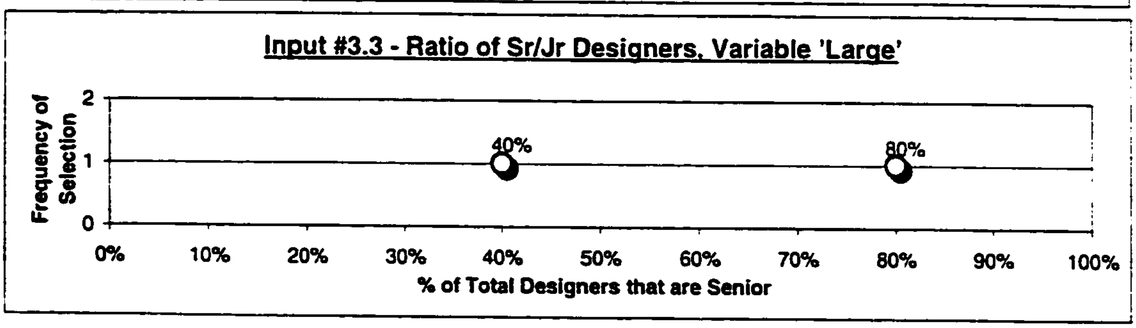

Input #3.2 - Number of Designers on Project, Variable 'Small'
Frequency of Selection
2
Number of Designers

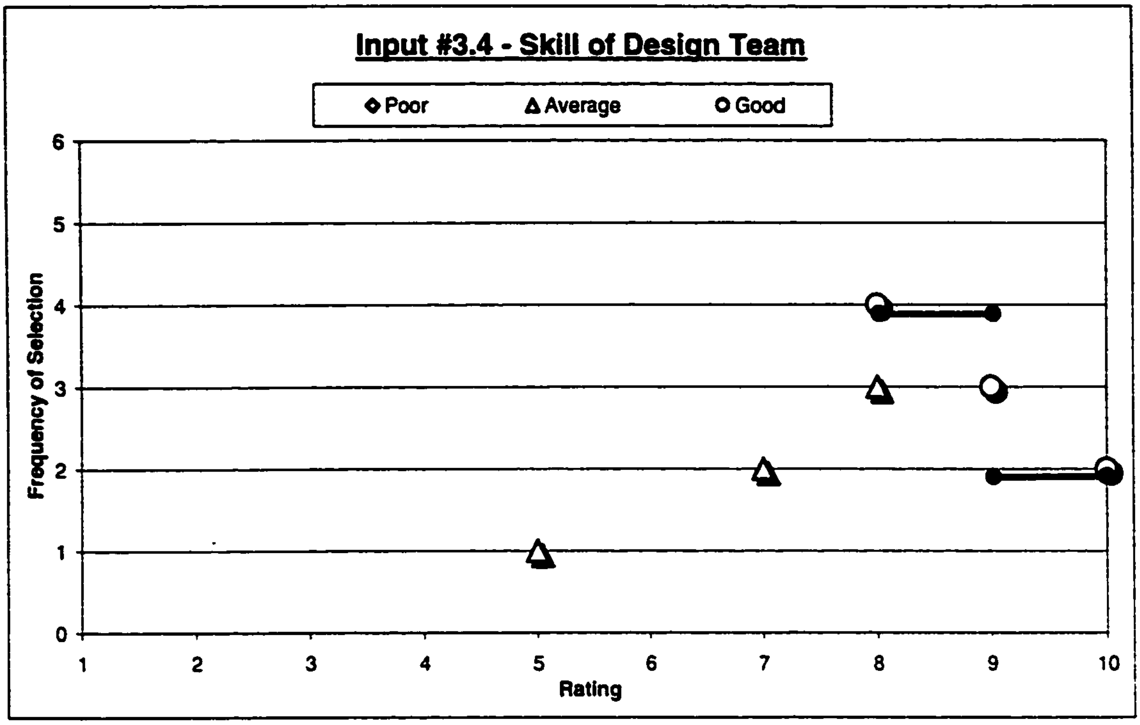

Input #3.2 - Number of Designers on Project, Variable 'Average'
Frequency of Selection
Number of Designers
6
3 5 4 7 8 15 20 25
120

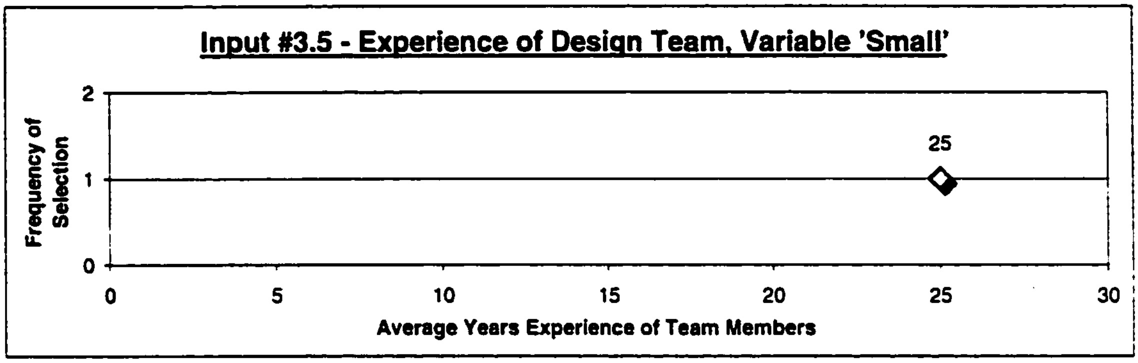

Input #3.2 - Number of Designers on Project, Variable 'Large'
Frequency of Selection
Number of Designers
7

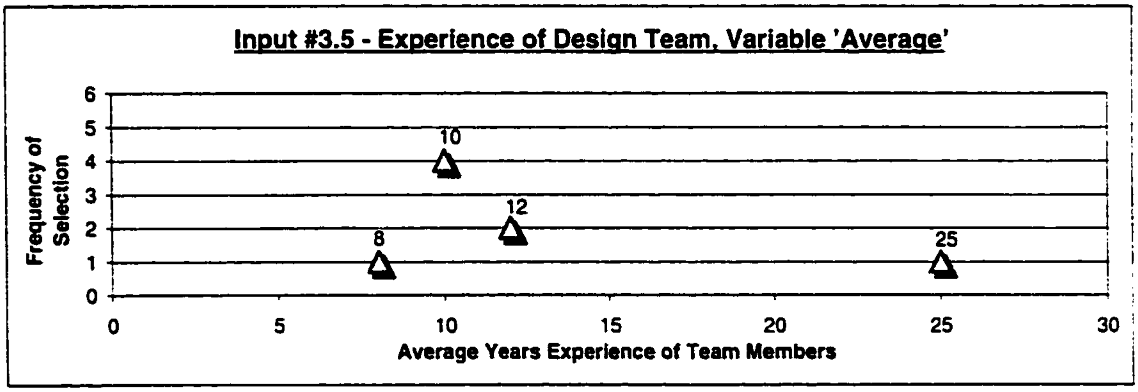

Input #3.3 - Ratio of Sr/Jr Designers, Variable 'Small'
Frequency of Selection
% of Total Designers that are Senior
50% 56% 67%

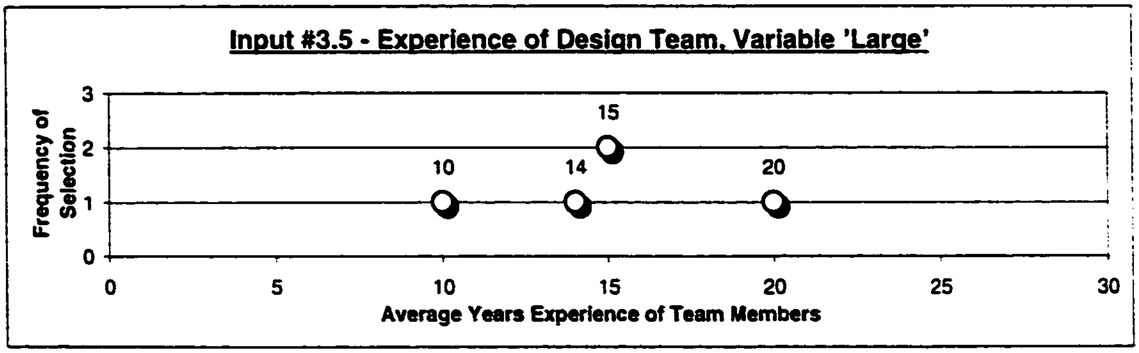

Input #3.3 - Ratio of Sr/Jr Designers, Variable 'Average'
Frequency of Selection
% of Total Designers that are Senior
50%
29% 33% 60% 67% 83%

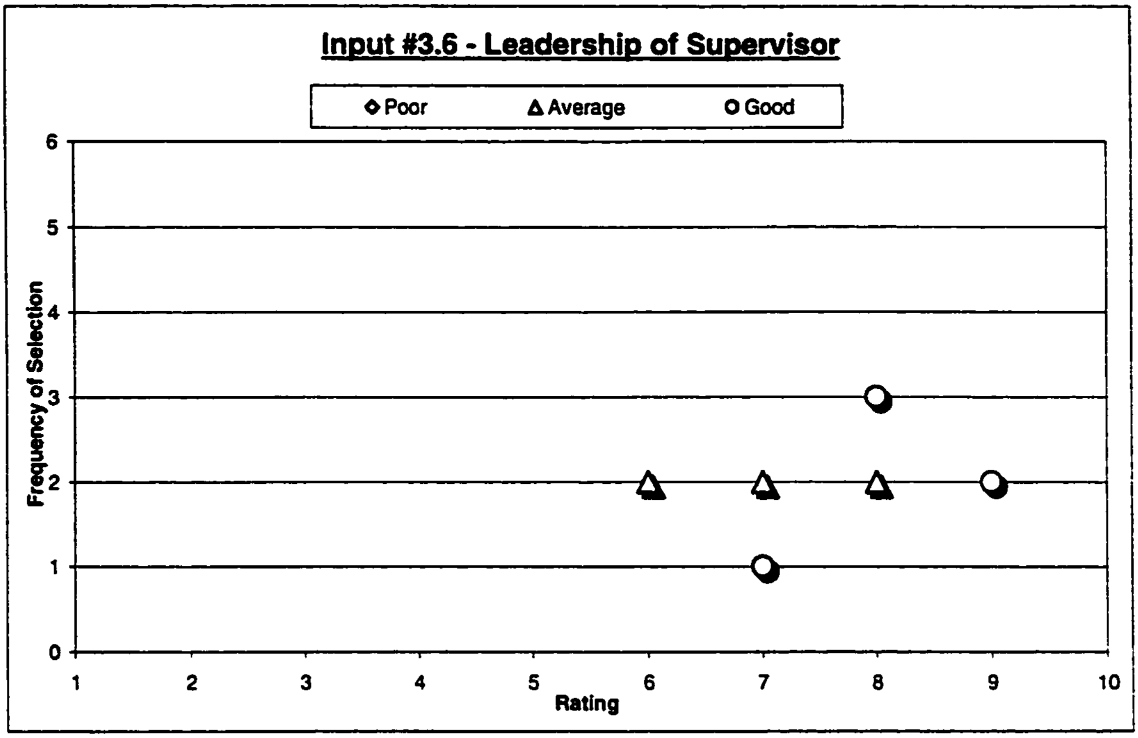

Input #3.3 - Ratio of Sr/Jr Designers, Variable 'Large'
Frequency of Selection
% of Total Designers that are Senior
40% 80%

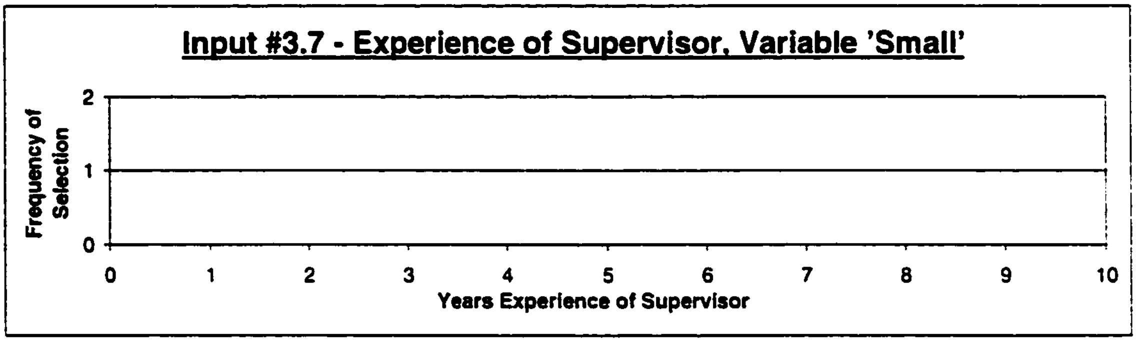

Input #3.4 - Skill of Design Team
Poor
Average
Good
Frequency of Selection
0
1
2
3
4
5
6
1
2
3
4
5
6
7
8
9
10
Rating

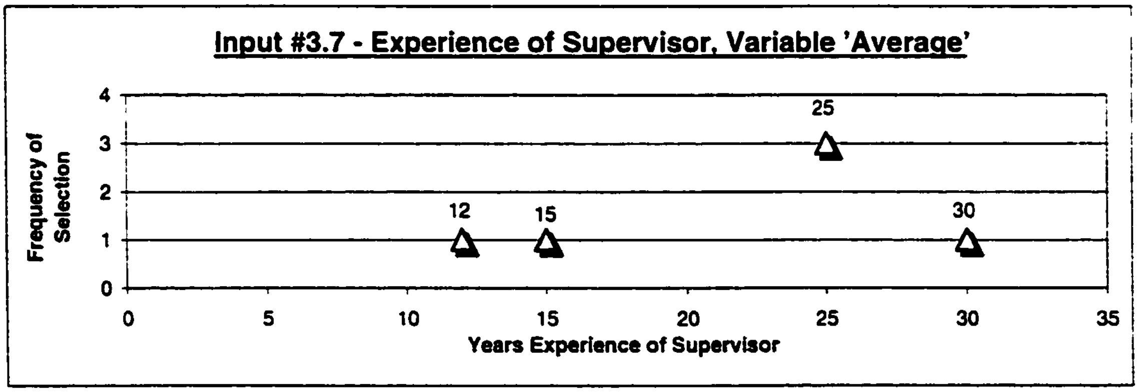

Input #3.5 - Experience of Design Team, Variable 'Small'
Frequency of Selection
0
1
2
25
0
5
10
15
20
25
30
Average Years Experience of Team Members

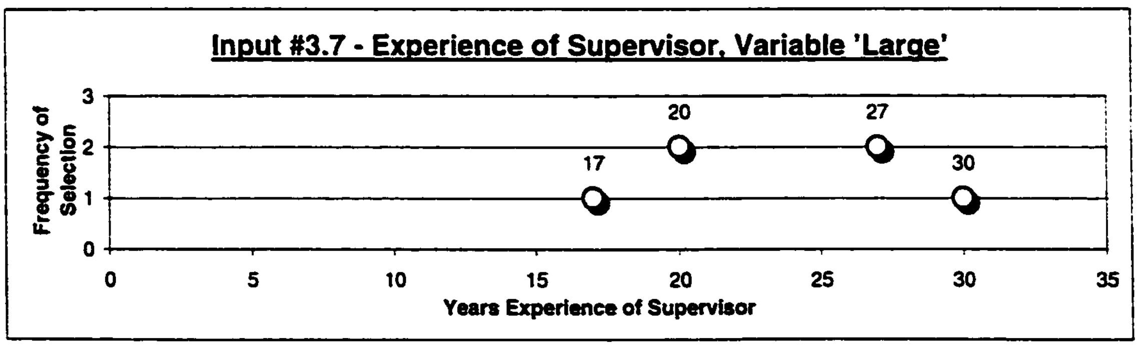

Input #3.5 - Experience of Design Team, Variable 'Average'
Frequency of Selection
0
1
2
3
4
5
6
8
10
12
25
0
5
10
15
20
25
30
Average Years Experience of Team Members

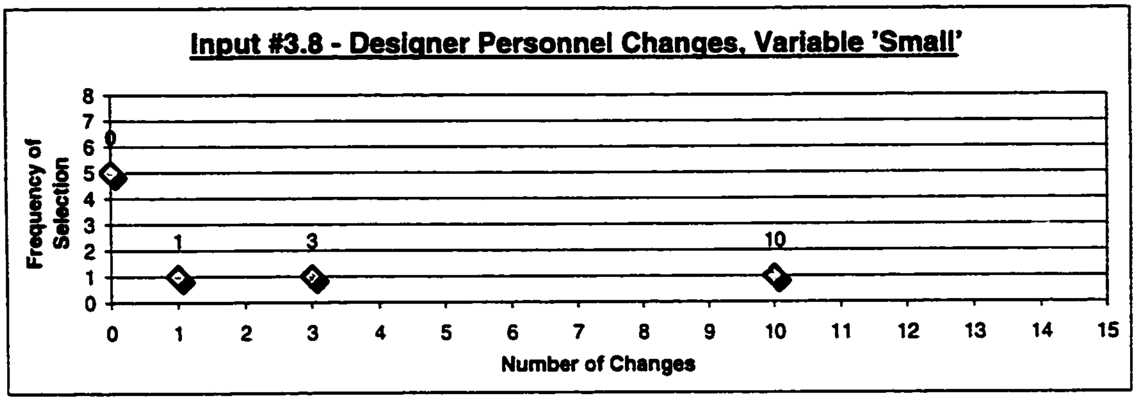

Input #3.5 - Experience of Design Team, Variable 'Large'
Frequency of Selection
0
1
2
3
10
14
15
20
0
5
10
15
20
25
30
Average Years Experience of Team Members


Input #3.6 - Leadership of Supervisor
Poor
Average
Good
Frequency of Selection
Rating

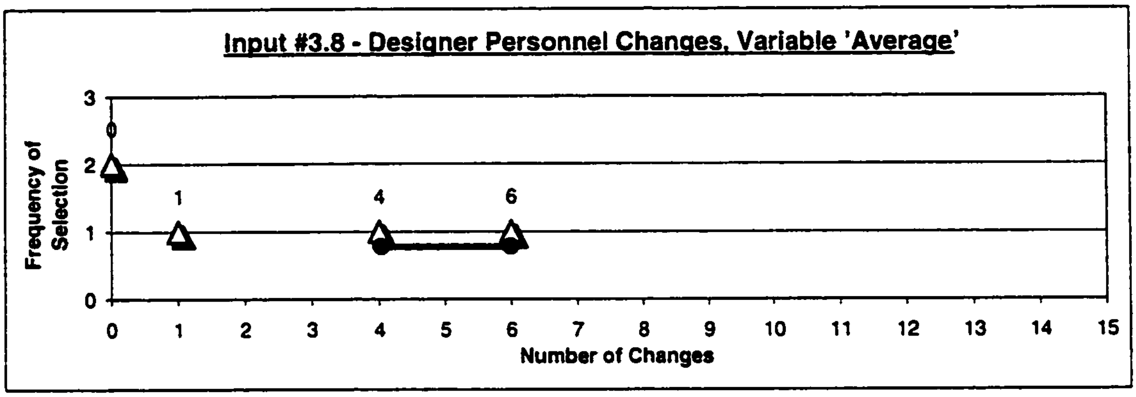

Input #3.7 - Experience of Supervisor, Variable 'Small'
Frequency of Selection
Years Experience of Supervisor


Input #3.7 - Experience of Supervisor, Variable 'Average'
Frequency of Selection
Years Experience of Supervisor

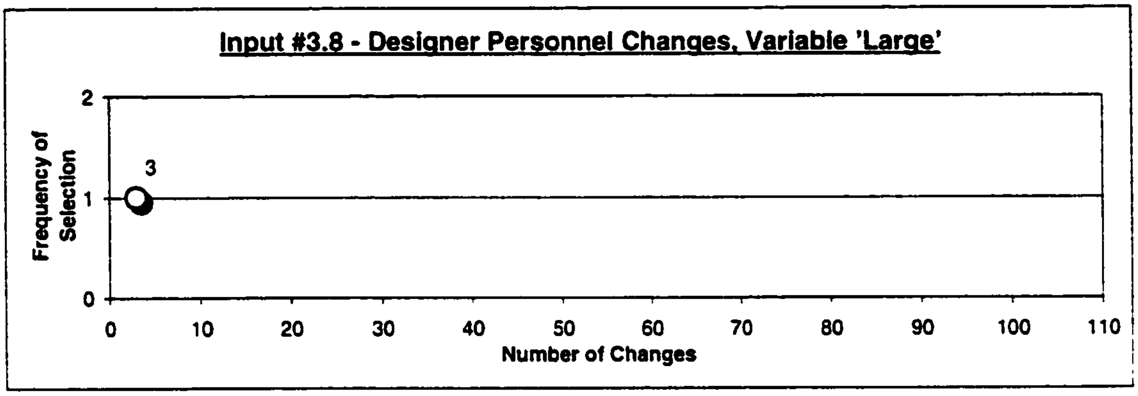

Input #3.7 - Experience of Supervisor, Variable 'Large'
Frequency of Selection
Years Experience of Supervisor



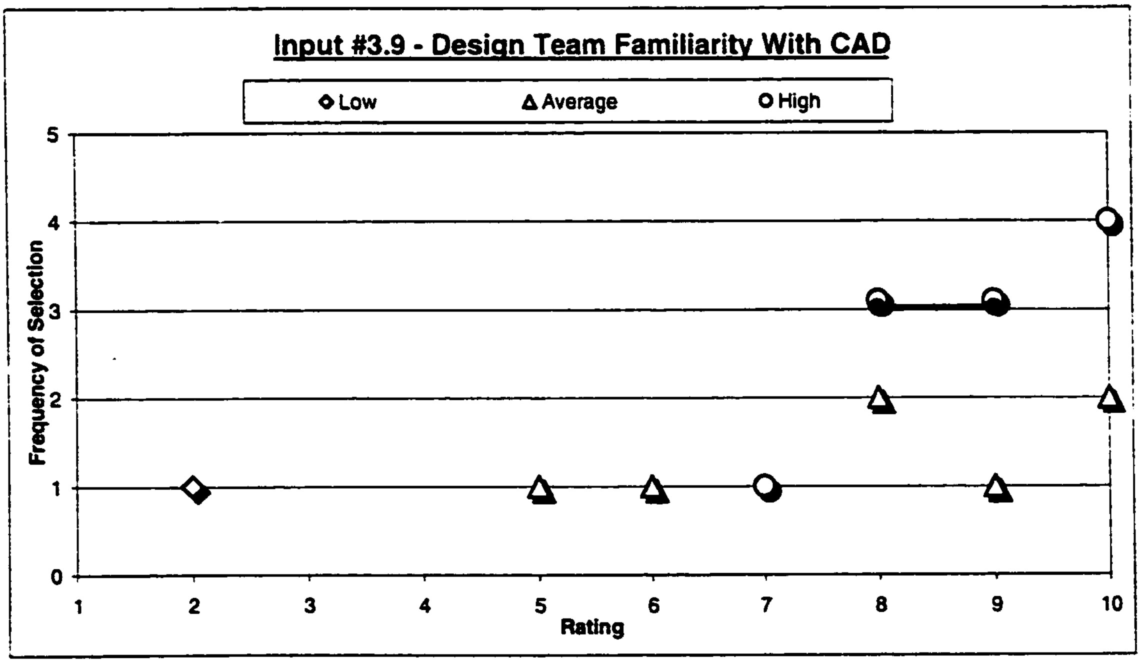
Input #3.8 - Designer Personnel Changes, Variable 'Small'
Frequency of Selection
Number of Changes
0
1
3
10



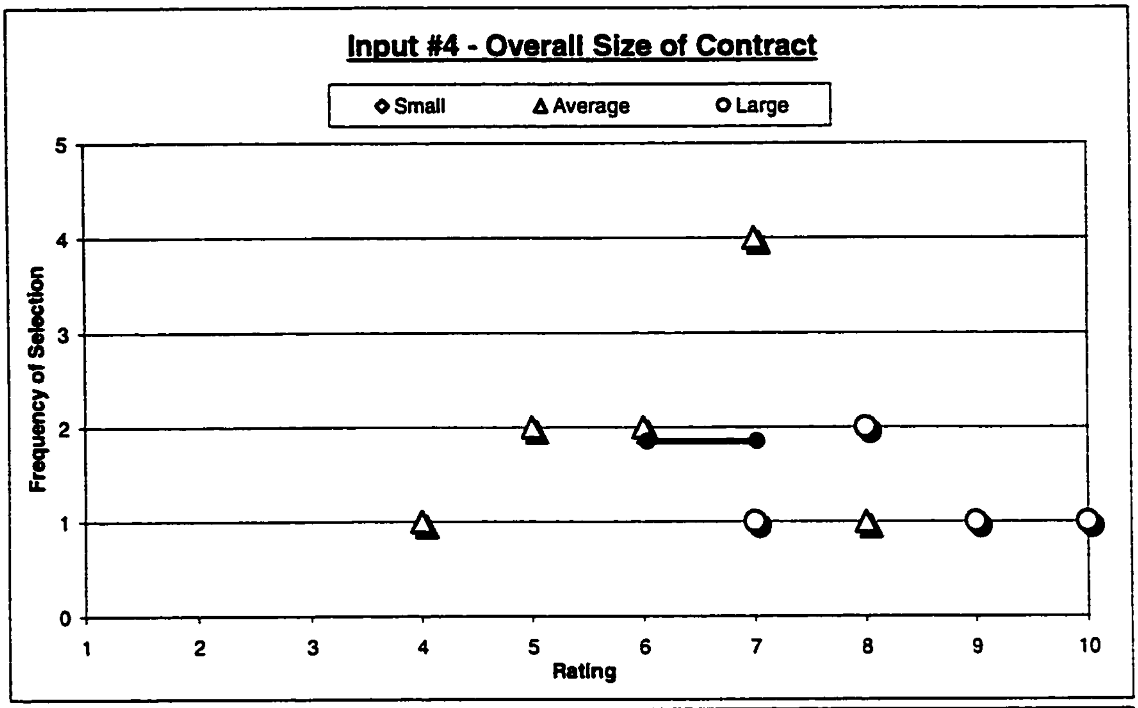
Input #3.8 - Designer Personnel Changes, Variable 'Average'
Frequency of Selection
Number of Changes
0
1
4
6



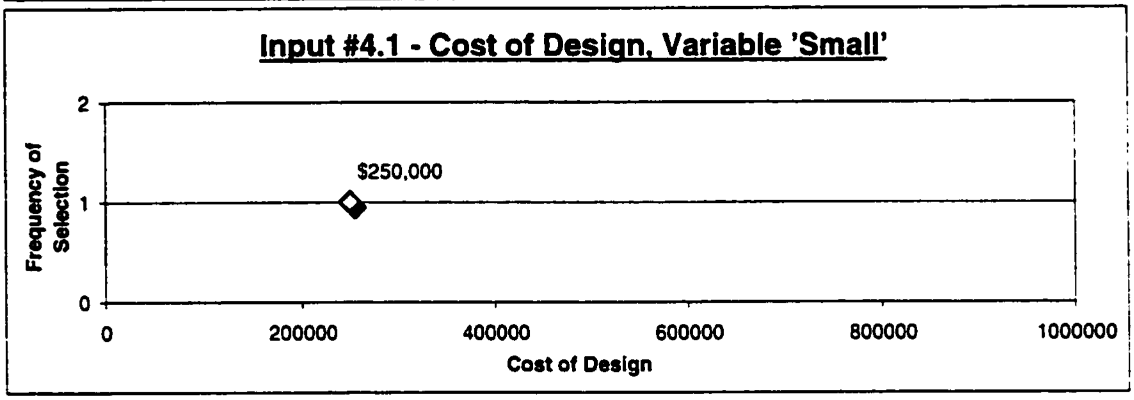
Input #3.8 - Designer Personnel Changes, Variable 'Large'
Frequency of Selection
Number of Changes
3



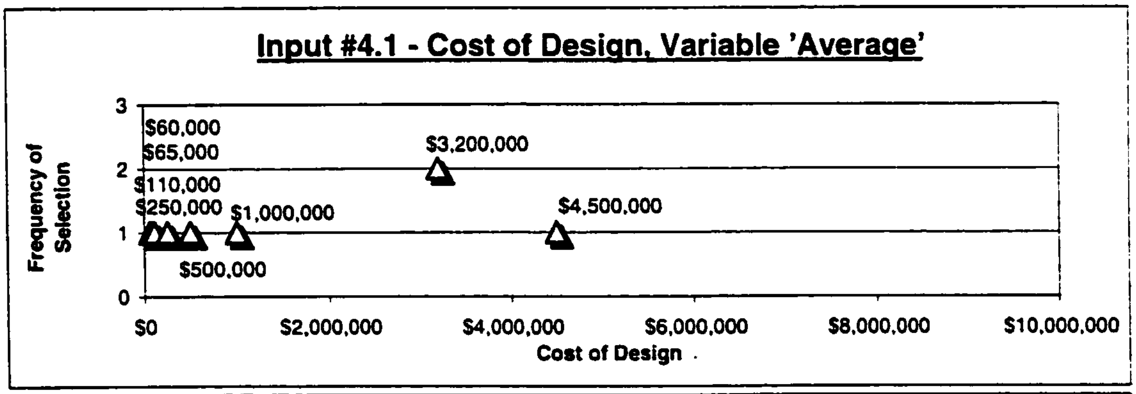
Input #3.9 - Design Team Familiarity With CAD
Low
Average
High
Frequency of Selection
Rating



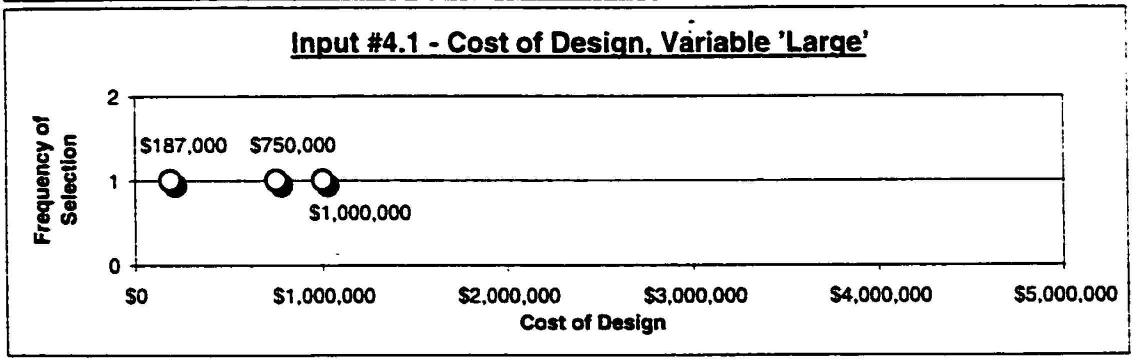
Input #4 - Overall Size of Contract
Small
Average
Large
Frequency of Selection
Rating



Input #4.1 - Cost of Design, Variable 'Small'
$250,000
Frequency of Selection
Cost of Design

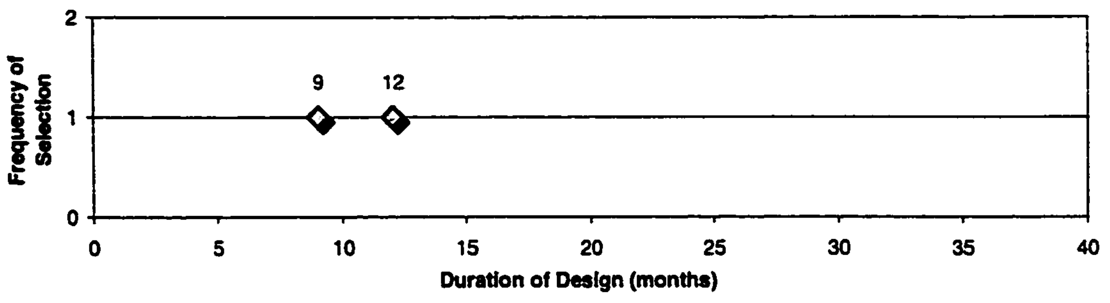


Input #4.1 - Cost of Design, Variable 'Average'
$60,000
$65,000
$110,000
$250,000
$1,000,000
$500,000
$3,200,000
$4,500,000
Frequency of Selection
Cost of Design

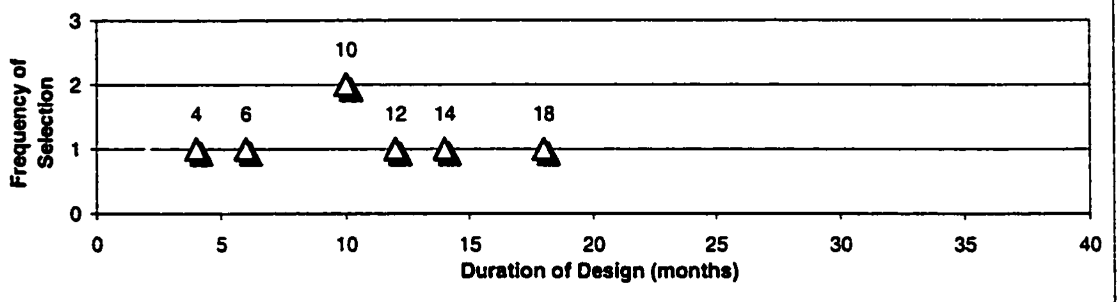


Input #4.1 - Cost of Design, Variable 'Large'
$187,000
$750,000
$1,000,000
Frequency of Selection
Cost of Design

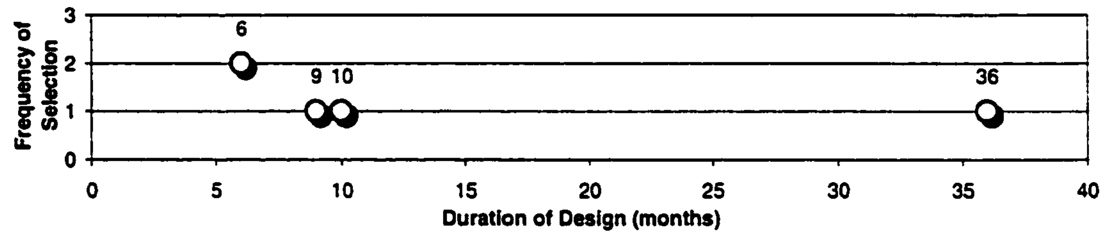
Input #4.2 - Duration of Design, Variable 'Short'



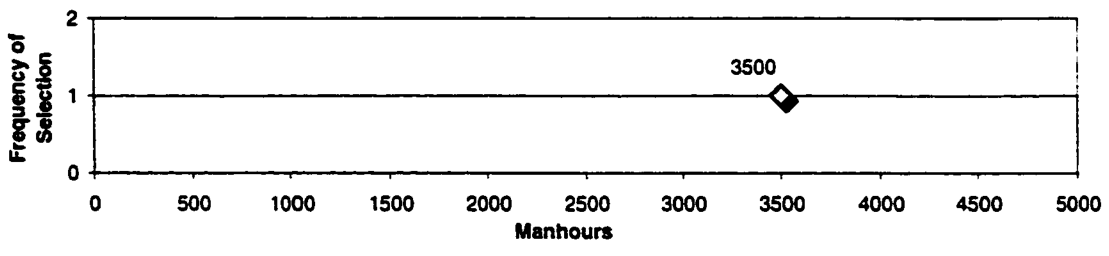
Input #4.2 - Duration of Design, Variable 'Average'



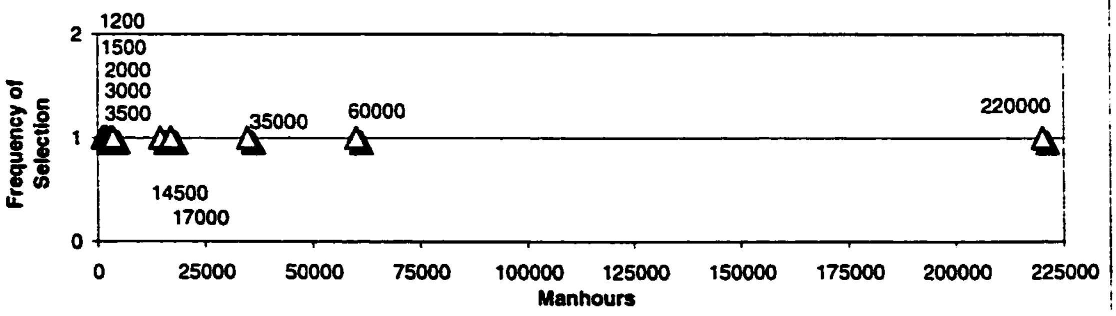
Input #4.2 - Duration of Design, Variable 'Long'

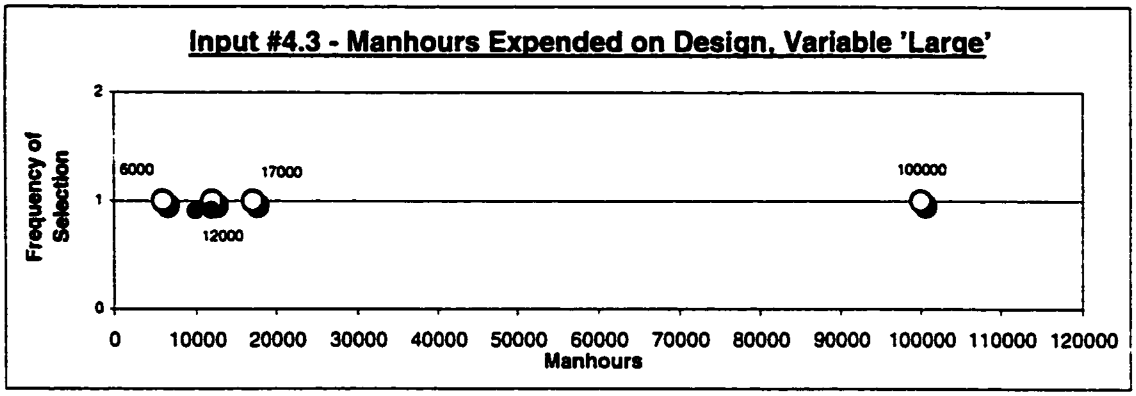


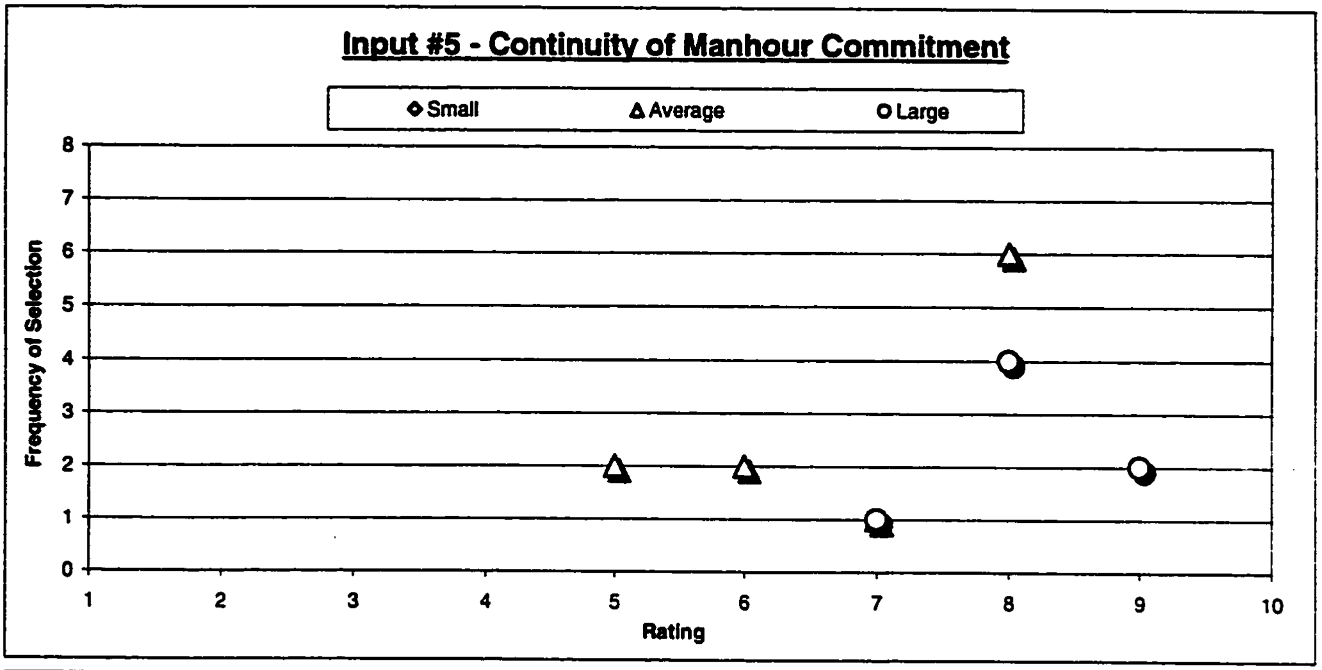
Input #4.3 - Manhours Expended on Design, Variable 'Small'



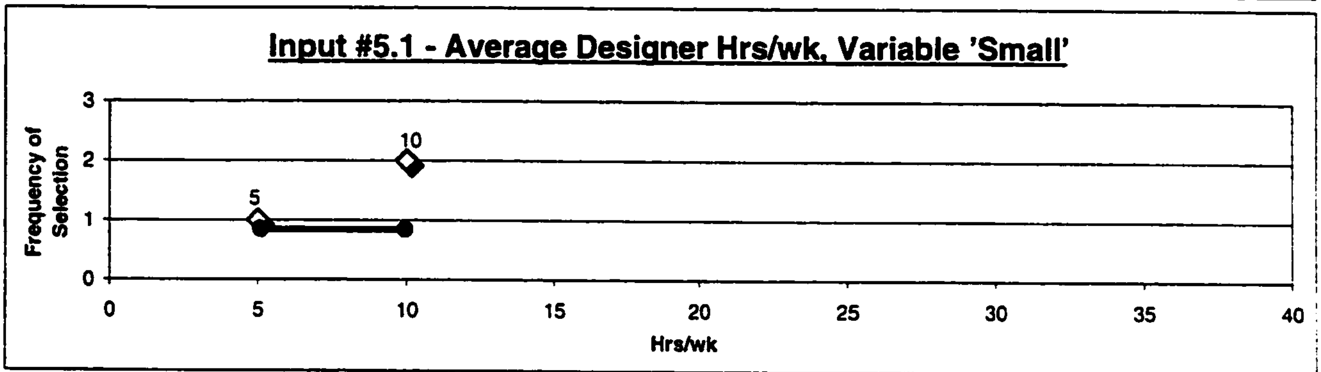
Input #4.3 - Manhours Expended on Design, Variable 'Average'



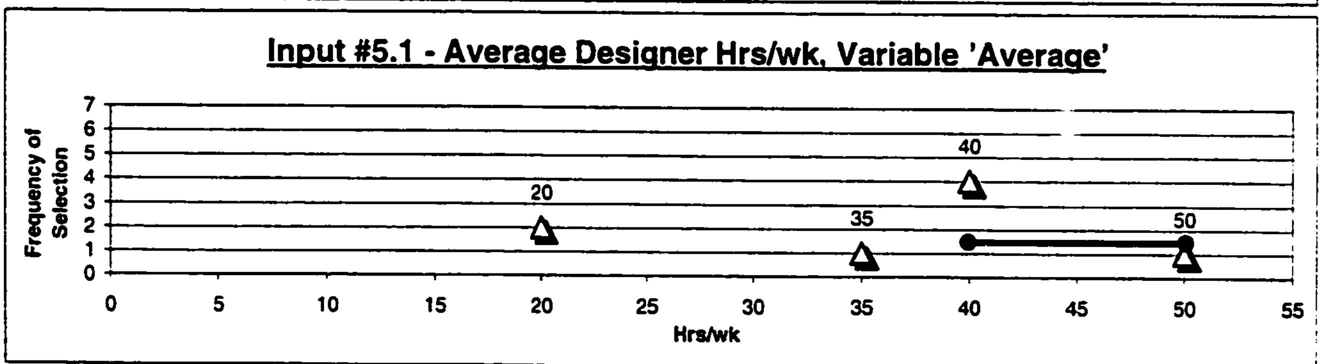

Input #4.3 - Manhours Expended on Design, Variable 'Large'
Frequency of Selection
2
1
0
6000
17000
12000
100000
0
10000
20000
30000
40000
50000
60000
70000
80000
90000
100000
110000
120000
Manhours



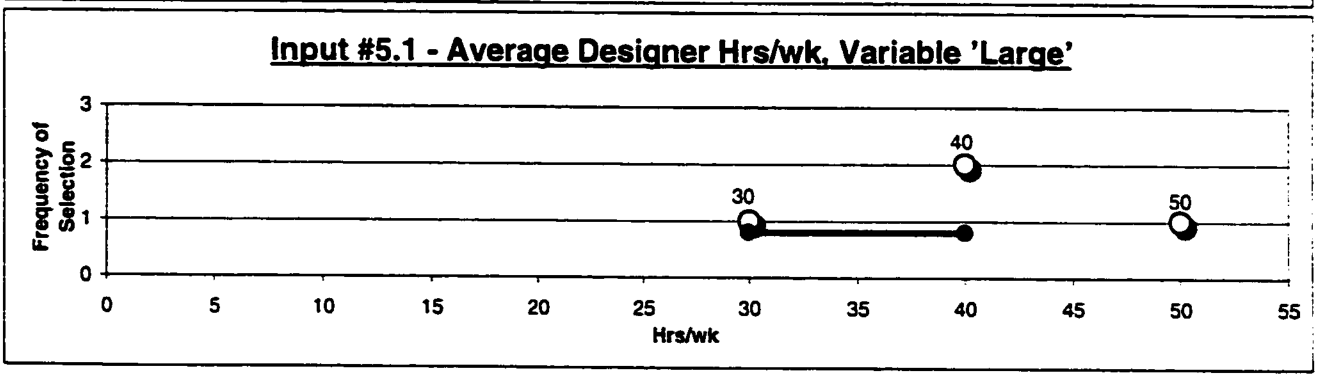
Input #5 - Continuity of Manhour Commitment
◇ Small
△ Average
○ Large
Frequency of Selection
Rating



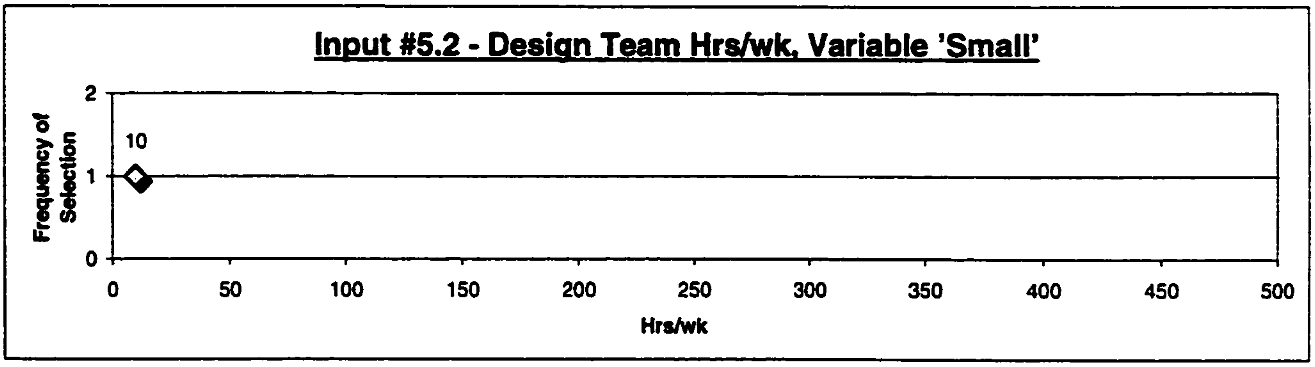
Input #5.1 - Average Designer Hrs/wk, Variable 'Small'
Frequency of Selection
Hrs/wk

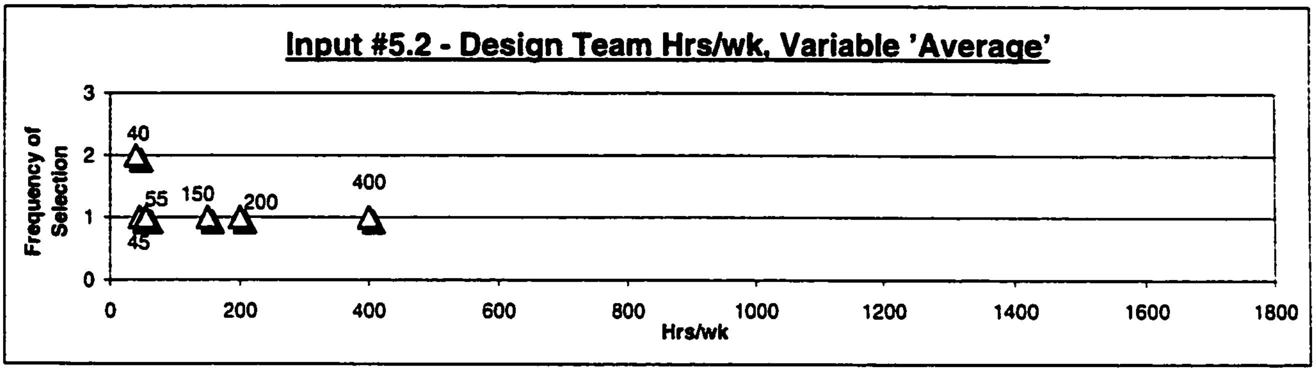


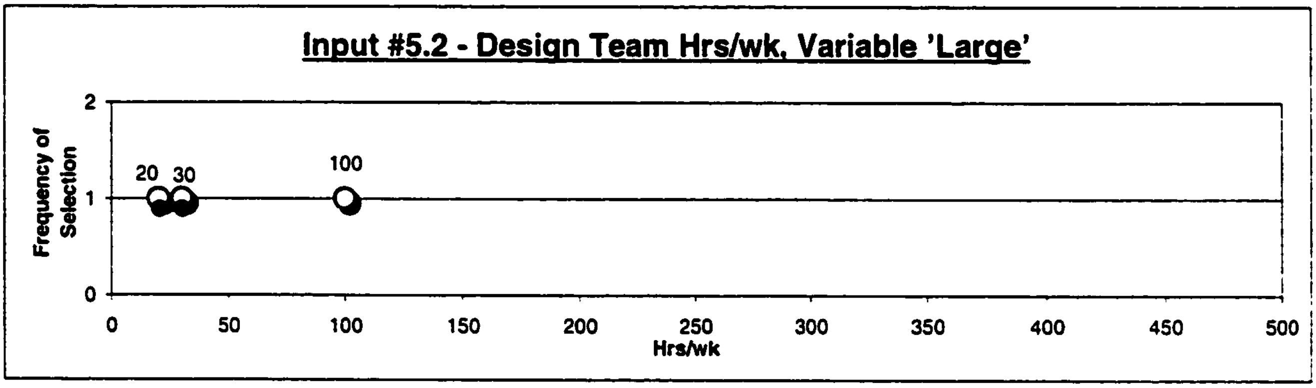
Input #5.1 - Average Designer Hrs/wk, Variable 'Average'
Frequency of Selection
Hrs/wk

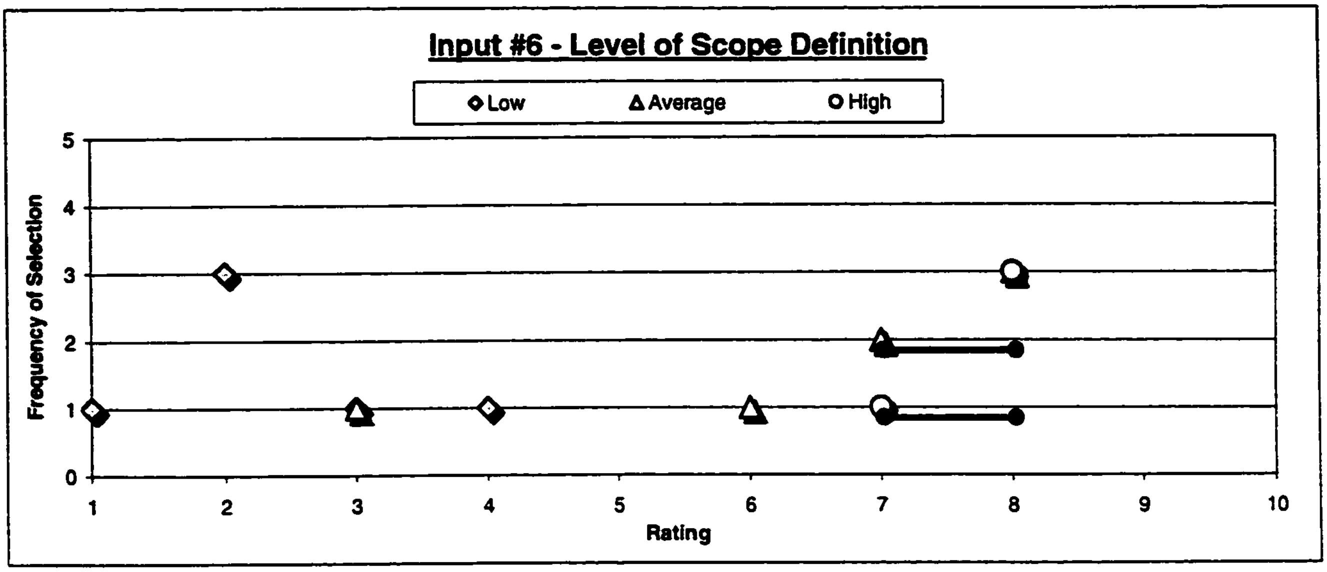


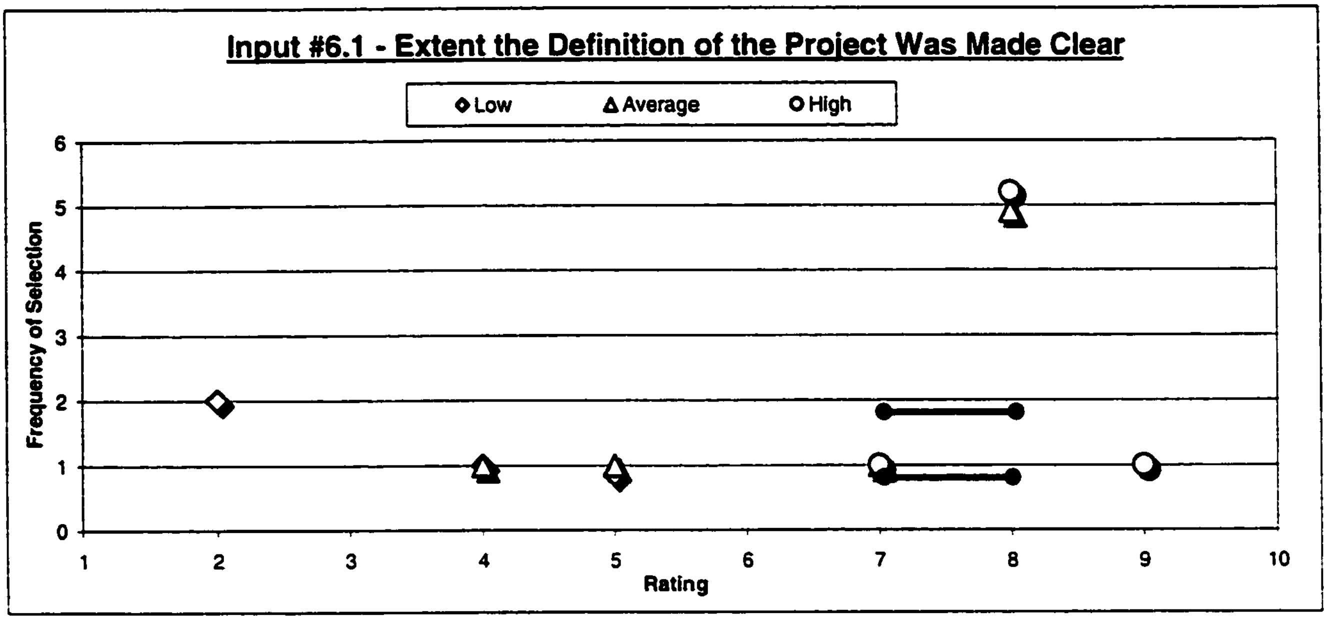
Input #5.1 - Average Designer Hrs/wk, Variable 'Large'
Frequency of Selection
Hrs/wk

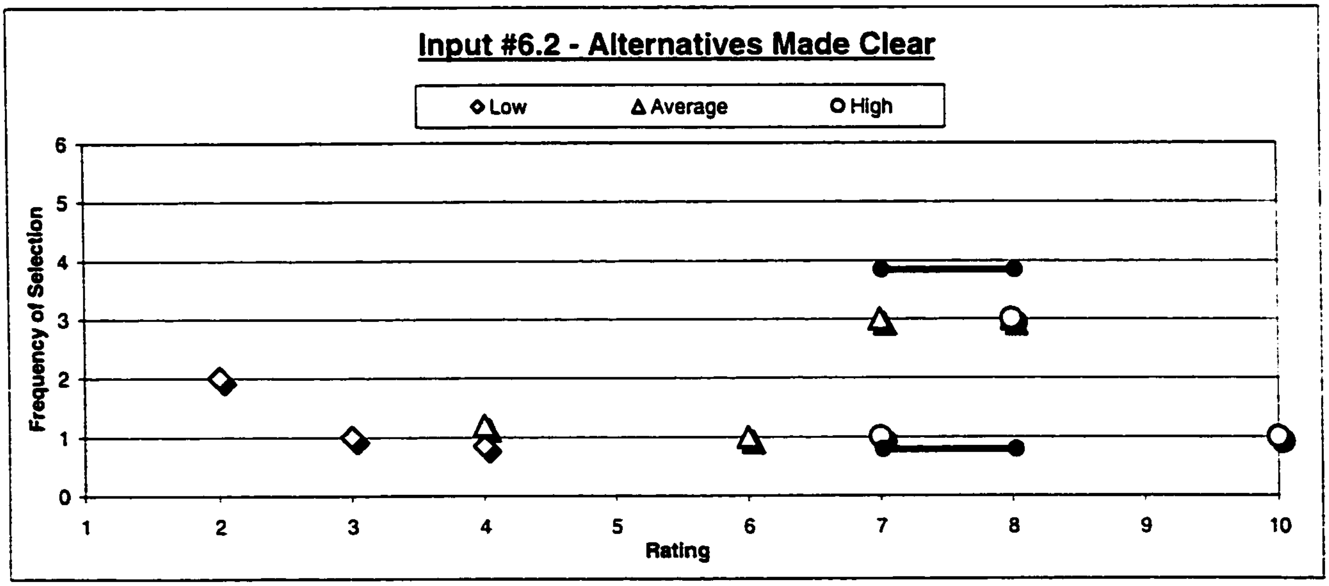

Input #5.2 - Design Team Hrs/wk, Variable 'Small'
Frequency of Selection
10
Hrs/wk
0
50
100
150
200
250
300
350
400
450
500
0
1
2

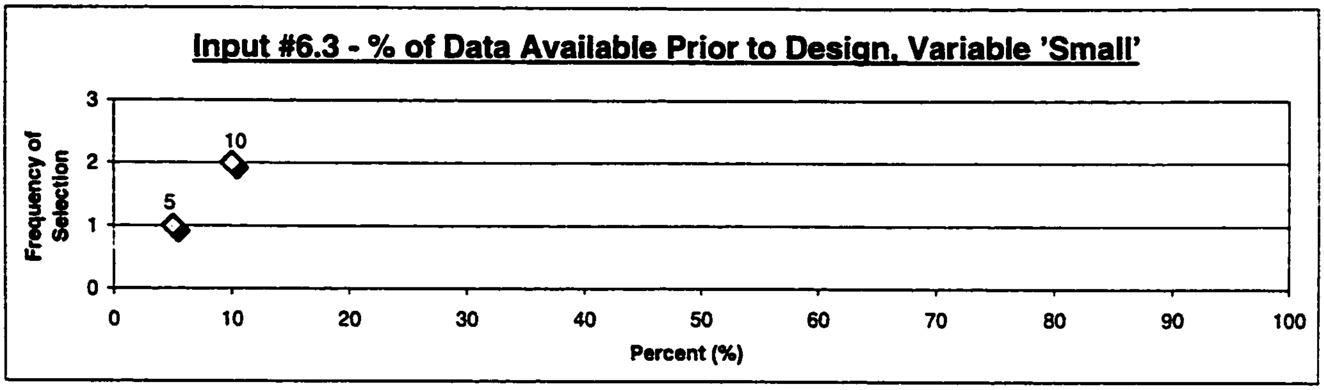

Input #5.2 - Design Team Hrs/wk, Variable 'Average'
Frequency of Selection
40
55
45
150
200
400
Hrs/wk
0
200
400
600
800
1000
1200
1400
1600
1800
0
1
2
3

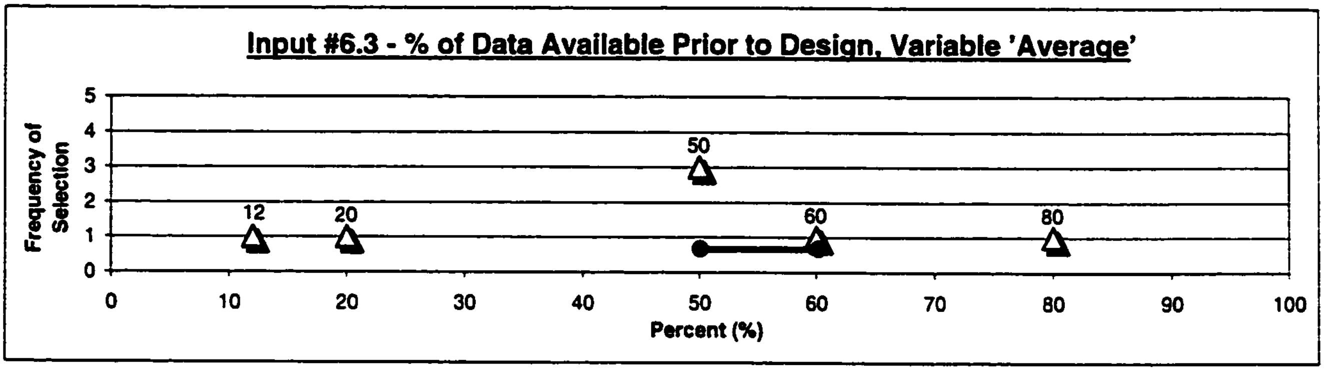

Input #5.2 - Design Team Hrs/wk, Variable 'Large'
Frequency of Selection
20
30
100
Hrs/wk
0
50
100
150
200
250
300
350
400
450
500
0
1
2

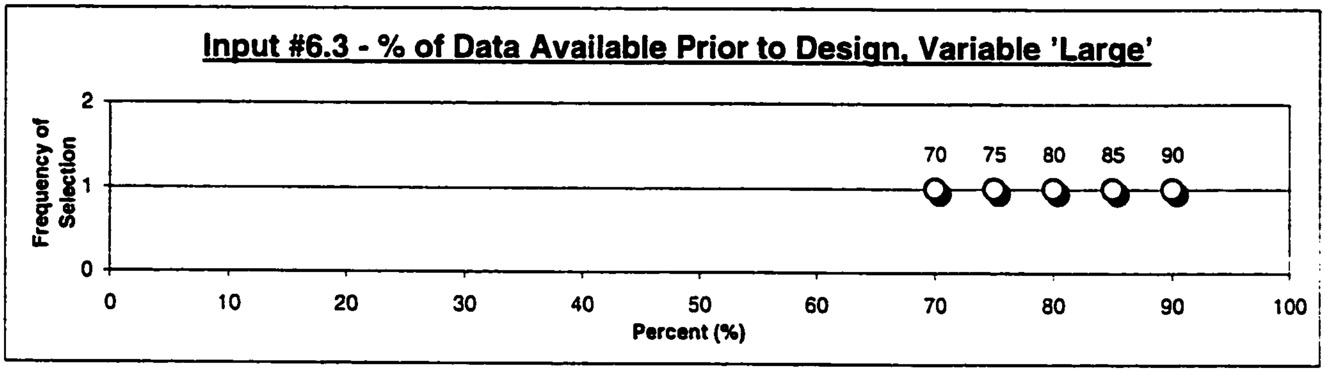

Input #6 - Level of Scope Definition
Low
Average
High
Frequency of Selection
Rating
0
1
2
3
4
5
1
2
3
4
5
6
7
8
9
10

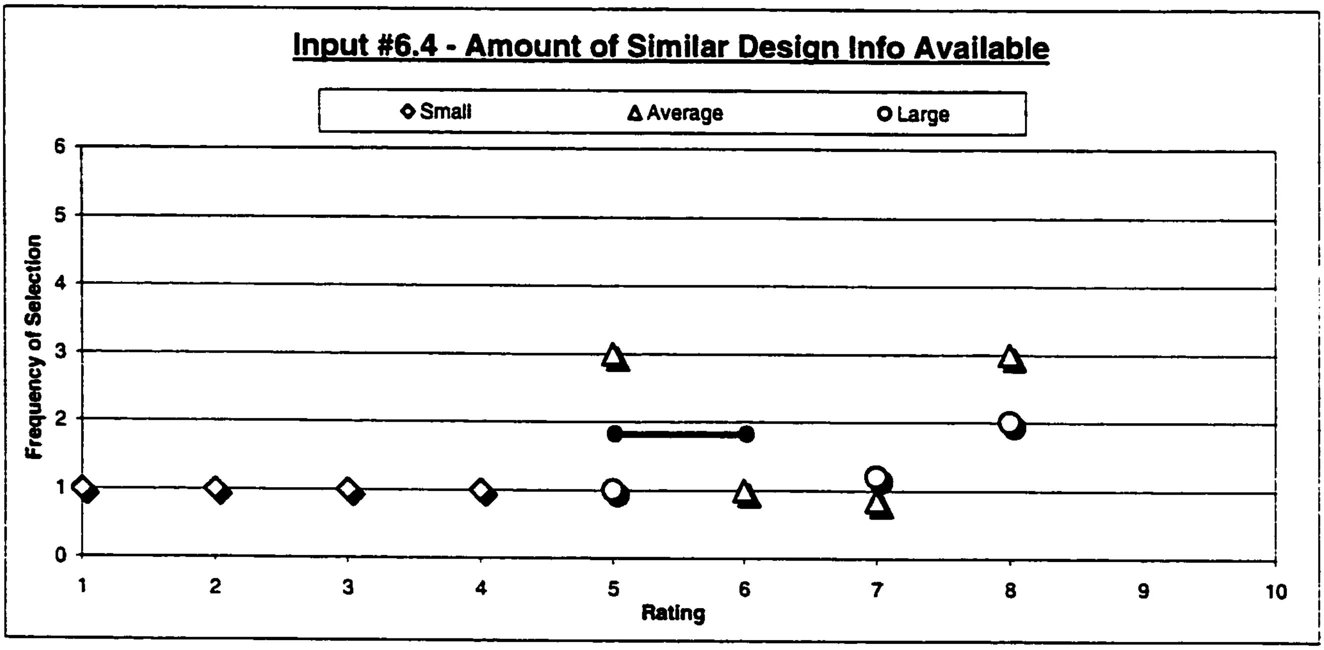

Input #6.1 - Extent the Definition of the Project Was Made Clear
Low
Average
High
Frequency of Selection
Rating
0
1
2
3
4
5
6
1
2
3
4
5
6
7
8
9
10

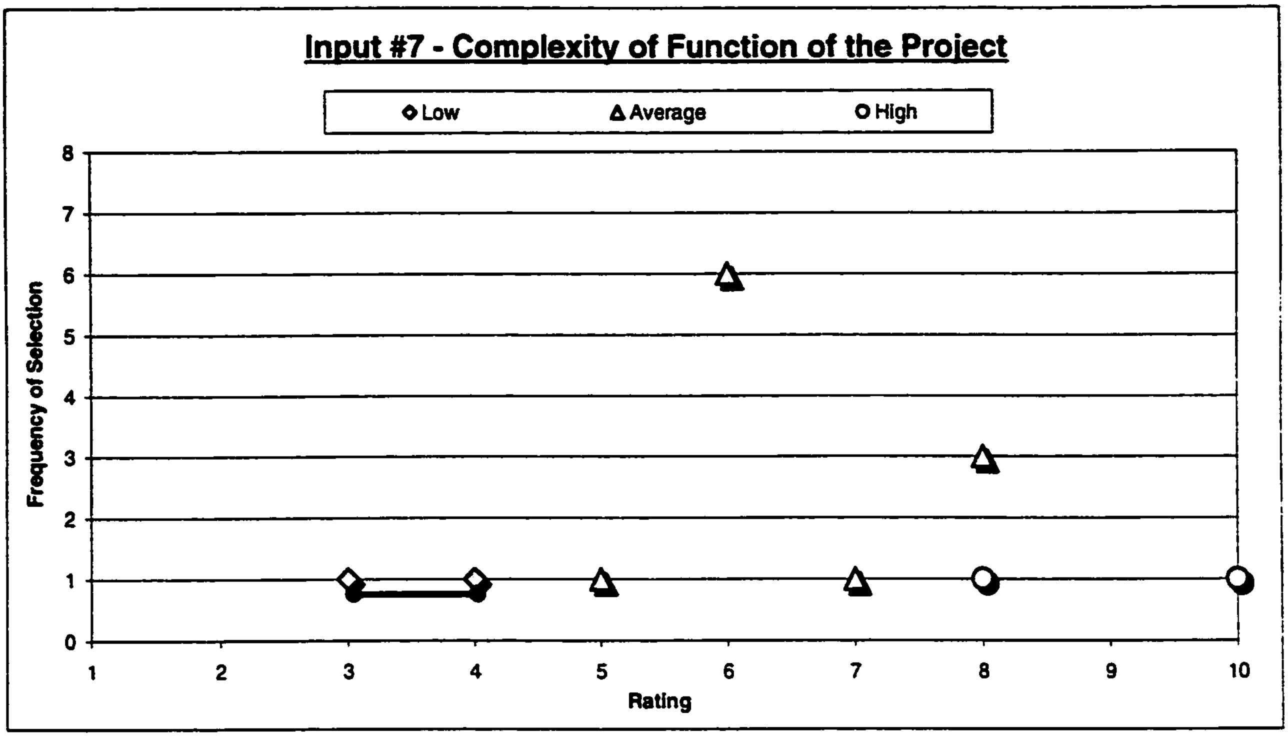

Input #6.2 - Alternatives Made Clear
Low
Average
High
Frequency of Selection
Rating
0
1
2
3
4
5
6
1
2
3
4
5
6
7
8
9
10

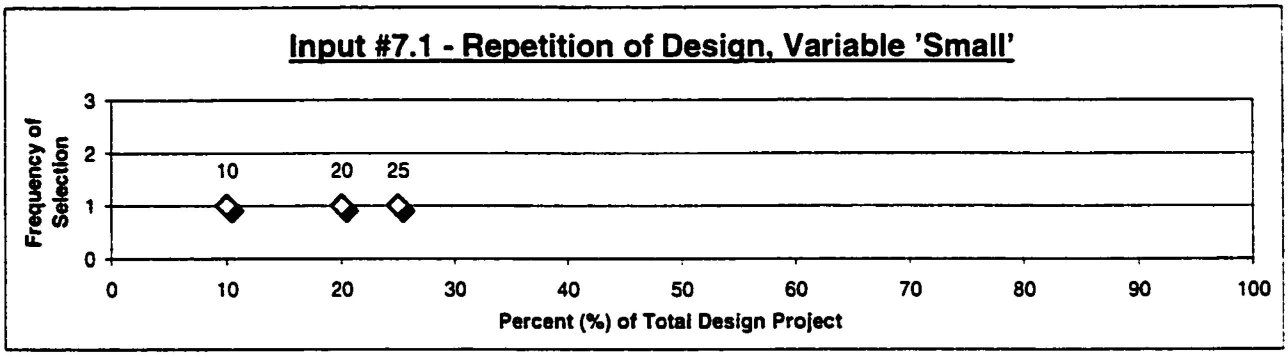

Input #6.3 - % of Data Available Prior to Design, Variable 'Small'
Frequency of Selection
Percent (%)
5
10

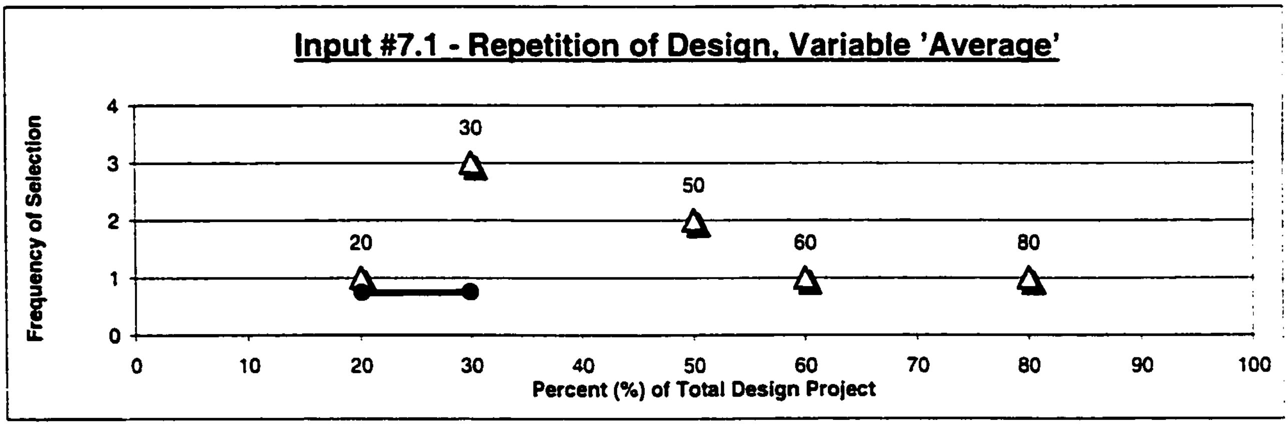

Input #6.3 - % of Data Available Prior to Design, Variable 'Average'
Frequency of Selection
Percent (%)
12
20
50
60
80

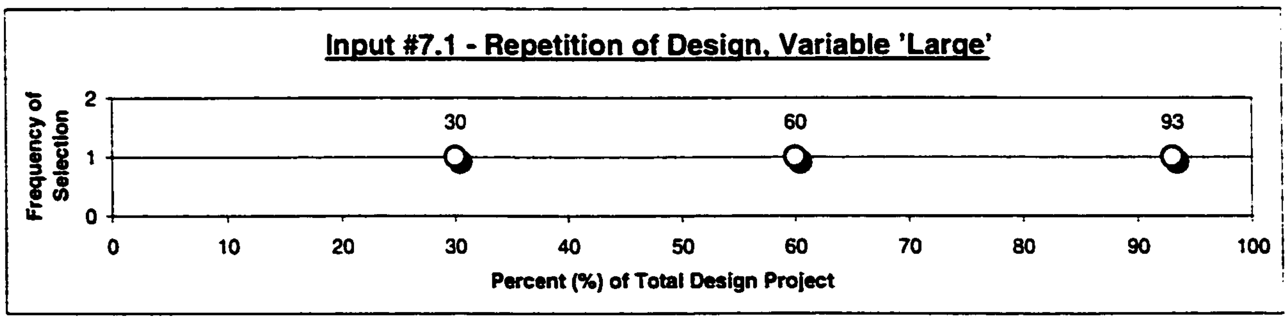

Input #6.3 - % of Data Available Prior to Design, Variable 'Large'
Frequency of Selection
Percent (%)
70
75
80
85
90

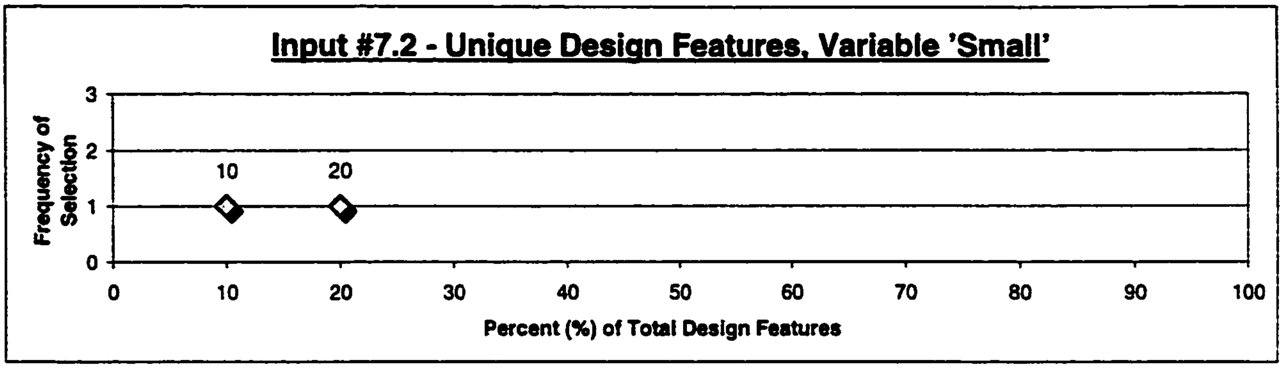

Input #6.4 - Amount of Similar Design Info Available
Small
Average
Large
Frequency of Selection
Rating

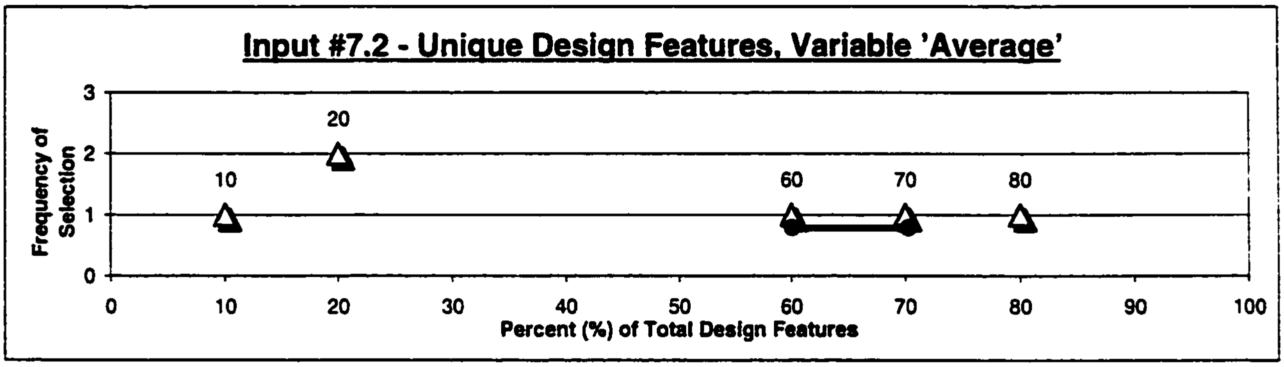

Input #7 - Complexity of Function of the Project
Low
Average
High
Frequency of Selection
Rating

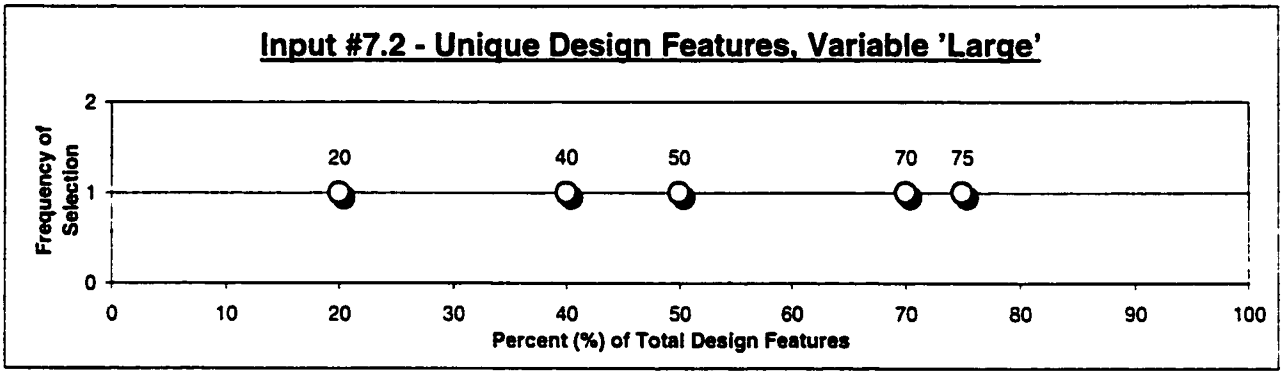

Input #7.1 - Repetition of Design, Variable 'Small'
Frequency of Selection
Percent (%) of Total Design Project
10
20
25

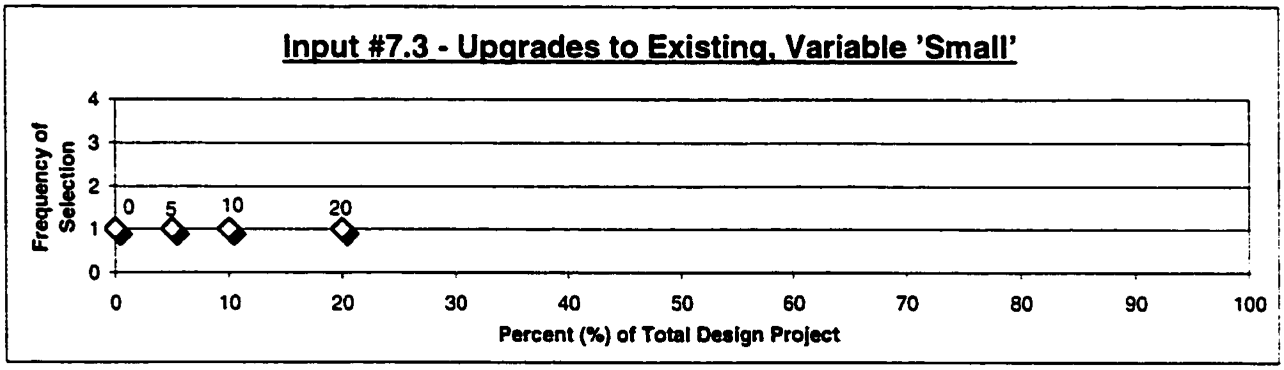

Input #7.1 - Repetition of Design, Variable 'Average'
Frequency of Selection
Percent (%) of Total Design Project
30
50
20
60
80

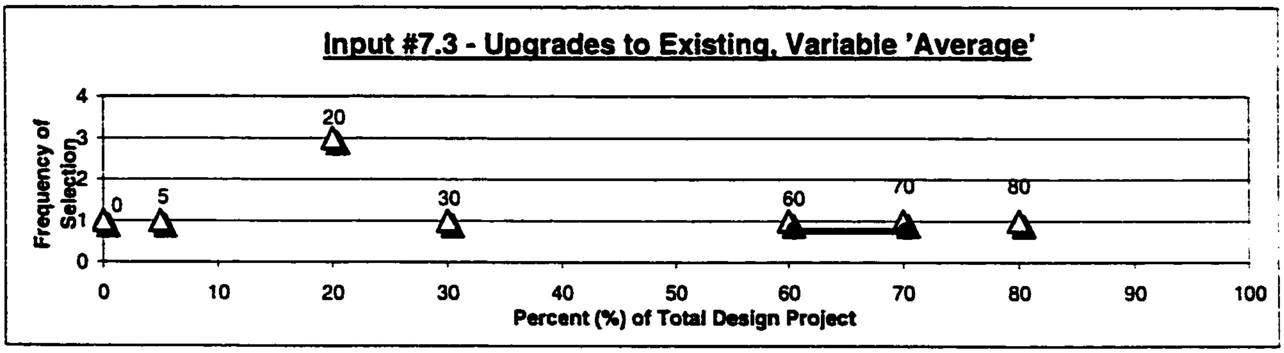

Input #7.1 - Repetition of Design, Variable 'Large'
Frequency of Selection
Percent (%) of Total Design Project
30
60
93

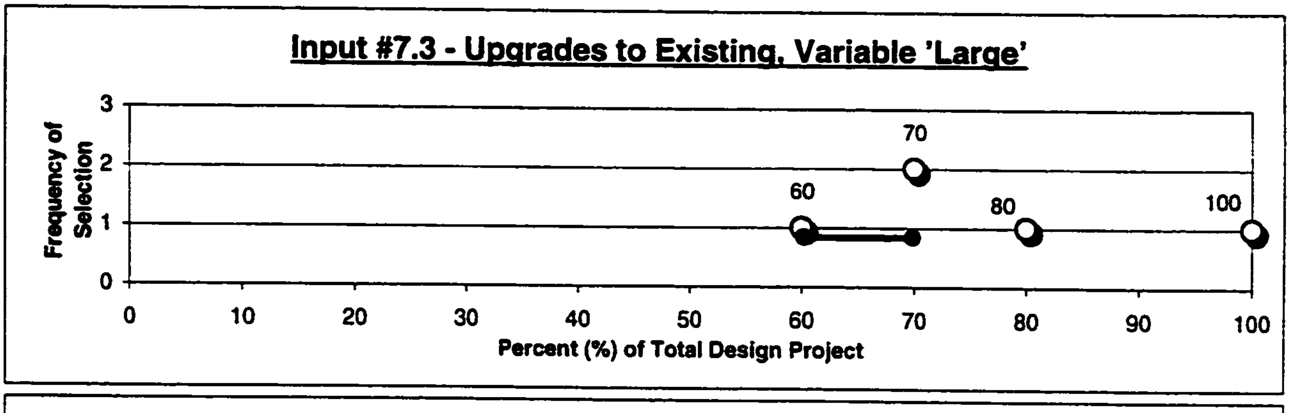

Input #7.2 - Unique Design Features, Variable 'Small'
Frequency of Selection
10
20
0 10 20 30 40 50 60 70 80 90 100
Percent (%) of Total Design Features

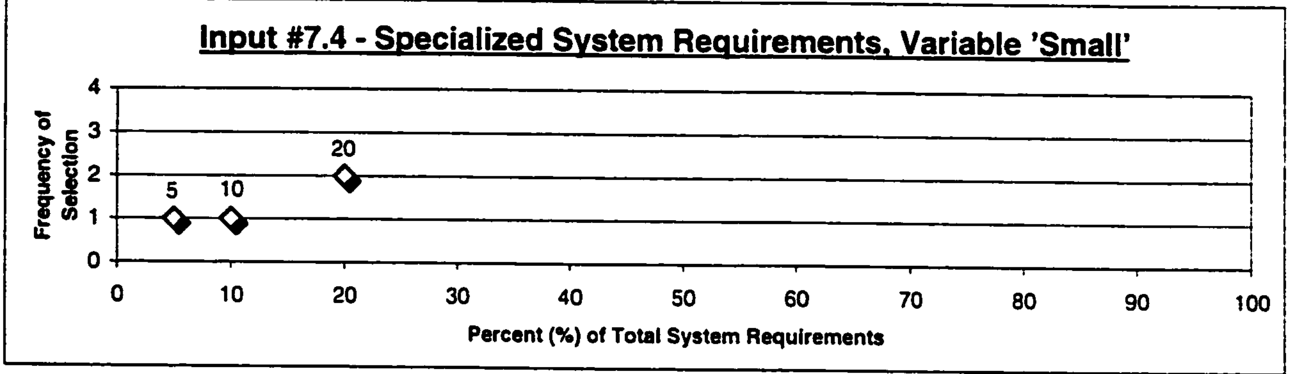

Input #7.2 - Unique Design Features, Variable 'Average'
Frequency of Selection
20
10
60
70
80
0 10 20 30 40 50 60 70 80 90 100
Percent (%) of Total Design Features

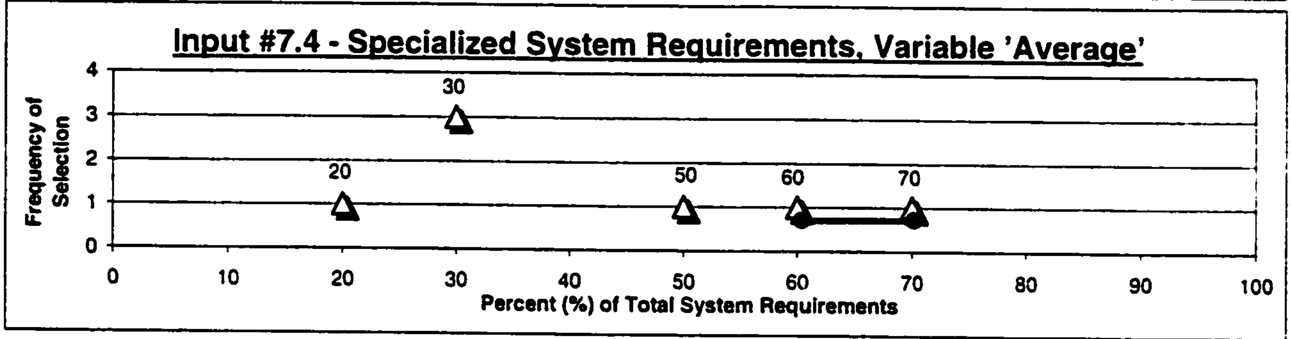

Input #7.2 - Unique Design Features, Variable 'Large'
Frequency of Selection
20
40
50
70
75
0 10 20 30 40 50 60 70 80 90 100
Percent (%) of Total Design Features

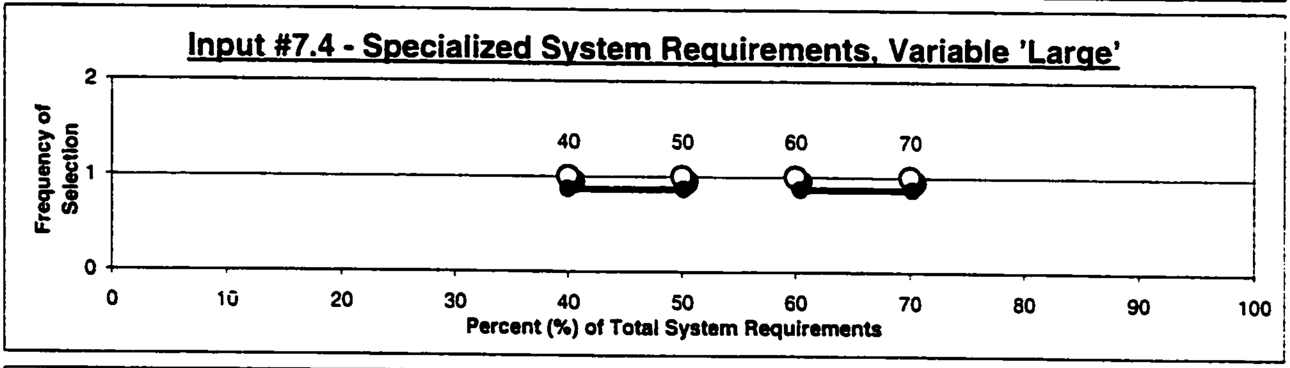

Input #7.3 - Upgrades to Existing, Variable 'Small'
Frequency of Selection
0
5
10
20
0 10 20 30 40 50 60 70 80 90 100
Percent (%) of Total Design Project

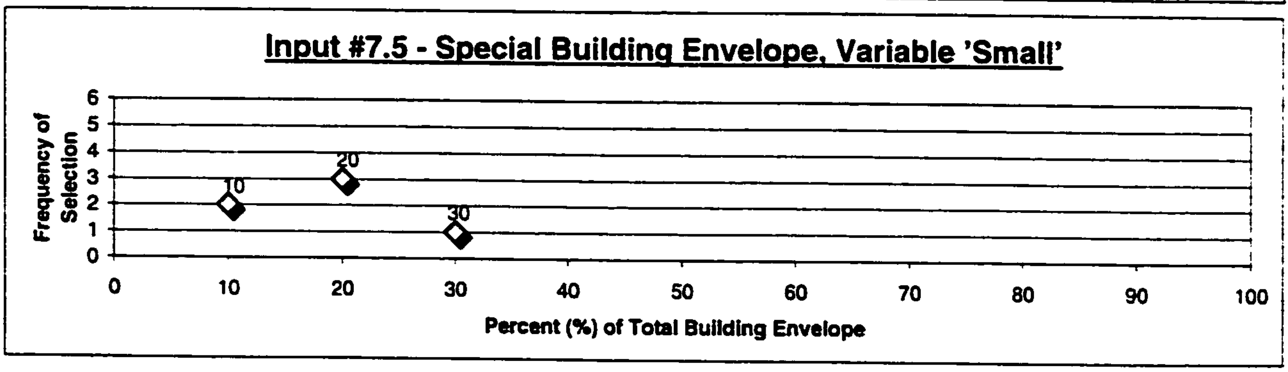

Input #7.3 - Upgrades to Existing, Variable 'Average'
Frequency of Selection
20
0
5
30
60
70
80
0 10 20 30 40 50 60 70 80 90 100
Percent (%) of Total Design Project

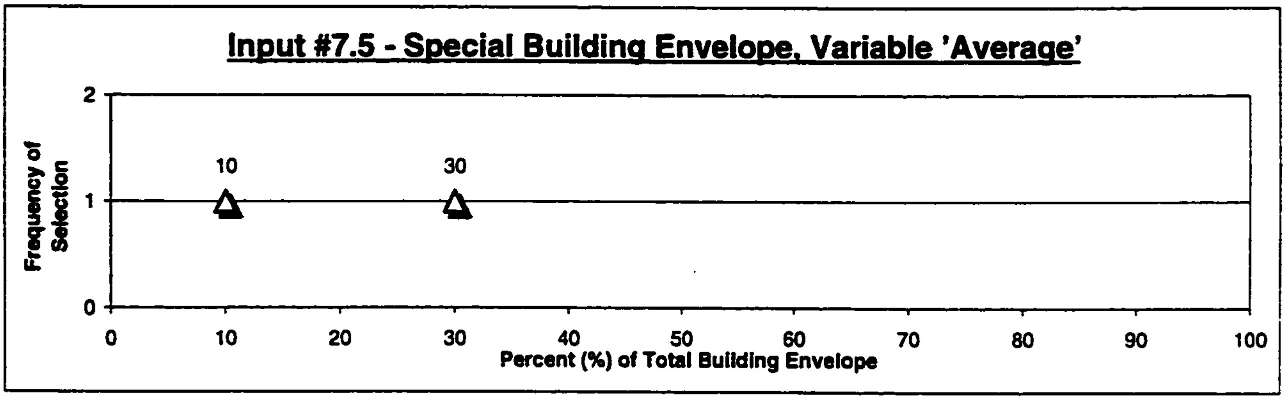

Input #7.3 - Upgrades to Existing, Variable 'Large'
Frequency of Selection
Percent (%) of Total Design Project
60
70
80
100

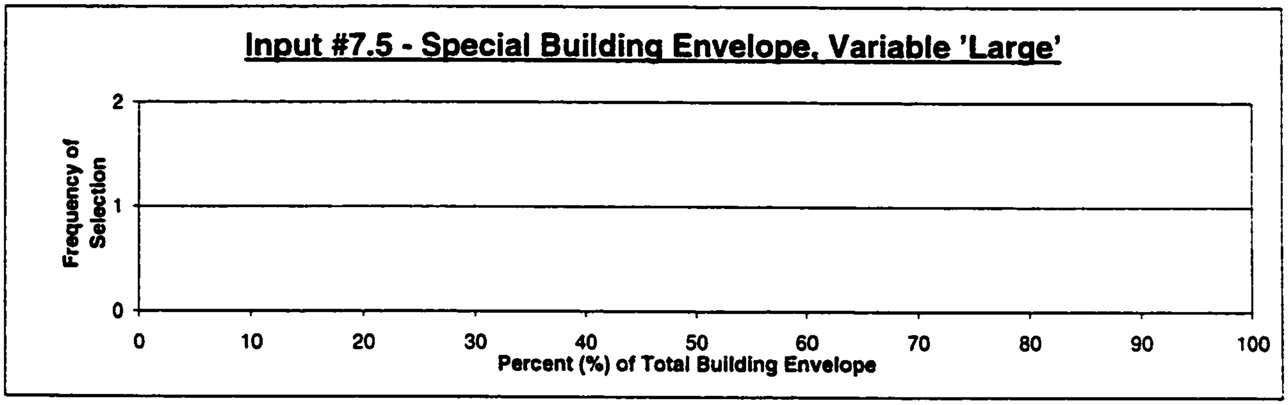

Input #7.4 - Specialized System Requirements, Variable 'Small'
Frequency of Selection
Percent (%) of Total System Requirements
5
10
20

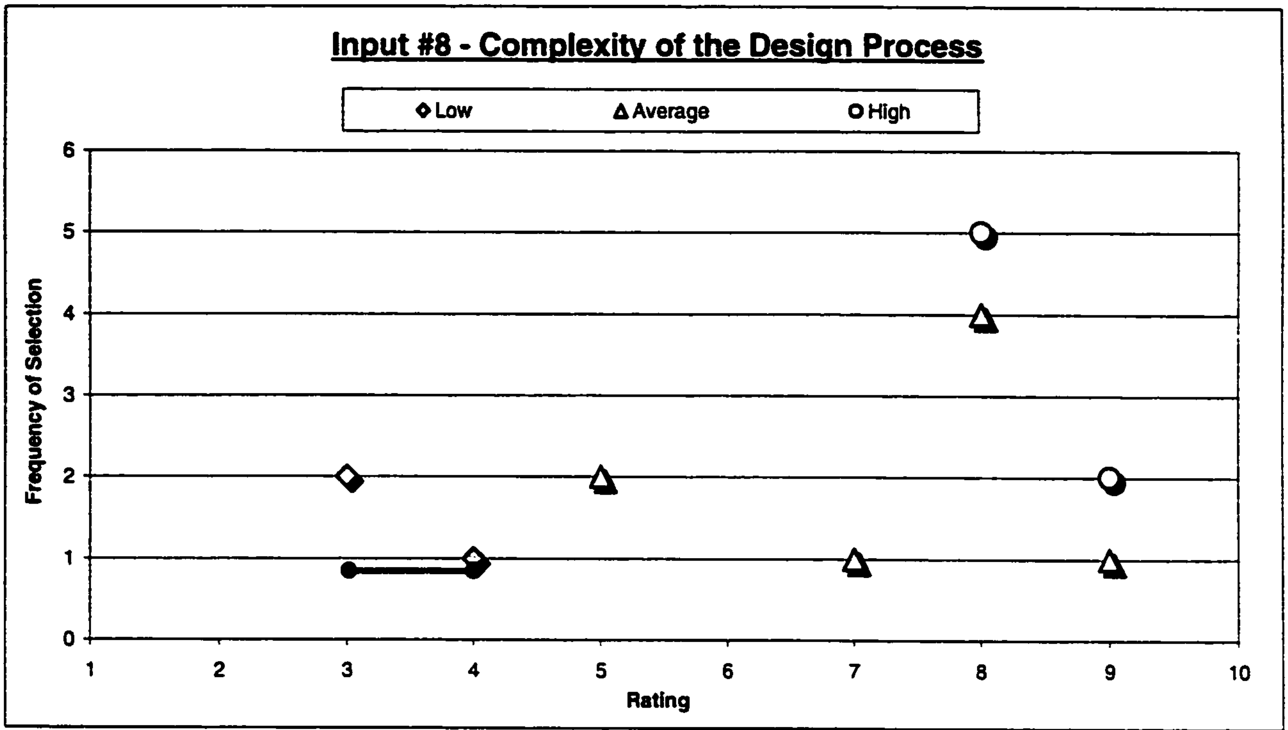

Input #7.4 - Specialized System Requirements, Variable 'Average'
Frequency of Selection
Percent (%) of Total System Requirements
30
20
50
60
70


Input #7.4 - Specialized System Requirements, Variable 'Large'
Frequency of Selection
Percent (%) of Total System Requirements
40
50
60
70

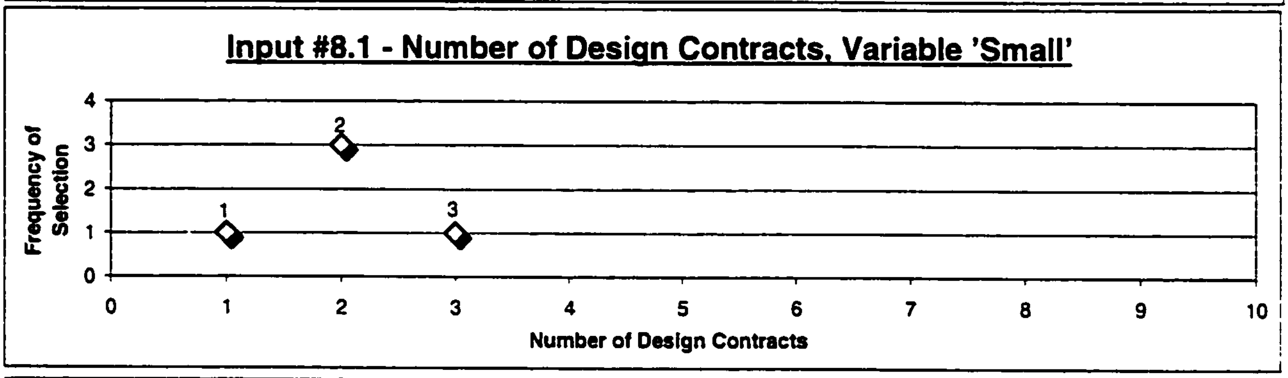

Input #7.5 - Special Building Envelope, Variable 'Small'
Frequency of Selection
Percent (%) of Total Building Envelope
10
20
30


Input #7.5 - Special Building Envelope, Variable 'Average'
Frequency of Selection
Percent (%) of Total Building Envelope
10
30
0
1
2
0
10
20
30
40
50
60
70
80
90
100

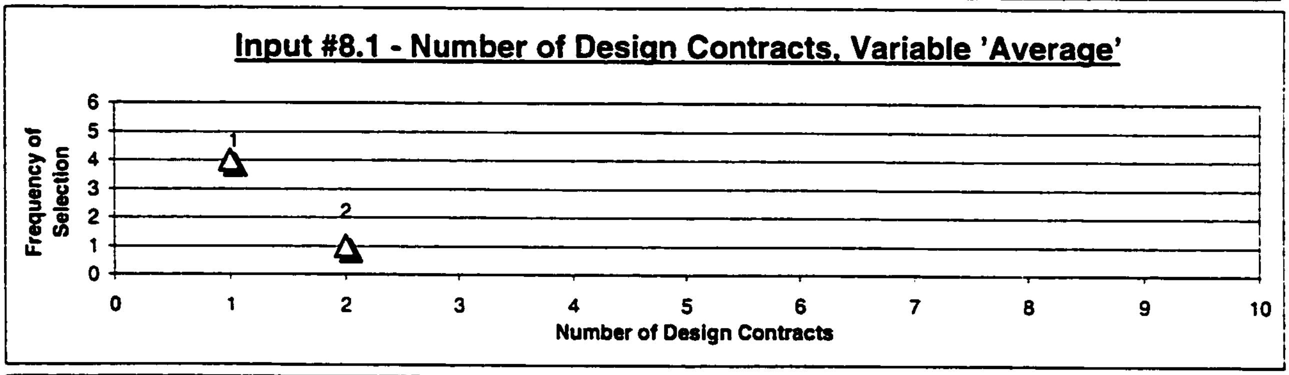

Input #7.5 - Special Building Envelope, Variable 'Large'
Frequency of Selection
Percent (%) of Total Building Envelope
0
1
2
0
10
20
30
40
50
60
70
80
90
100



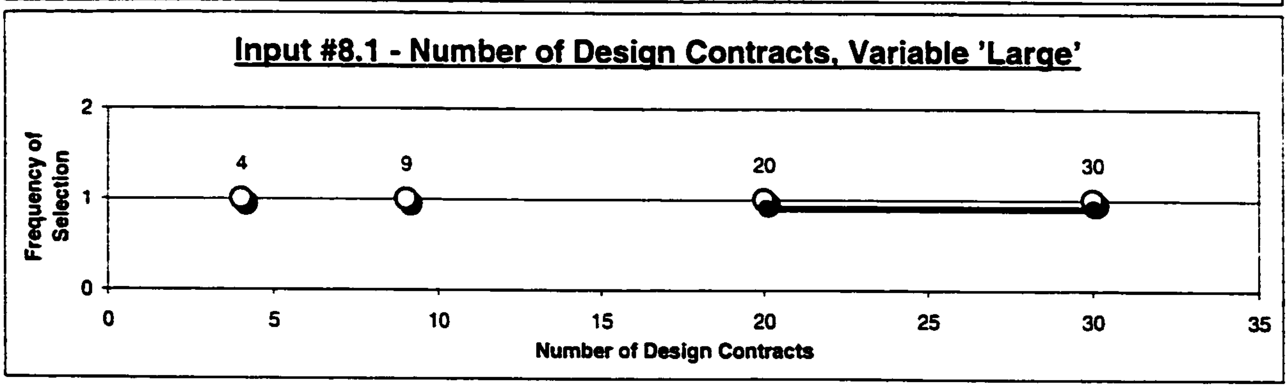
Input #8 - Complexity of the Design Process
Low
Average
High
Frequency of Selection
Rating
0
1
2
3
4
5
6
1
2
3
4
5
6
7
8
9
10



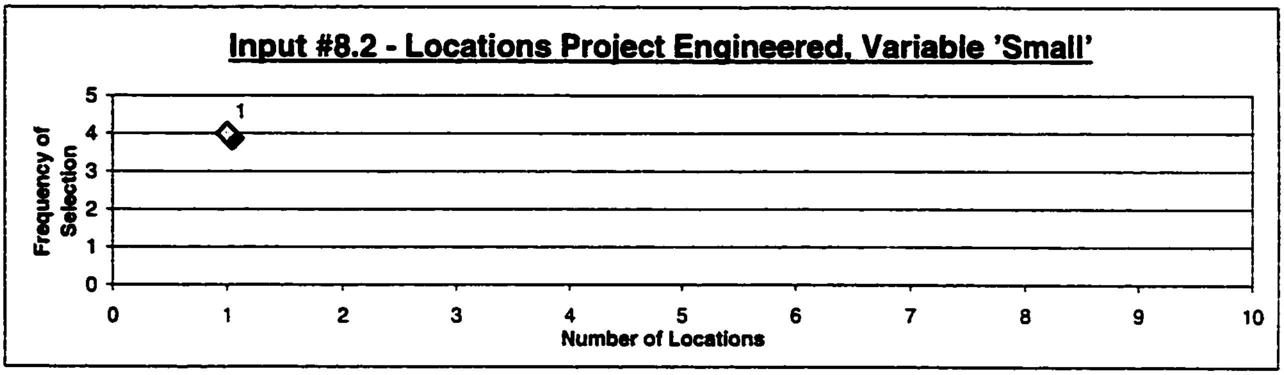
Input #8.1 - Number of Design Contracts, Variable 'Small'
Frequency of Selection
Number of Design Contracts
1
2
3
0
1
2
3
4
0
1
2
3
4
5
6
7
8
9
10

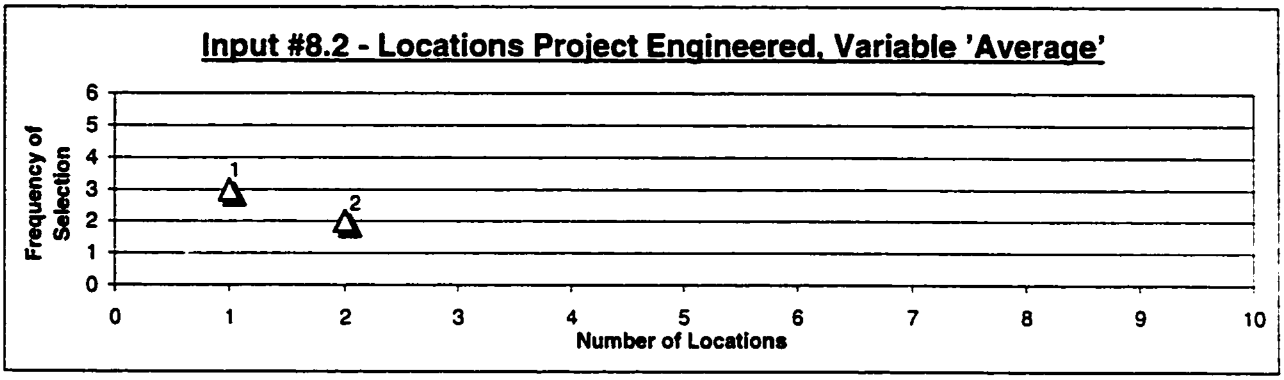


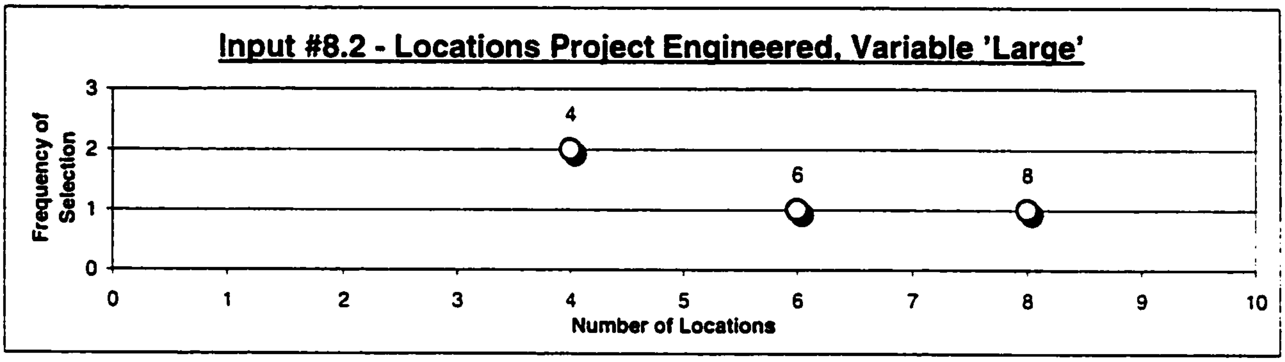
Input #8.1 - Number of Design Contracts, Variable 'Average'
Frequency of Selection
Number of Design Contracts
1
2
0
1
2
3
4
5
6
0
1
2
3
4
5
6
7
8
9
10

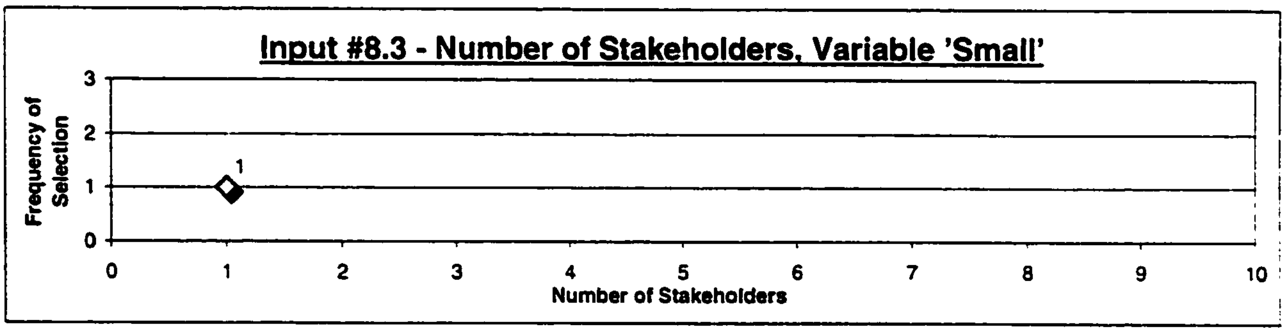


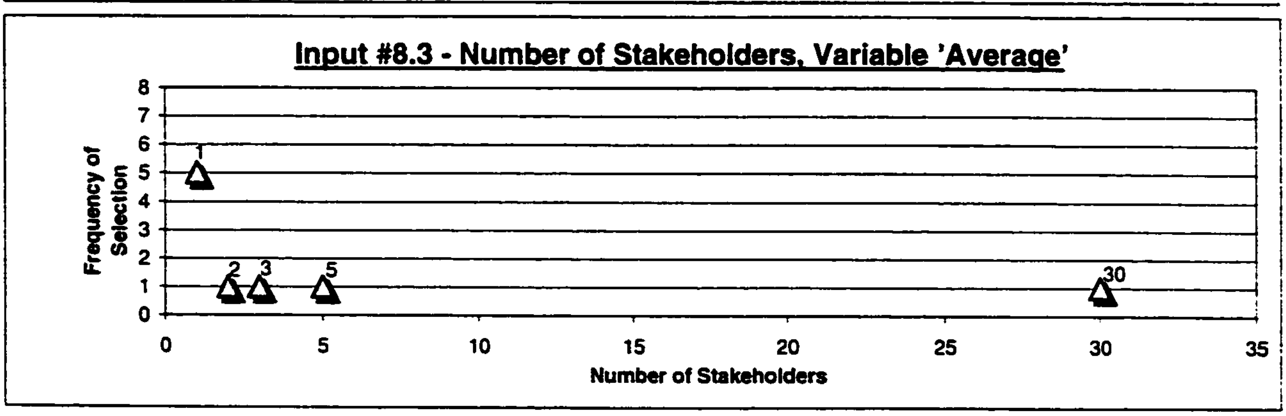
Input #8.1 - Number of Design Contracts, Variable 'Large'
Frequency of Selection
Number of Design Contracts
4
9
20
30
0
1
2
0
5
10
15
20
25
30
35

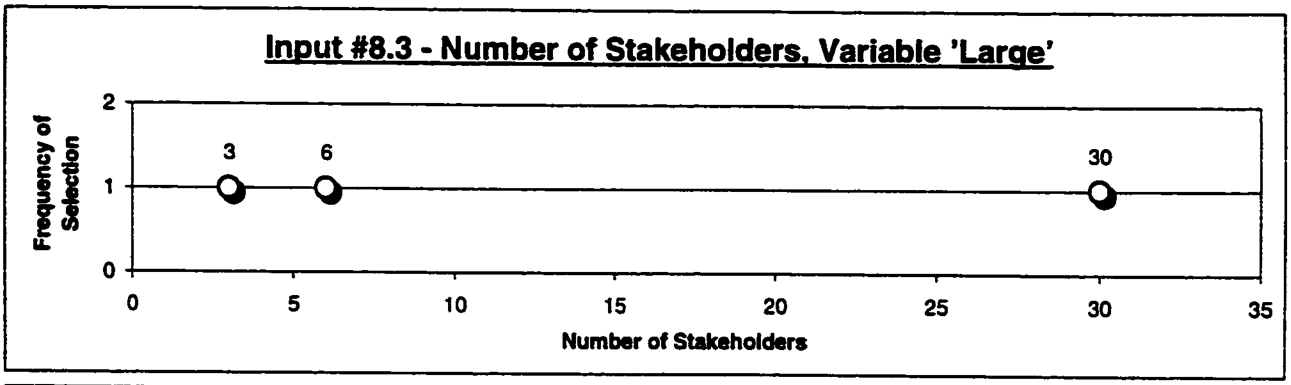

Input #8.2 - Locations Project Engineered, Variable 'Small'
Frequency of Selection
Number of Locations
1

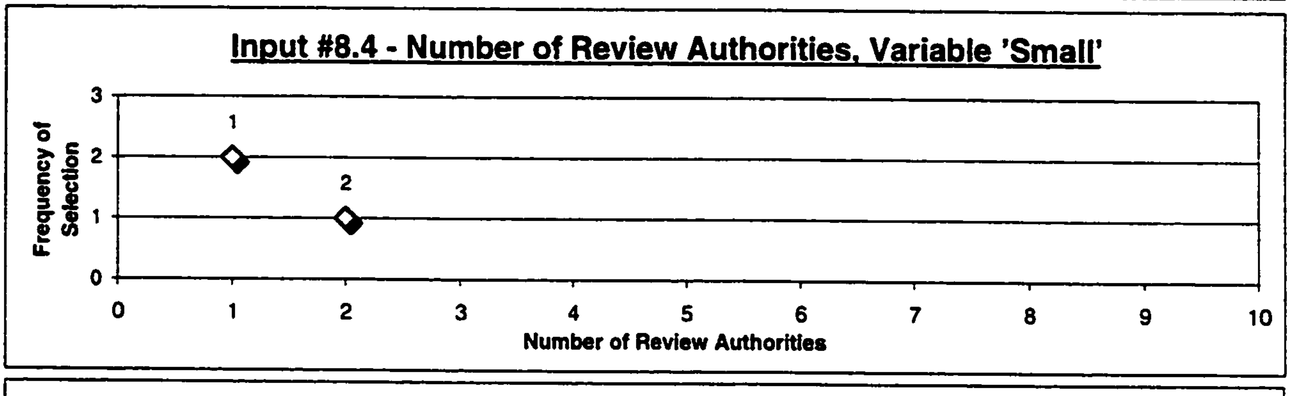

Input #8.2 - Locations Project Engineered, Variable 'Average'
Frequency of Selection
Number of Locations
1
2

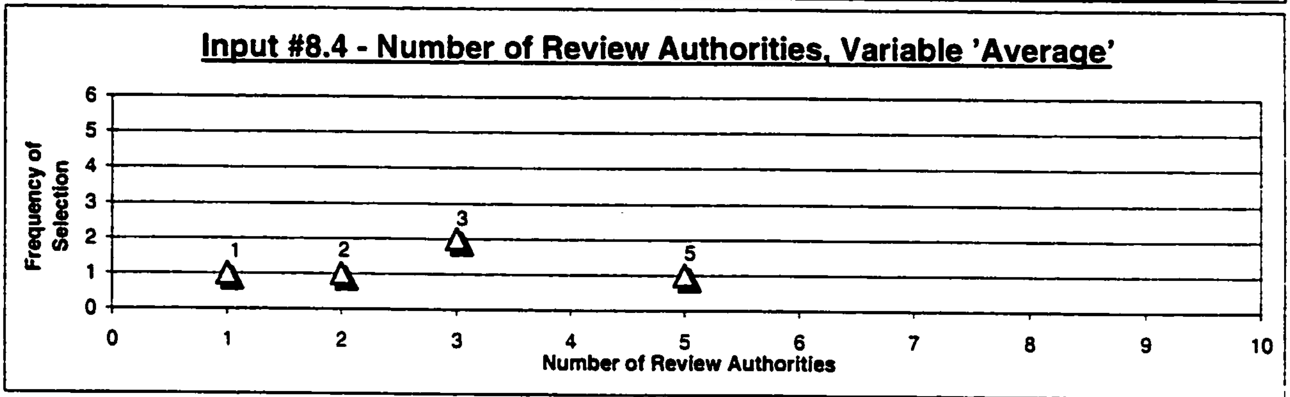

Input #8.2 - Locations Project Engineered, Variable 'Large'
Frequency of Selection
Number of Locations
4
6
8

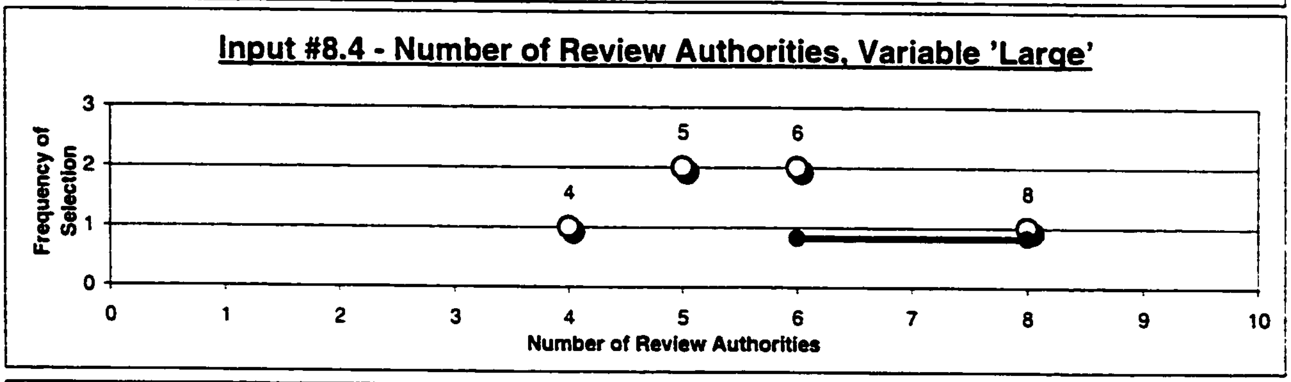

Input #8.3 - Number of Stakeholders, Variable 'Small'
Frequency of Selection
Number of Stakeholders
1

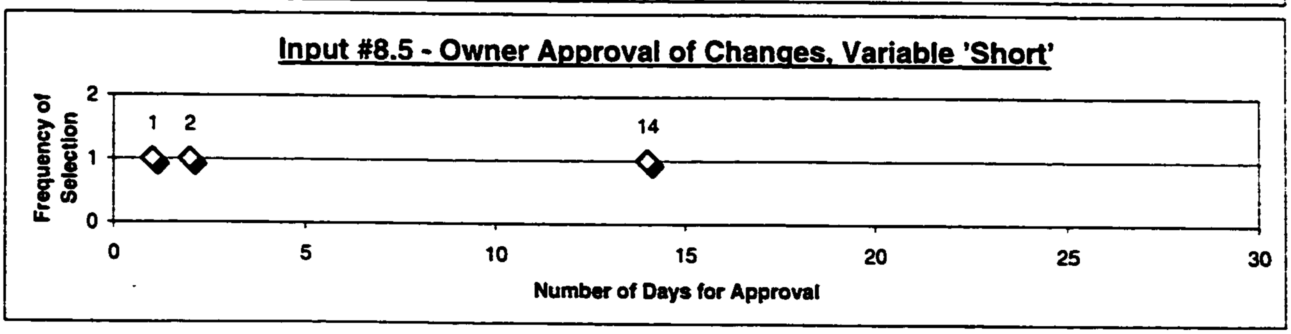

Input #8.3 - Number of Stakeholders, Variable 'Average'
Frequency of Selection
Number of Stakeholders
1
2
3
5
30

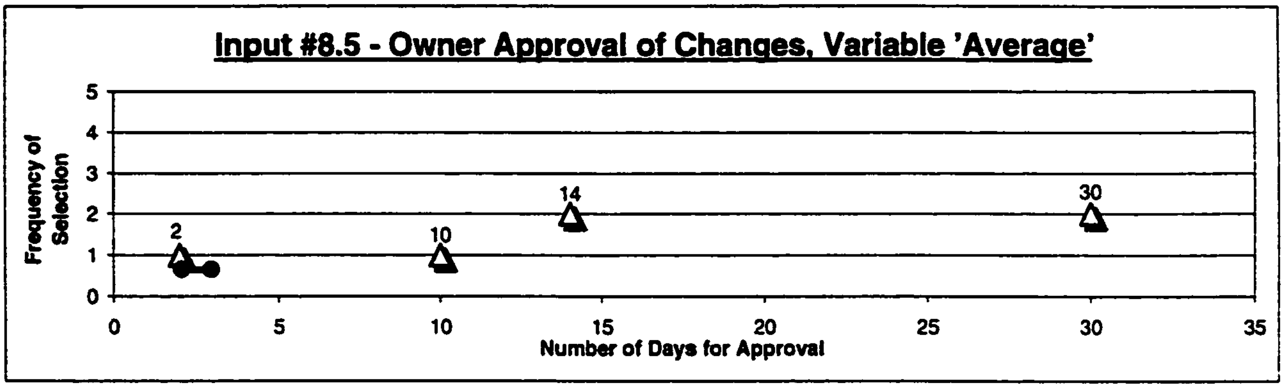

Input #8.3 - Number of Stakeholders, Variable 'Large'
Frequency of Selection
Number of Stakeholders
3
6
30
0
5
10
15
20
25
30
35
0
1
2

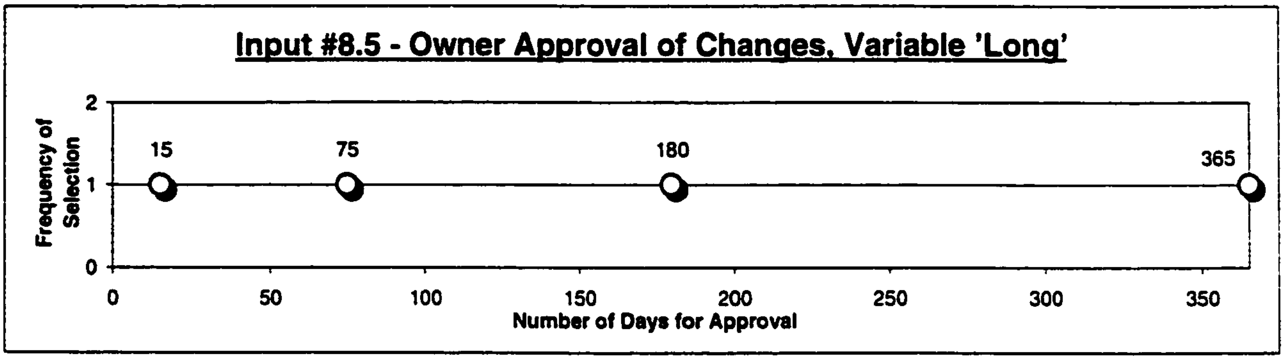

Input #8.4 - Number of Review Authorities, Variable 'Small'
Frequency of Selection
Number of Review Authorities
1
2
0
1
2
3
4
5
6
7
8
9
10

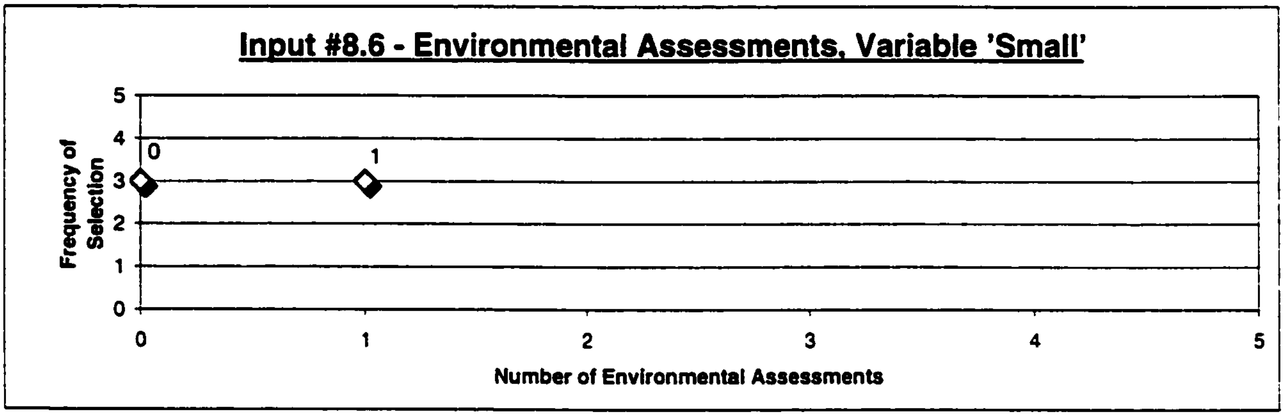

Input #8.4 - Number of Review Authorities, Variable 'Average'
Frequency of Selection
Number of Review Authorities
1
2
3
5
0
1
2
3
4
5
6
7
8
9
10

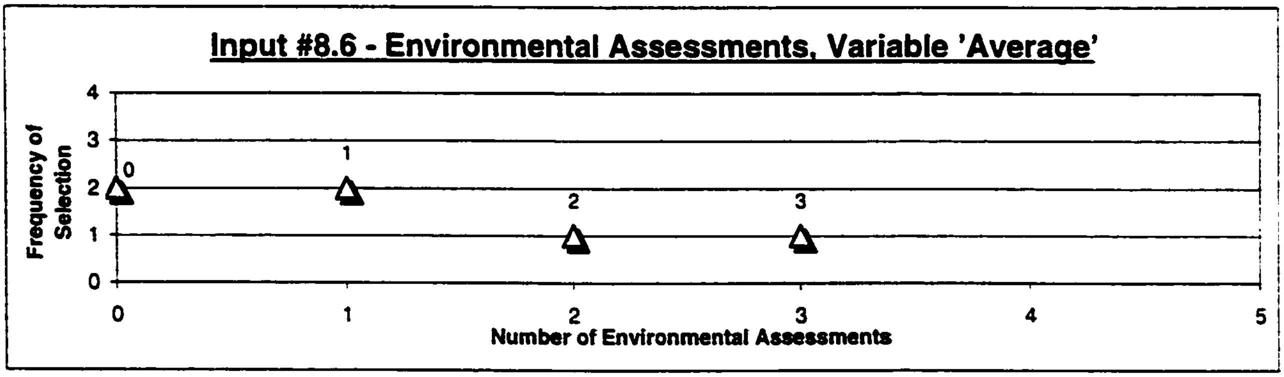

Input #8.4 - Number of Review Authorities, Variable 'Large'
Frequency of Selection
Number of Review Authorities
4
5
6
8
0
1
2
3
4
5
6
7
8
9
10

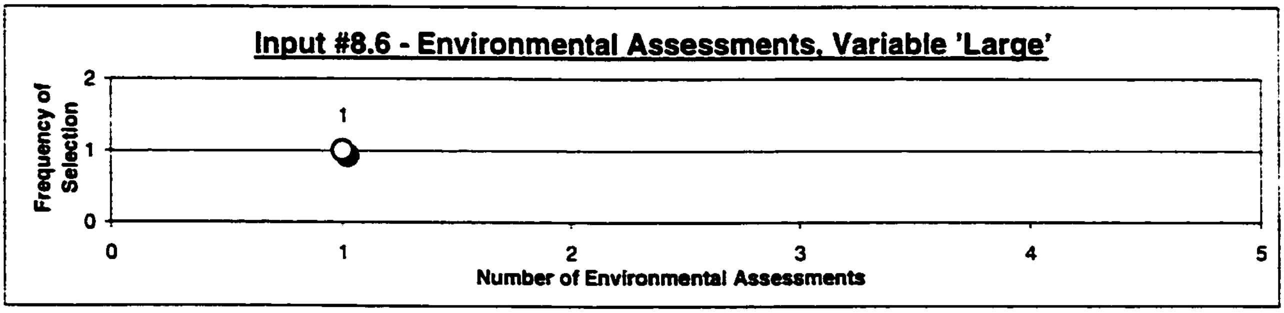

Input #8.5 - Owner Approval of Changes, Variable 'Short'
Frequency of Selection
Number of Days for Approval
1
2
14
0
5
10
15
20
25
30

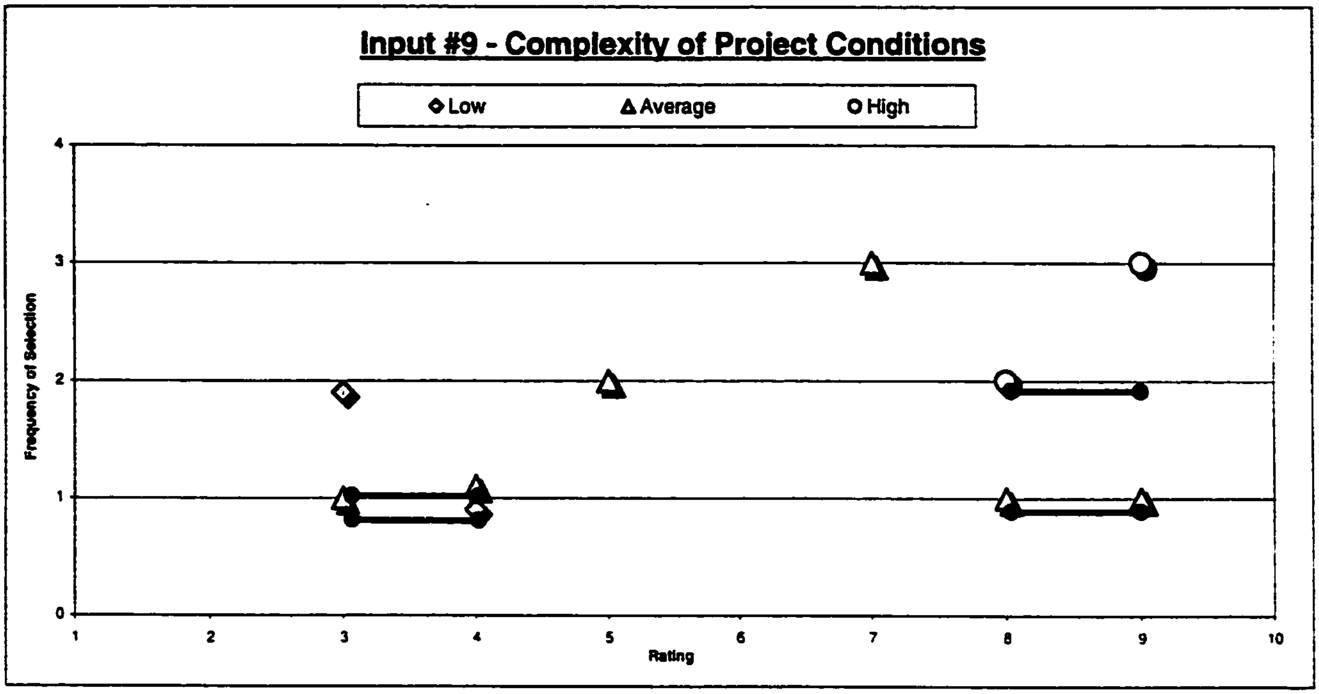

Input #8.5 - Owner Approval of Changes, Variable 'Average'
Frequency of Selection
Number of Days for Approval
2
10
14
30


Input #8.5 - Owner Approval of Changes, Variable 'Long'
Frequency of Selection
Number of Days for Approval
15
75
180
365

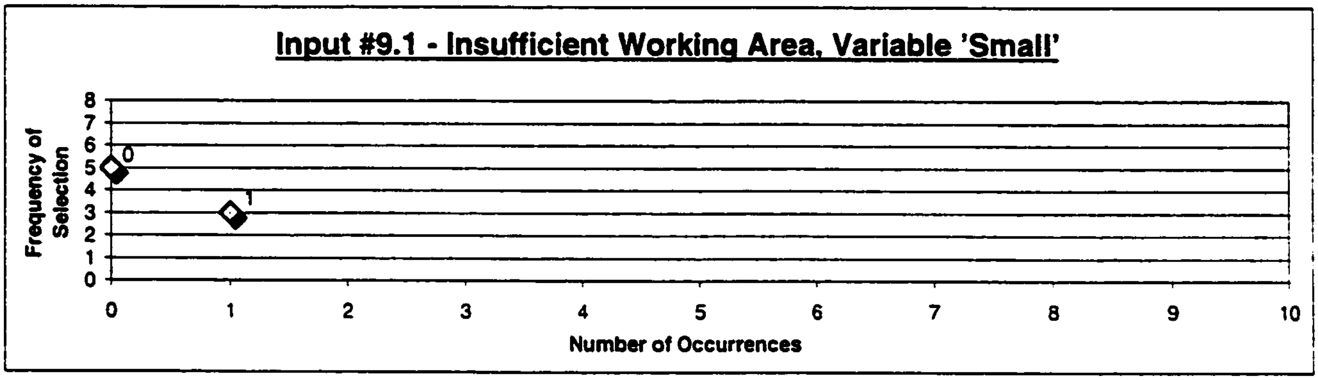

Input #8.6 - Environmental Assessments, Variable 'Small'
Frequency of Selection
Number of Environmental Assessments
0
1


Input #8.6 - Environmental Assessments, Variable 'Average'
Frequency of Selection
Number of Environmental Assessments
0
1
2
3

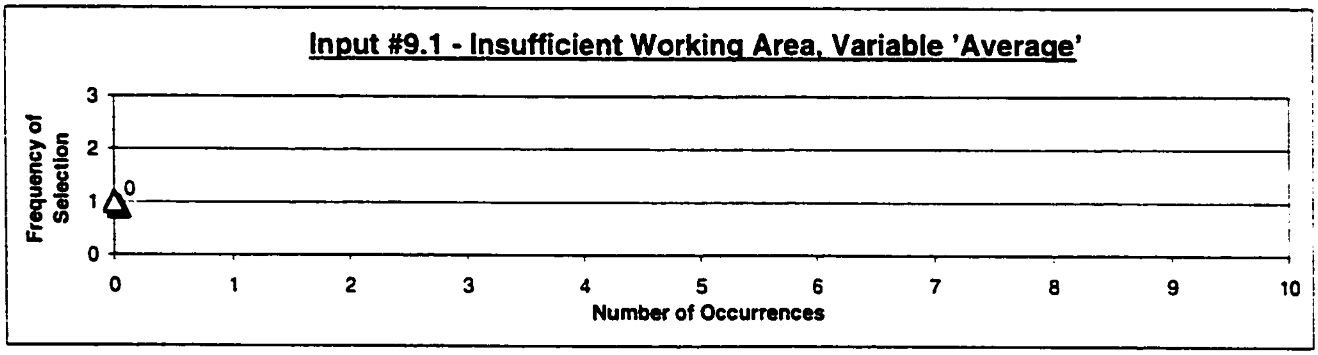

Input #8.6 - Environmental Assessments, Variable 'Large'
Frequency of Selection
Number of Environmental Assessments
1



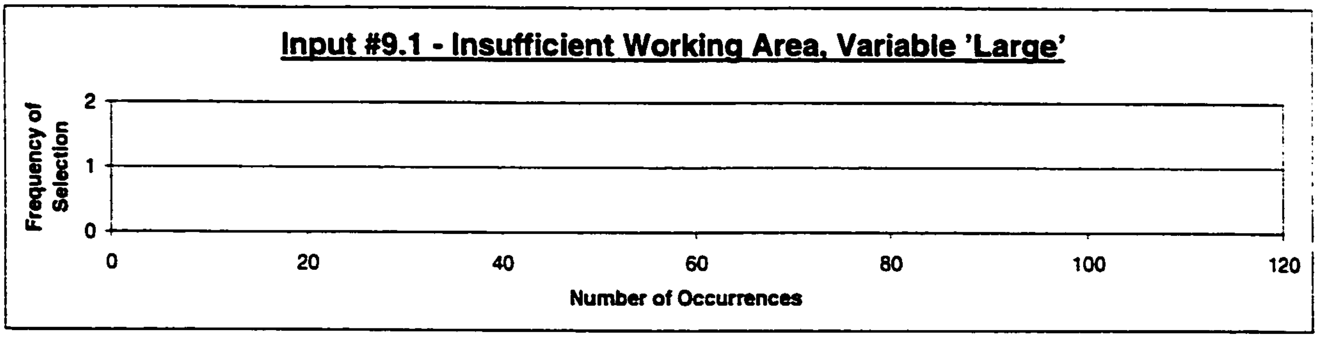
Input #9 - Complexity of Project Conditions
Low
Average
High
Frequency of Selection
Rating



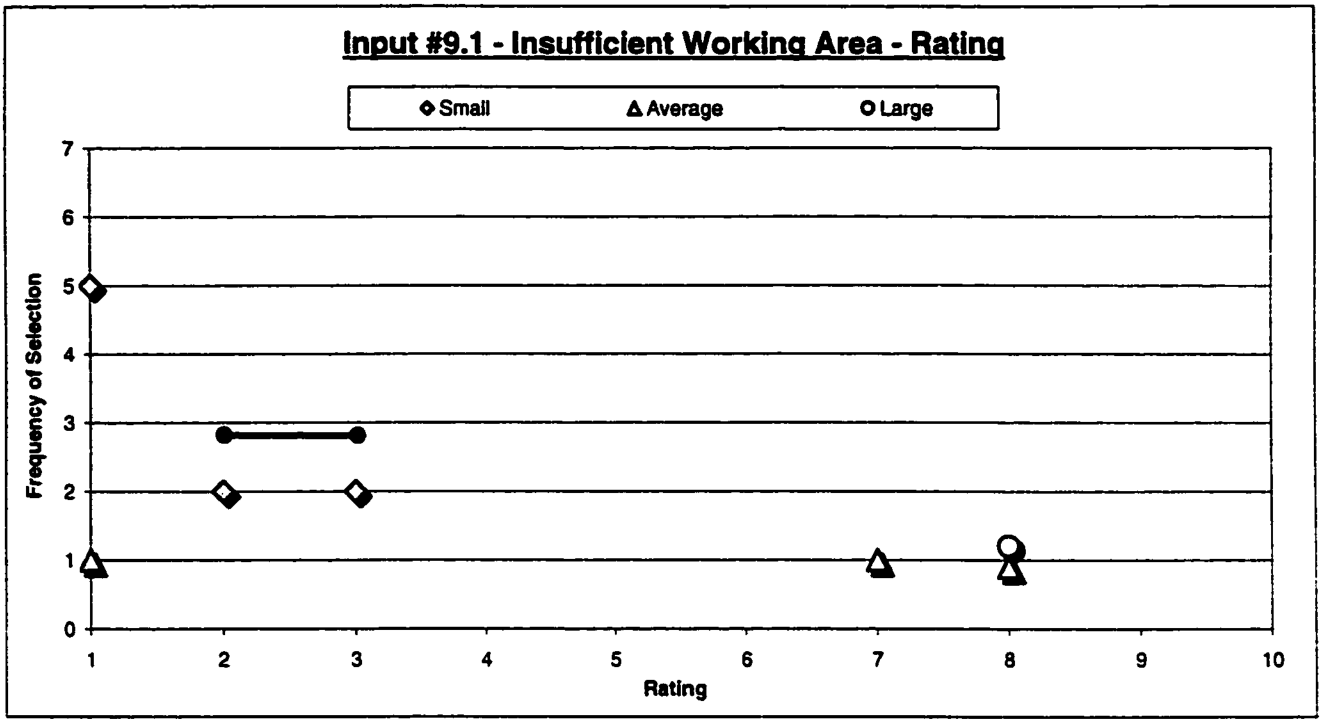
Input #9.1 - Insufficient Working Area, Variable 'Small'
Frequency of Selection
Number of Occurrences

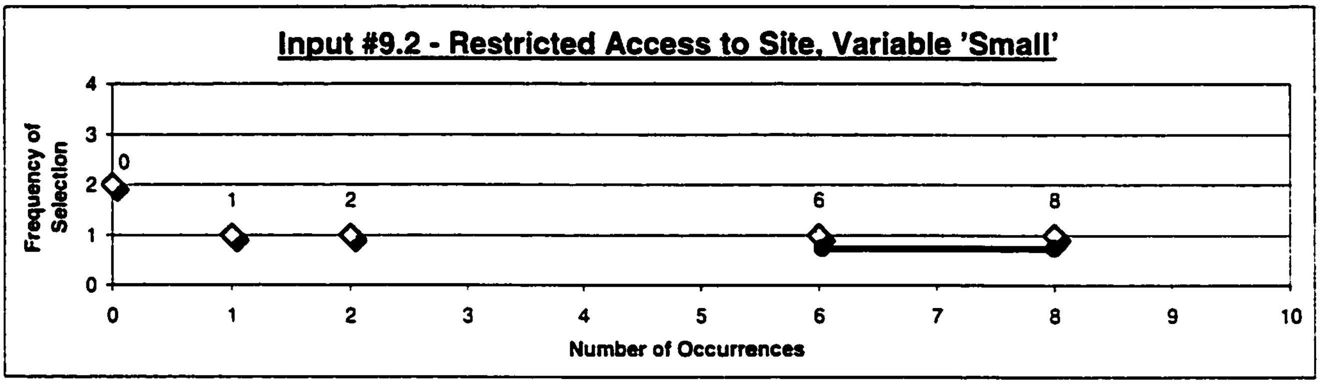


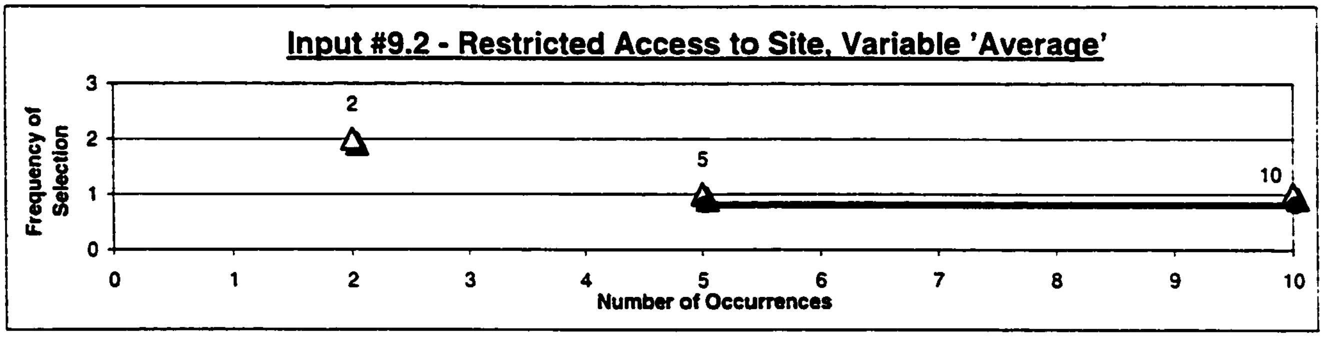
Input #9.1 - Insufficient Working Area, Variable 'Average'
Frequency of Selection
Number of Occurrences

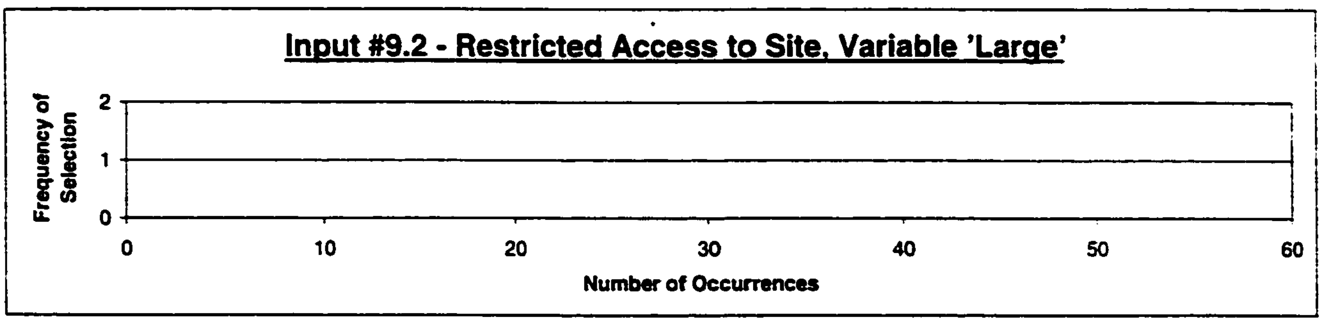


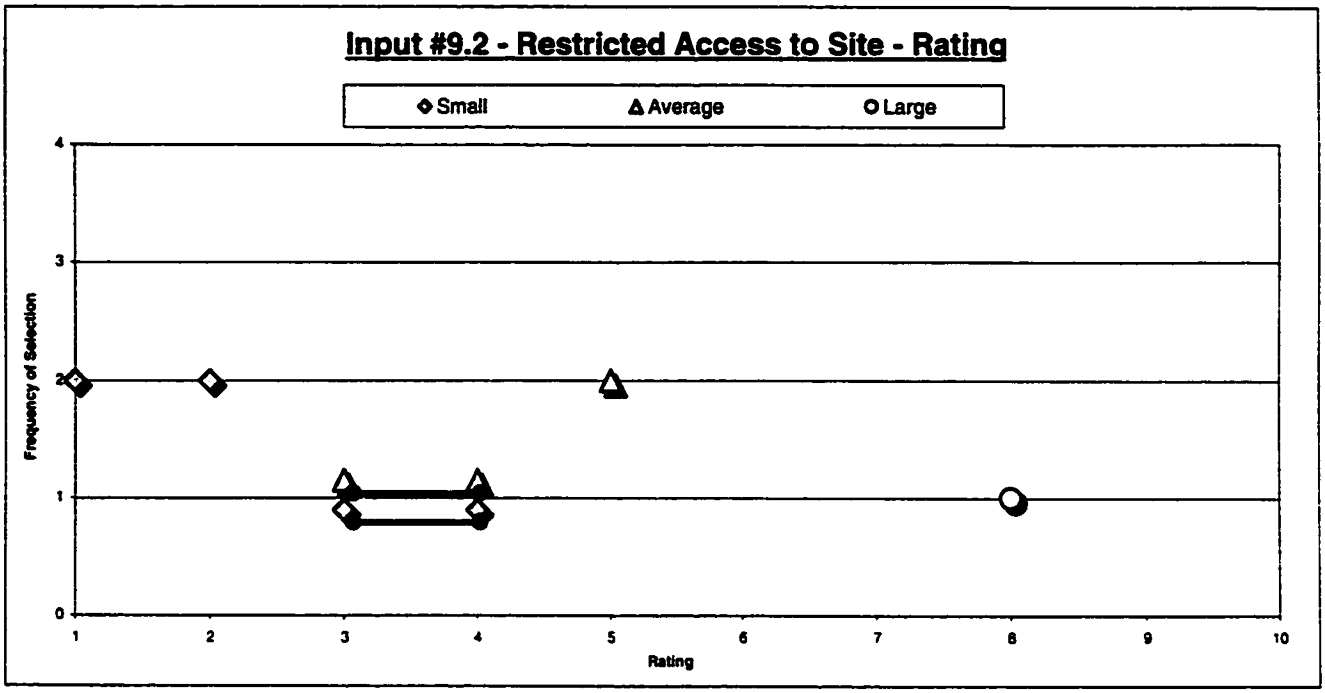
Input #9.1 - Insufficient Working Area, Variable 'Large'
Frequency of Selection
Number of Occurrences


Input #9.1 - Insufficient Working Area - Rating
Small
Average
Large
Frequency of Selection
Rating
0
1
2
3
4
5
6
7
1
2
3
4
5
6
7
8
9
10

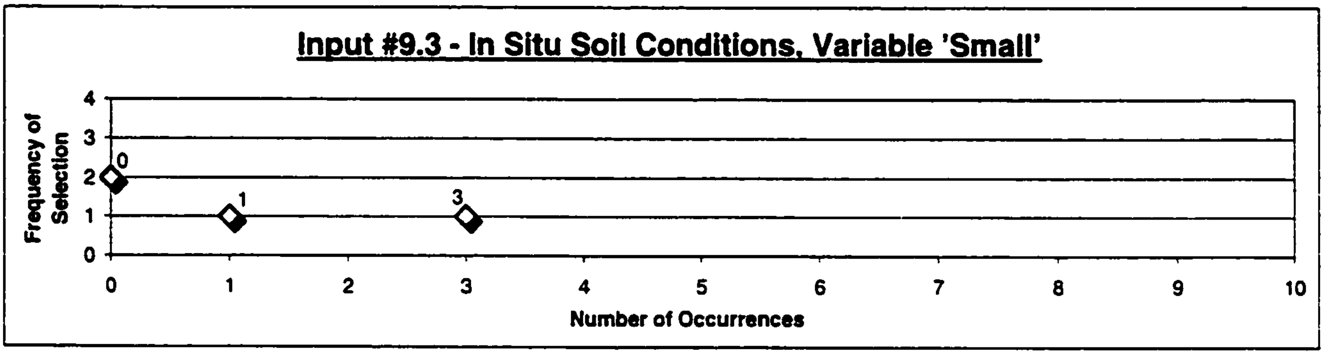

Input #9.2 - Restricted Access to Site, Variable 'Small'
Frequency of Selection
Number of Occurrences
0
1
2
6
8


Input #9.2 - Restricted Access to Site, Variable 'Average'
Frequency of Selection
Number of Occurrences
2
5
10

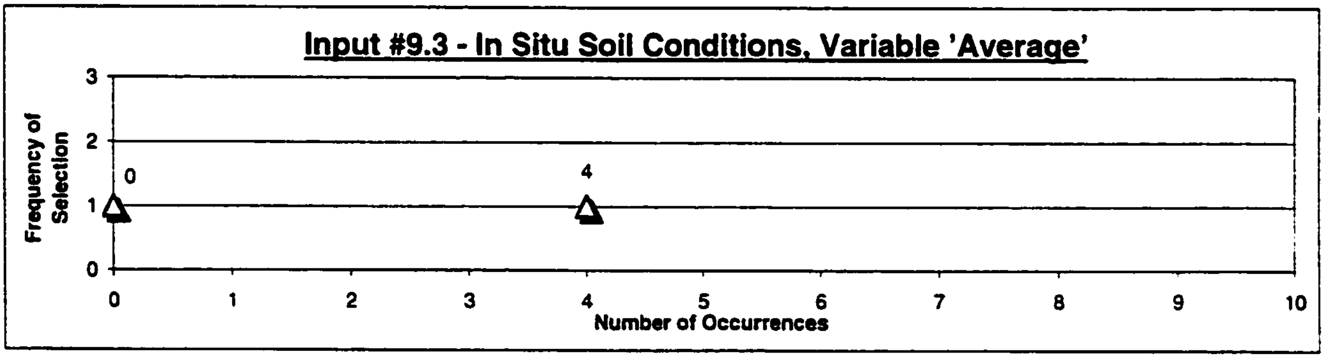

Input #9.2 - Restricted Access to Site, Variable 'Large'
Frequency of Selection
Number of Occurrences
0
10
20
30
40
50
60



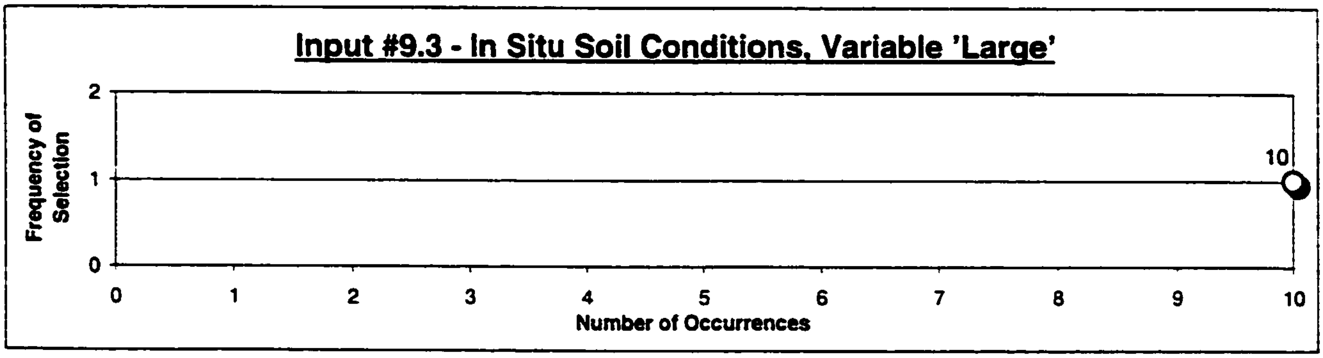
Input #9.2 - Restricted Access to Site - Rating
Small
Average
Large
Frequency of Selection
Rating



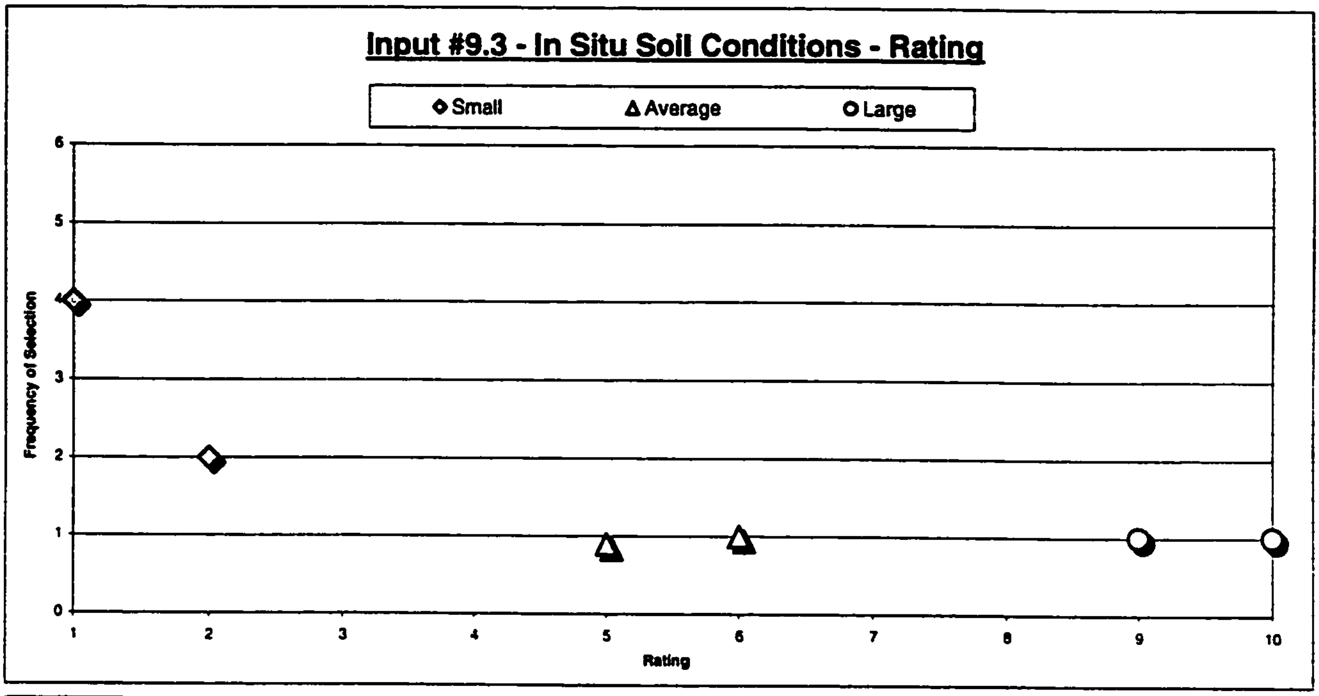
Input #9.3 - In Situ Soil Conditions, Variable 'Small'
Frequency of Selection
Number of Occurrences

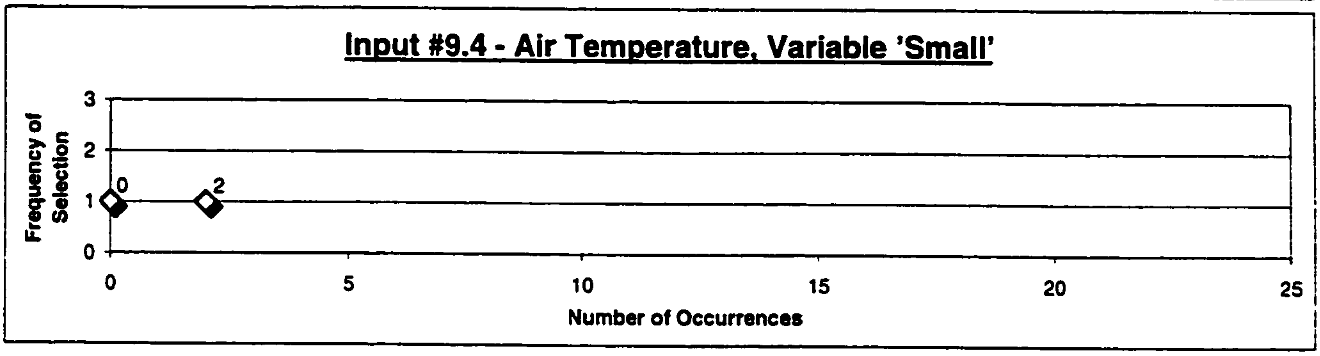


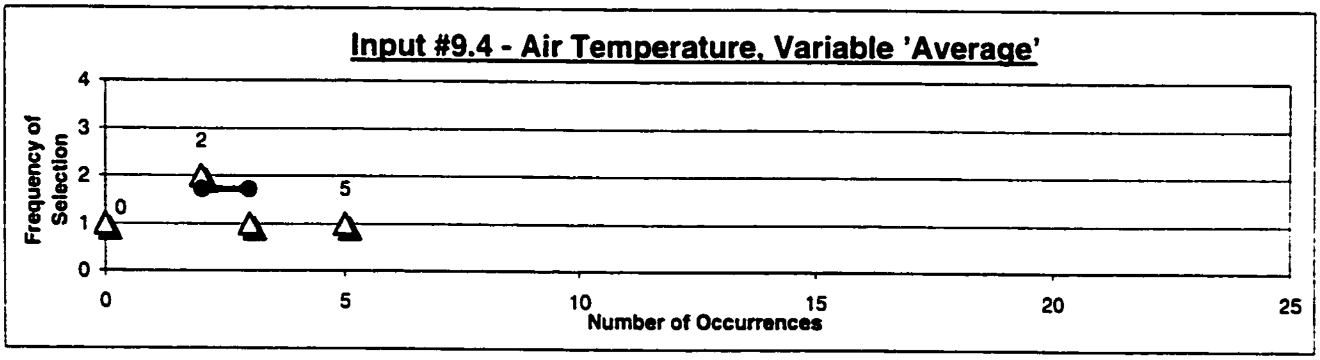
Input #9.3 - In Situ Soil Conditions, Variable 'Average'
Frequency of Selection
Number of Occurrences

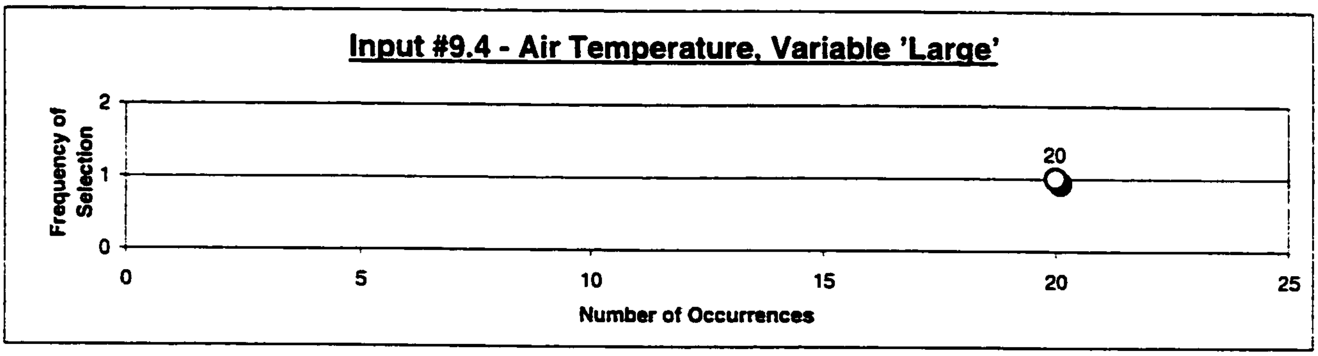


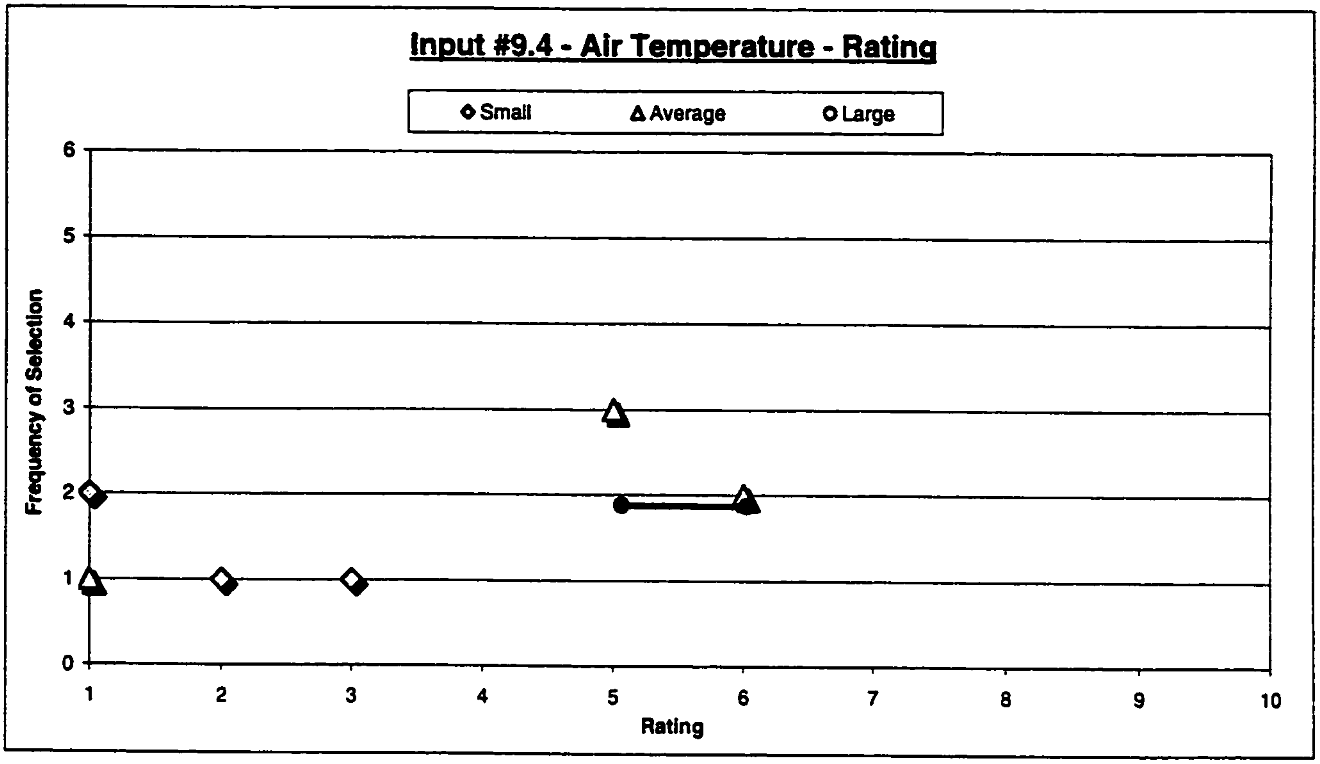
Input #9.3 - In Situ Soil Conditions, Variable 'Large'
Frequency of Selection
Number of Occurrences

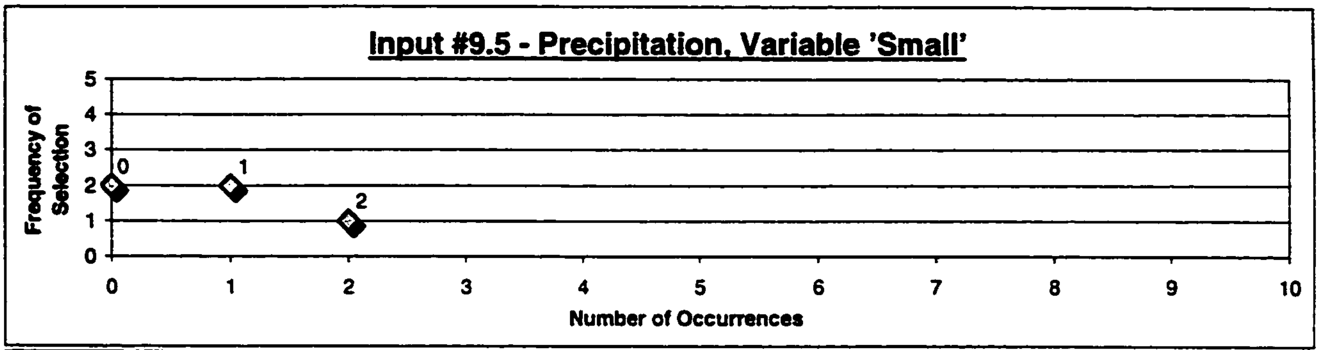

Input #9.3 - In Situ Soil Conditions - Rating
Small
Average
Large
Frequency of Selection
Rating
0
1
2
3
4
5
6
1
2
3
4
5
6
7
8
9
10


Input #9.4 - Air Temperature, Variable 'Small'
Frequency of Selection
Number of Occurrences
0
1
2
3
0
5
10
15
20
25
0
2

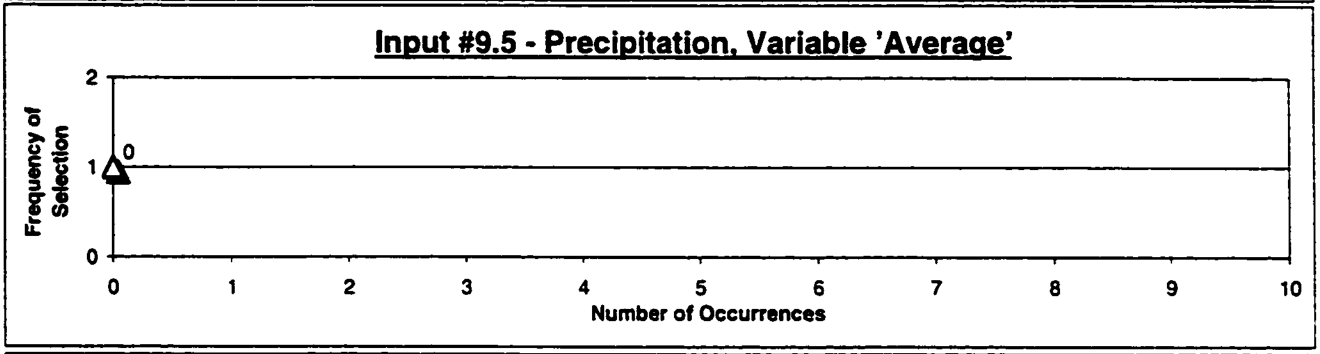

Input #9.4 - Air Temperature, Variable 'Average'
Frequency of Selection
Number of Occurrences
0
1
2
3
4
0
5
10
15
20
25
0
2
5


Input #9.4 - Air Temperature, Variable 'Large'
Frequency of Selection
Number of Occurrences
0
1
2
0
5
10
15
20
25
20

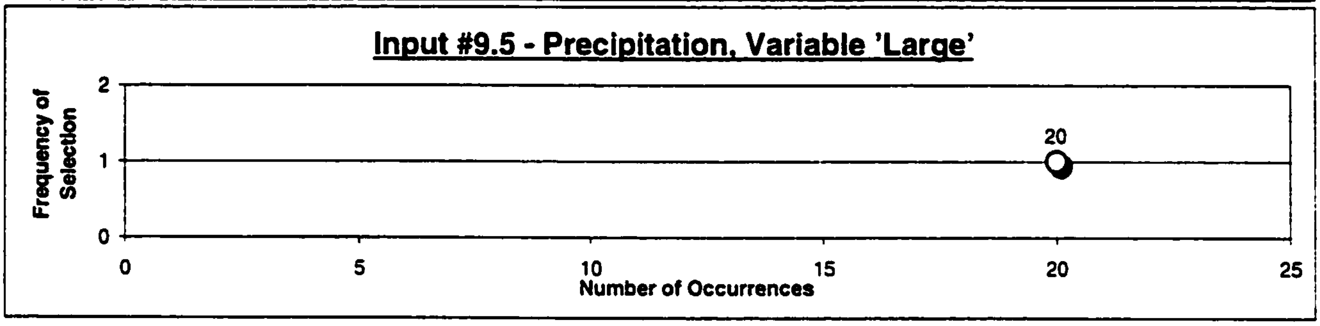

Input #9.4 - Air Temperature - Rating
Small
Average
Large
Frequency of Selection
Rating
0
1
2
3
4
5
6
1
2
3
4
5
6
7
8
9
10



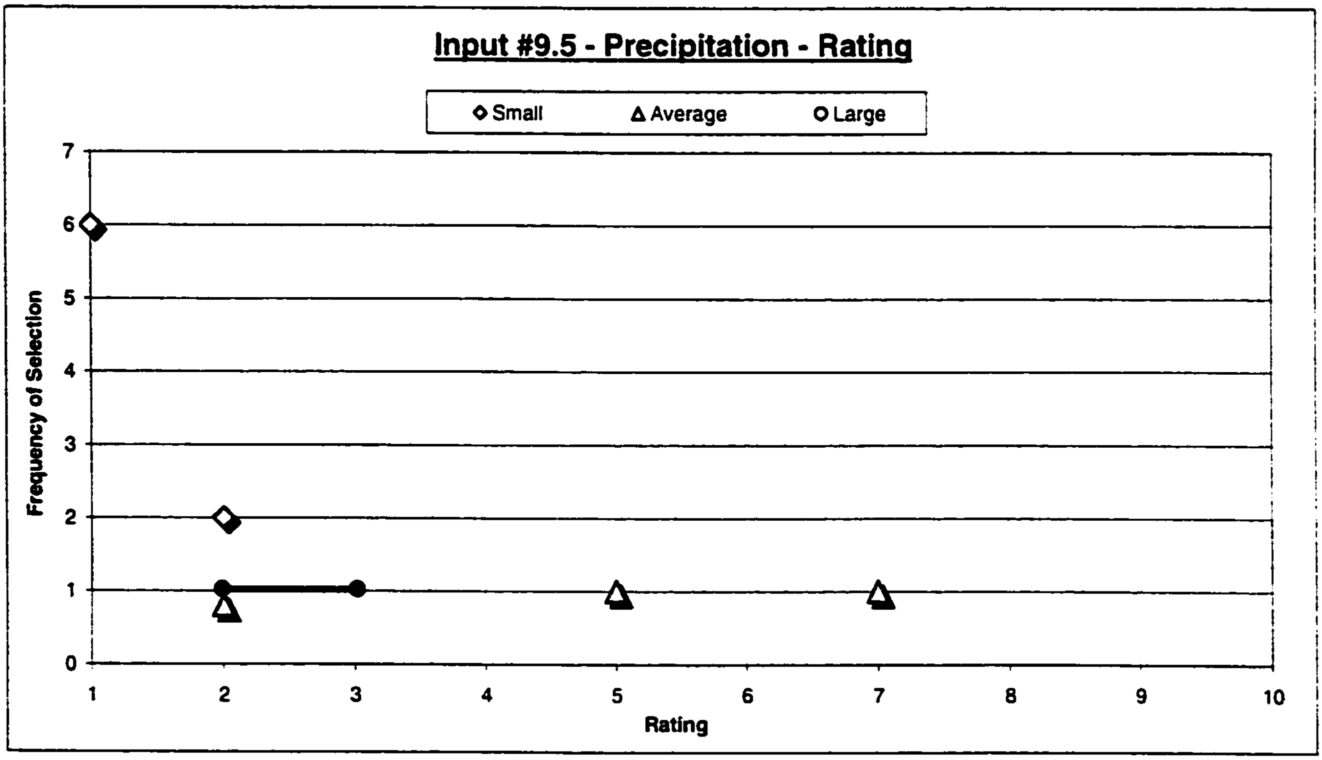
Input #9.5 - Precipitation, Variable 'Small'
Frequency of Selection
Number of Occurrences



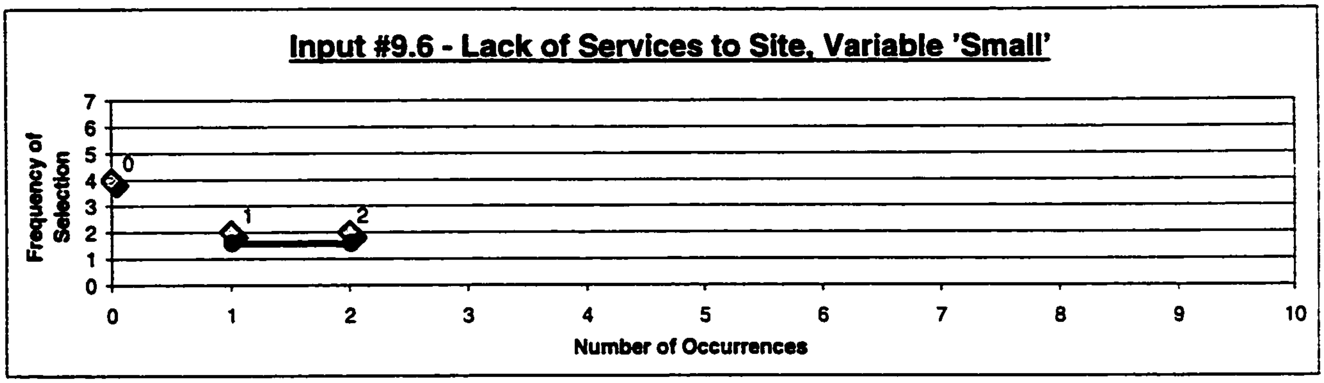
Input #9.5 - Precipitation, Variable 'Average'
Frequency of Selection
Number of Occurrences

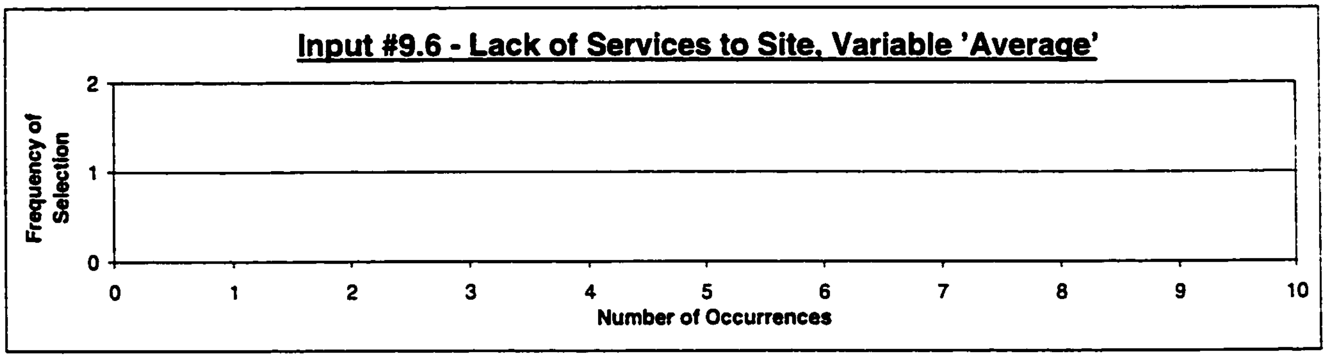


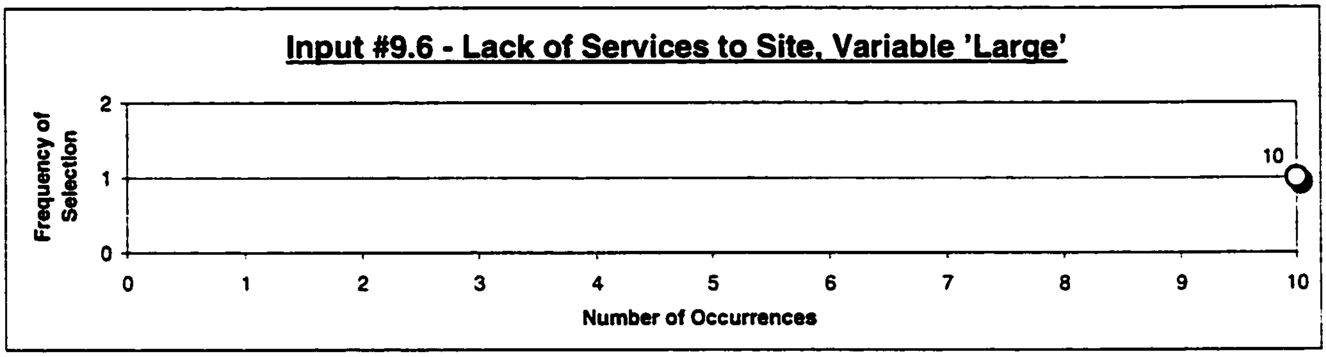
Input #9.5 - Precipitation, Variable 'Large'
Frequency of Selection
Number of Occurrences

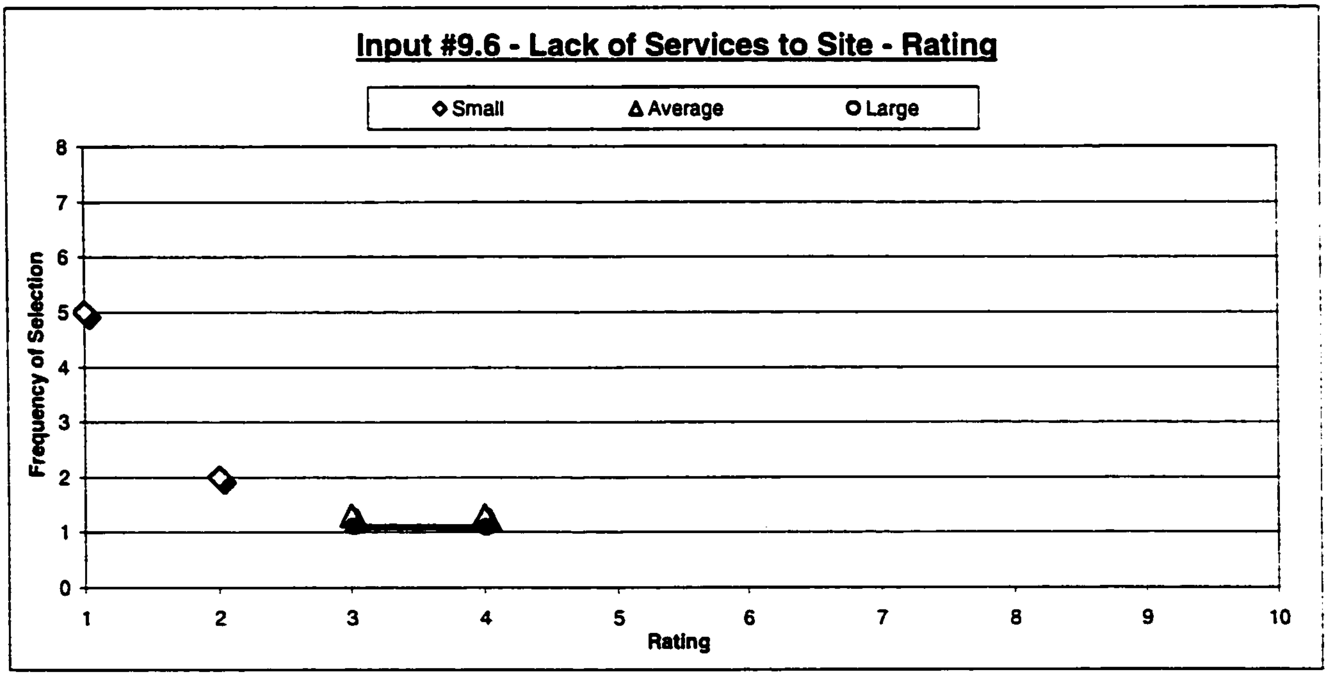


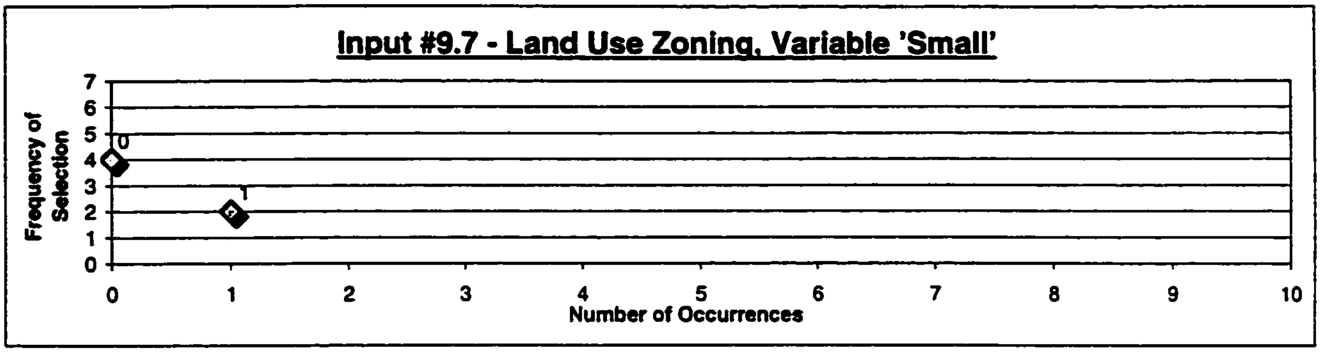
Input #9.5 - Precipitation - Rating
Small
Average
Large
Frequency of Selection
Rating

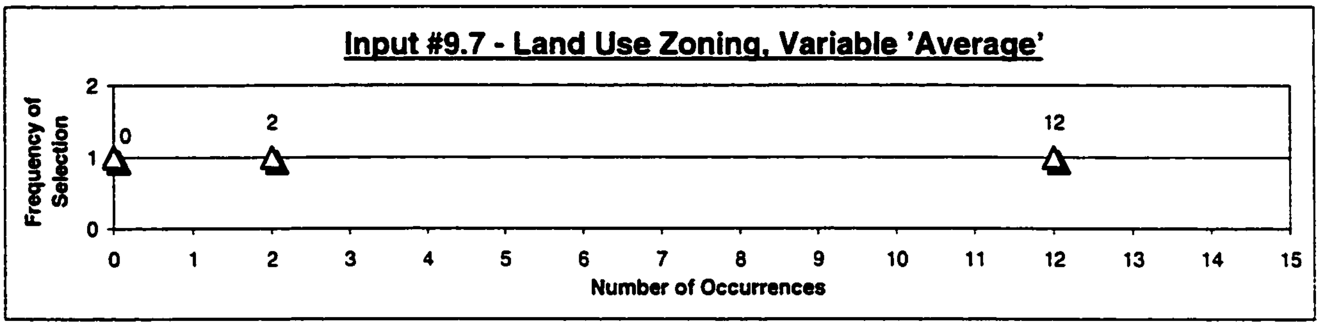

Input #9.6 - Lack of Services to Site, Variable 'Small'
Frequency of Selection
Number of Occurrences

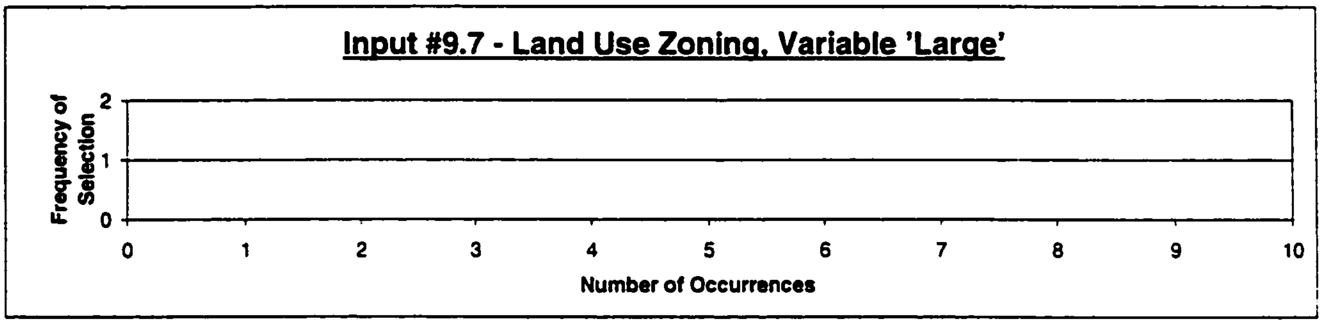

Input #9.6 - Lack of Services to Site, Variable 'Average'
Frequency of Selection
Number of Occurrences

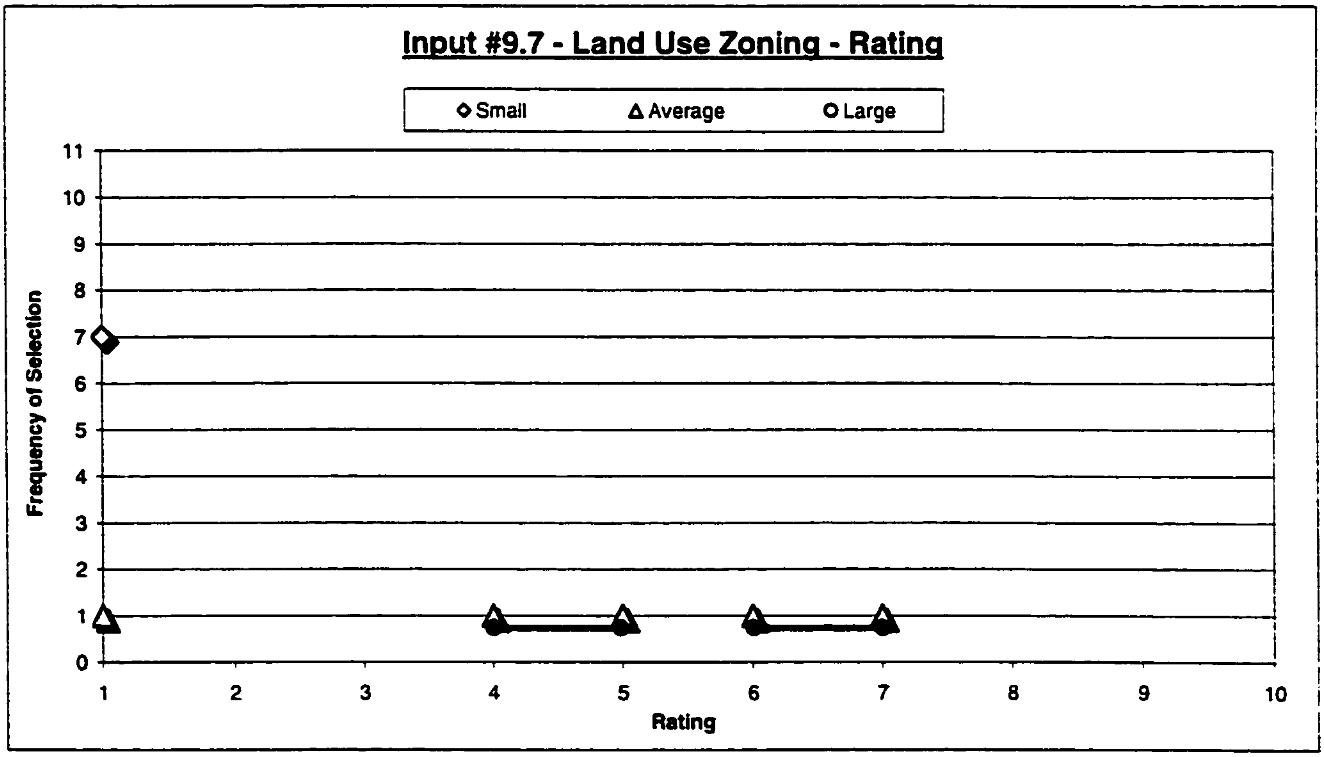

Input #9.6 - Lack of Services to Site, Variable 'Large'
Frequency of Selection
Number of Occurrences

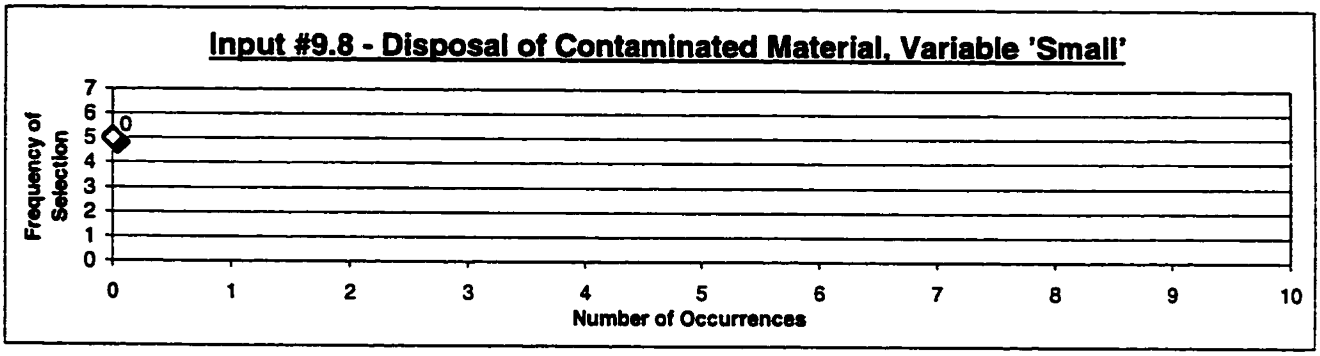

Input #9.6 - Lack of Services to Site - Rating
Small
Average
Large
Frequency of Selection
Rating

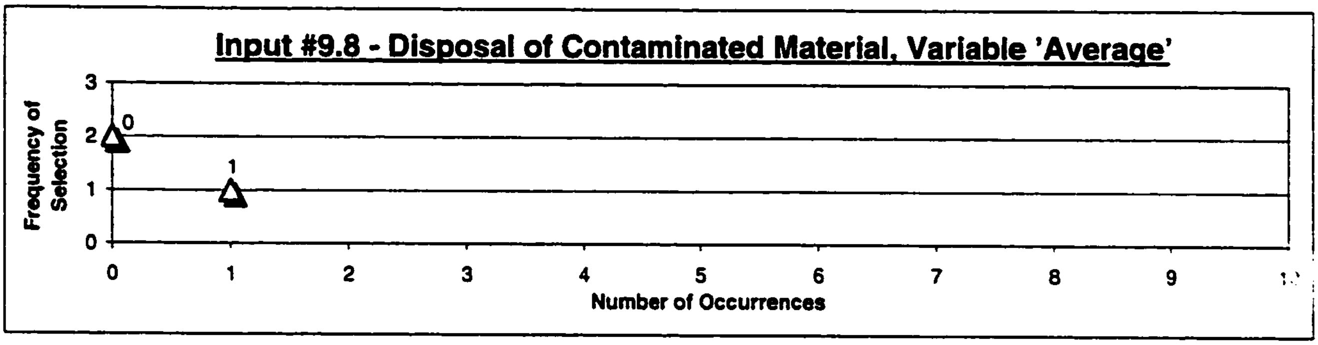

Input #9.7 - Land Use Zoning, Variable 'Small'
Frequency of Selection
Number of Occurrences

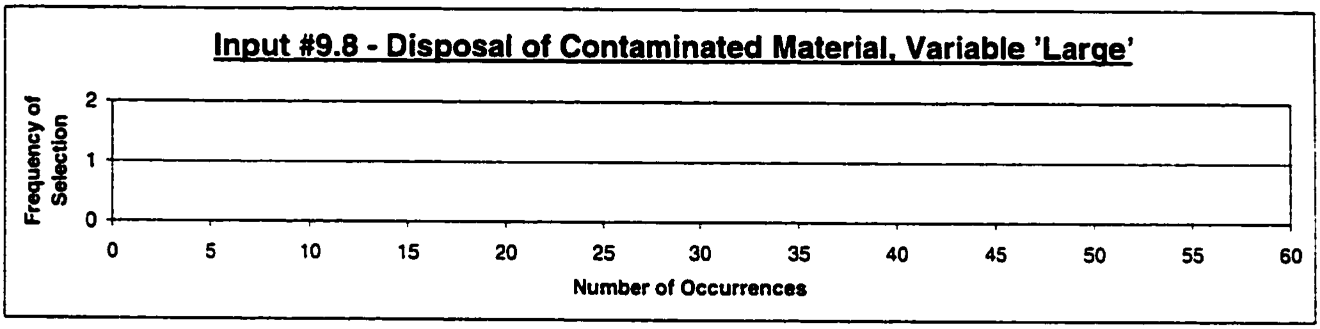

Input #9.7 - Land Use Zoning, Variable 'Average'
Frequency of Selection
Number of Occurrences

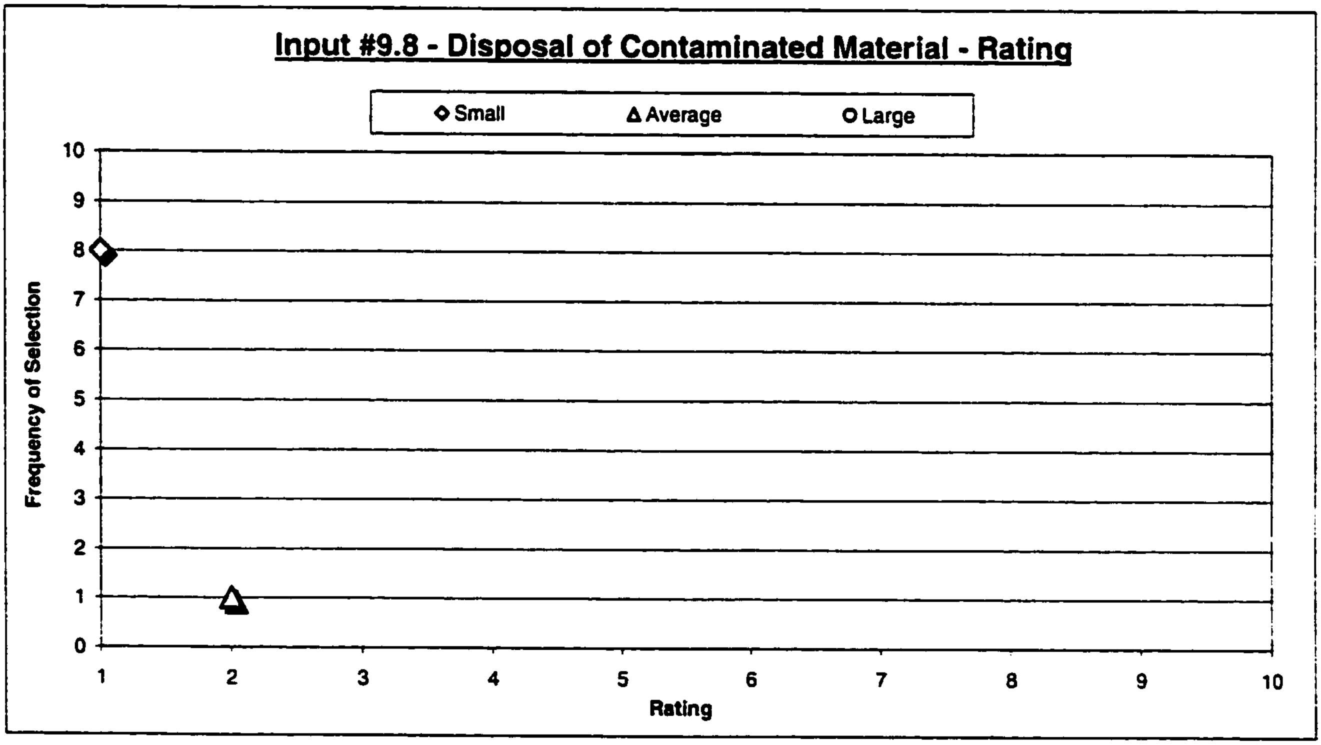

Input #9.7 - Land Use Zoning, Variable 'Large'
Frequency of Selection
Number of Occurrences

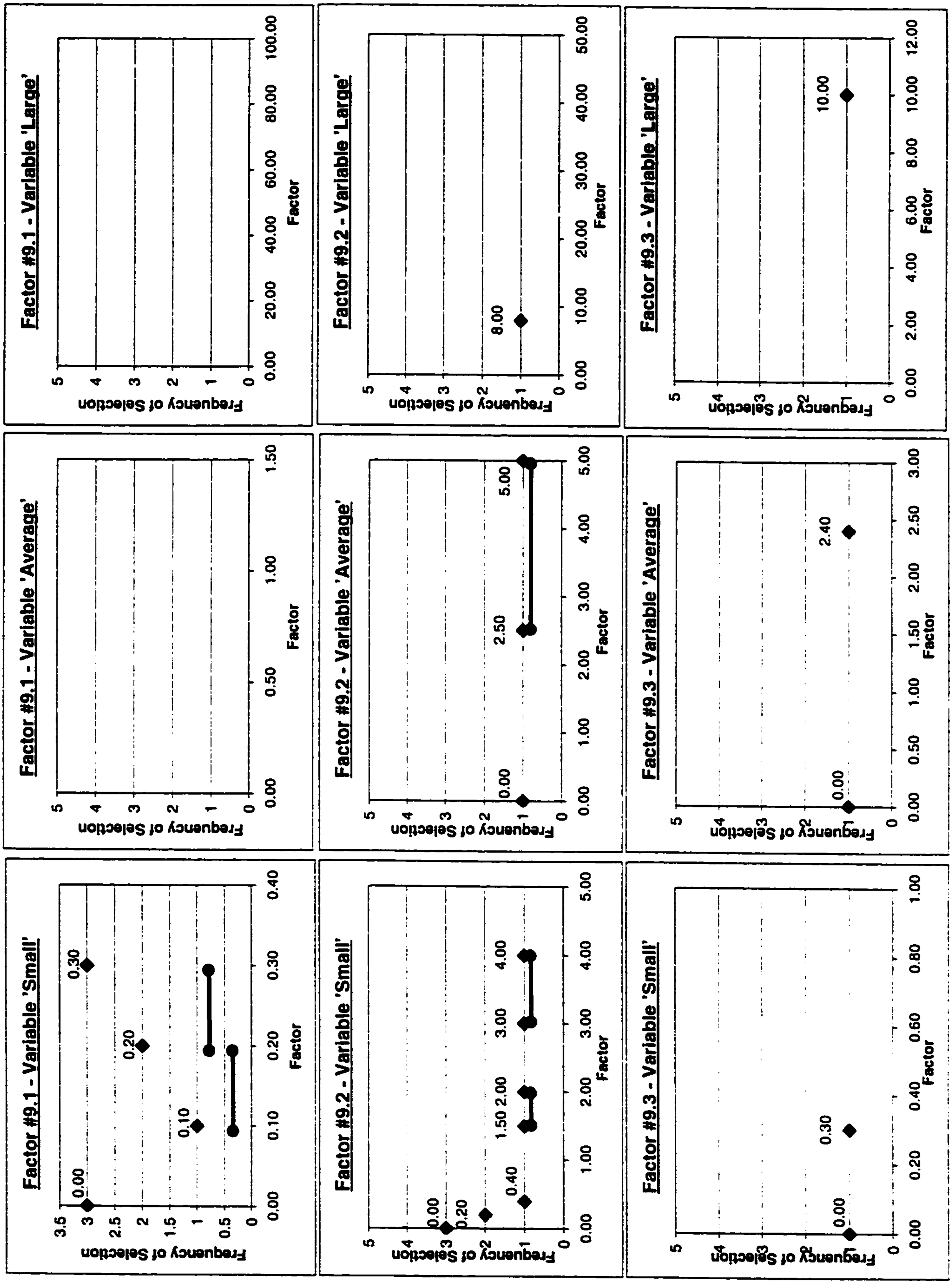

Input #9.7 - Land Use Zoning - Rating
Small
Average
Large
Frequency of Selection
Rating


Input #9.8 - Disposal of Contaminated Material, Variable 'Small'
Frequency of Selection
Number of Occurrences
0 1 2 3 4 5 6 7
0 1 2 3 4 5 6 7 8 9 10
0


Input #9.8 - Disposal of Contaminated Material, Variable 'Average'
Frequency of Selection
Number of Occurrences
0 1 2 3
0 1 2 3 4 5 6 7 8 9
0
1

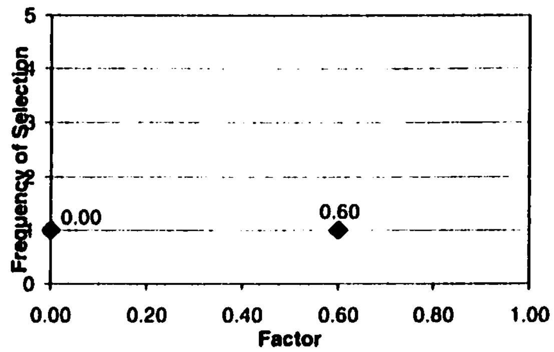

Input #9.8 - Disposal of Contaminated Material, Variable 'Large'
Frequency of Selection
Number of Occurrences
0 1 2
0 5 10 15 20 25 30 35 40 45 50 55 60


Input #9.8 - Disposal of Contaminated Material - Rating
Small
Average
Large
Frequency of Selection
Rating
0 1 2 3 4 5 6 7 8 9 10
1 2 3 4 5 6 7 8 9 10

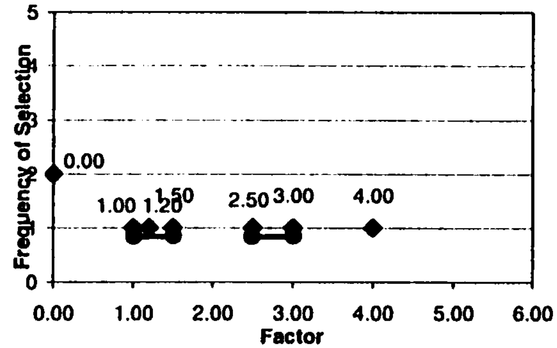


Factor #9.1 - Variable 'Small'
Factor #9.1 - Variable 'Average'
Factor #9.1 - Variable 'Large'
Factor #9.2 - Variable 'Small'
Factor #9.2 - Variable 'Average'
Factor #9.2 - Variable 'Large'
Factor #9.3 - Variable 'Small'
Factor #9.3 - Variable 'Average'
Factor #9.3 - Variable 'Large'
Frequency of Selection
Factor
0.00
0.10
0.20
0.30
0.20
0.40
1.50 2.00
3.00
4.00
2.50
5.00
8.00
2.40
10.00

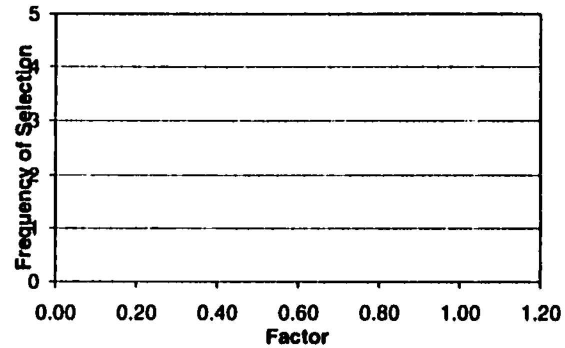
Factor #9.4 - Variable 'Small'



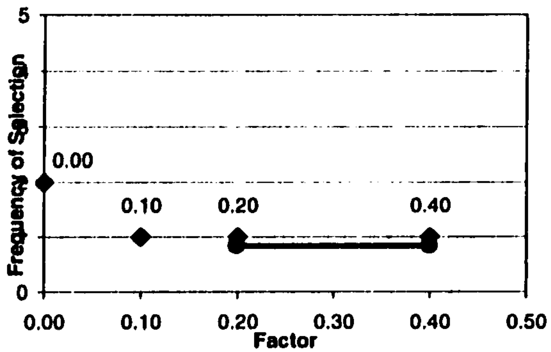
Factor #9.4 - Variable 'Average'



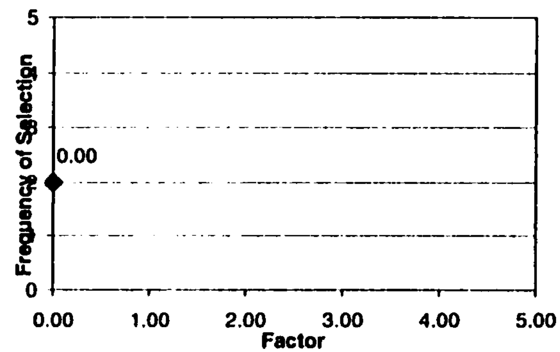
Factor #9.4 - Variable 'Large'



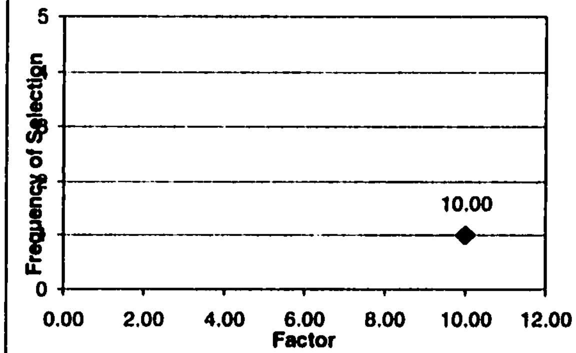
Factor #9.5 - Variable 'Small'



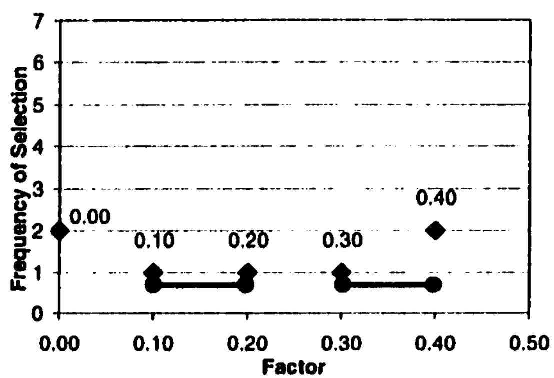
Factor #9.5 - Variable 'Average'



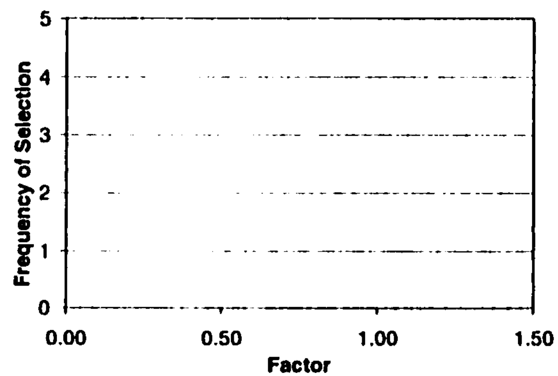
Factor #9.5 - Variable 'Large'



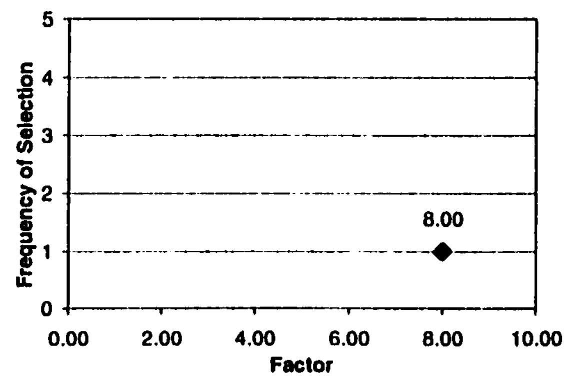
Factor #9.6 - Variable 'Small'

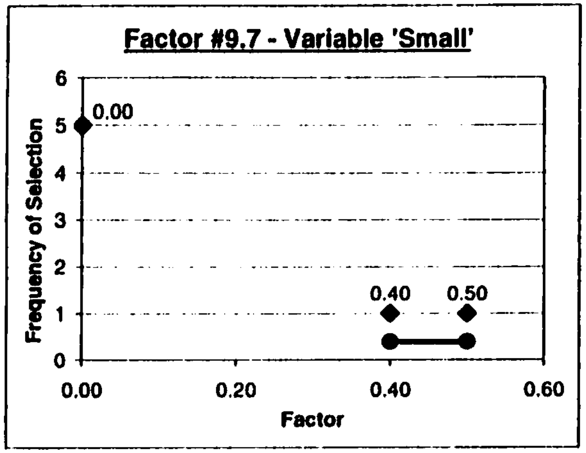


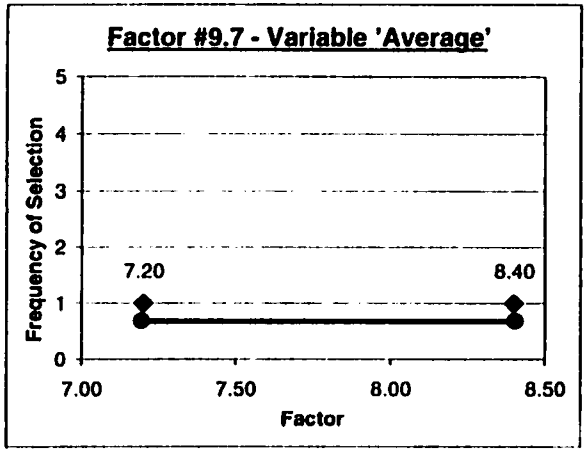
Factor #9.6 - Variable 'Average'

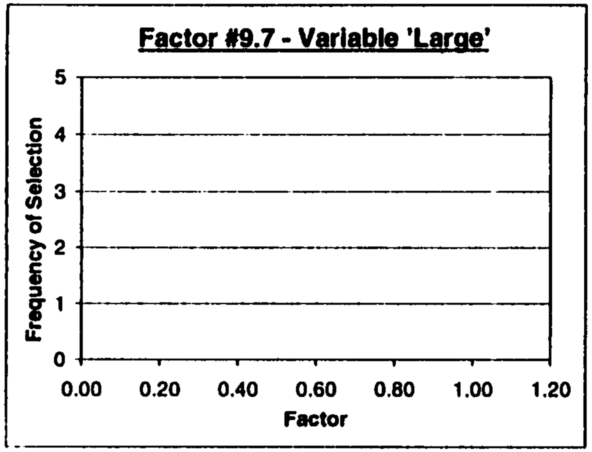


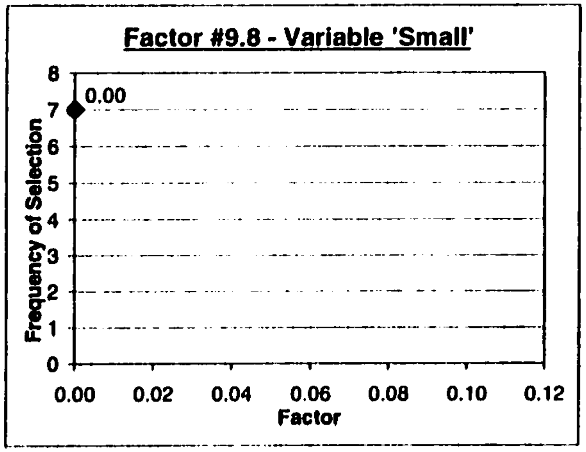
Factor #9.6 - Variable 'Large'

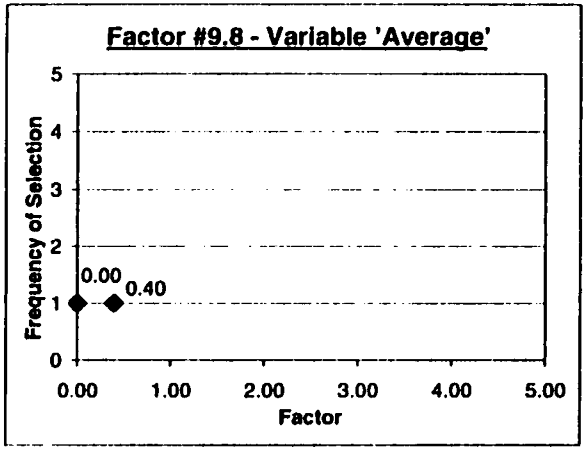


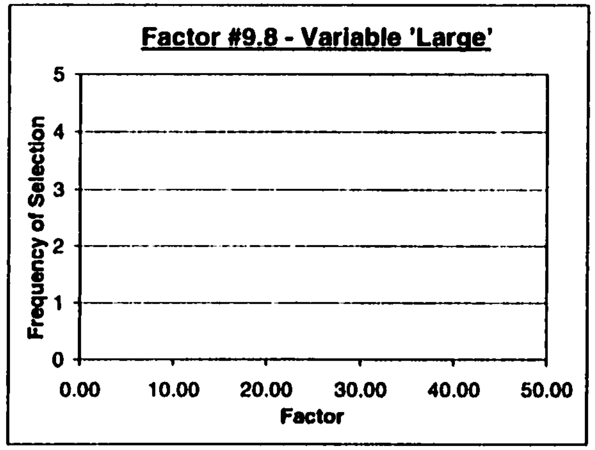

Factor #9.7 - Variable 'Small'
Frequency of Selection
Factor
0.00
0.40
0.50


Factor #9.7 - Variable 'Average'
Frequency of Selection
Factor
7.20
8.40

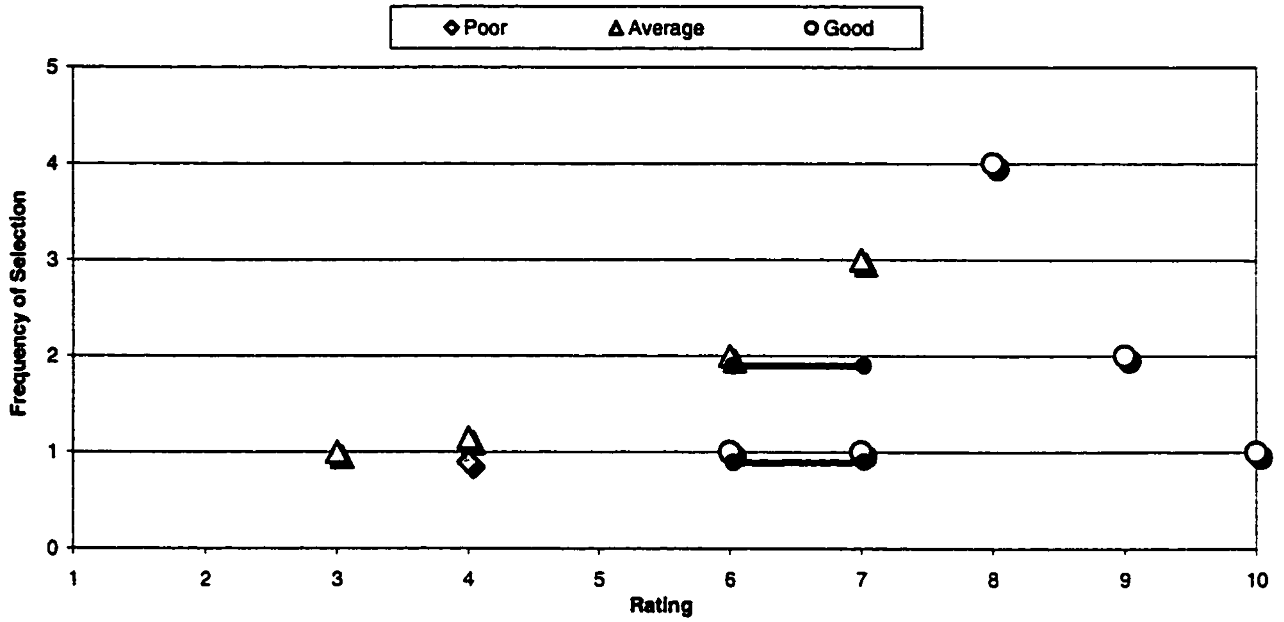

Factor #9.7 - Variable 'Large'
Frequency of Selection
Factor


Factor #9.8 - Variable 'Small'
Frequency of Selection
Factor
0.00

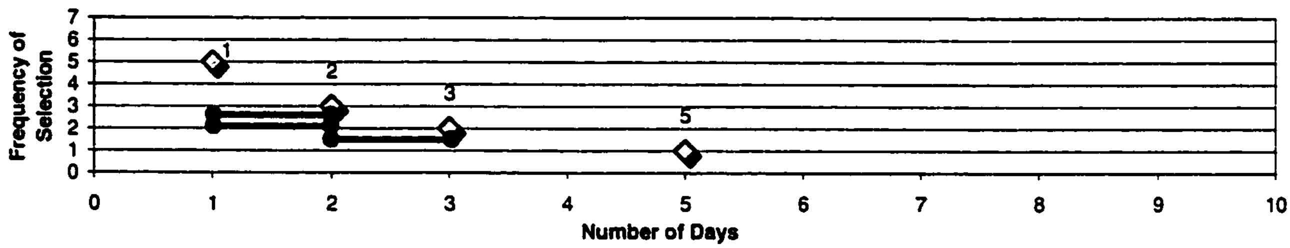

Factor #9.8 - Variable 'Average'
Frequency of Selection
Factor
0.00
0.40


Factor #9.8 - Variable 'Large'
Frequency of Selection
Factor

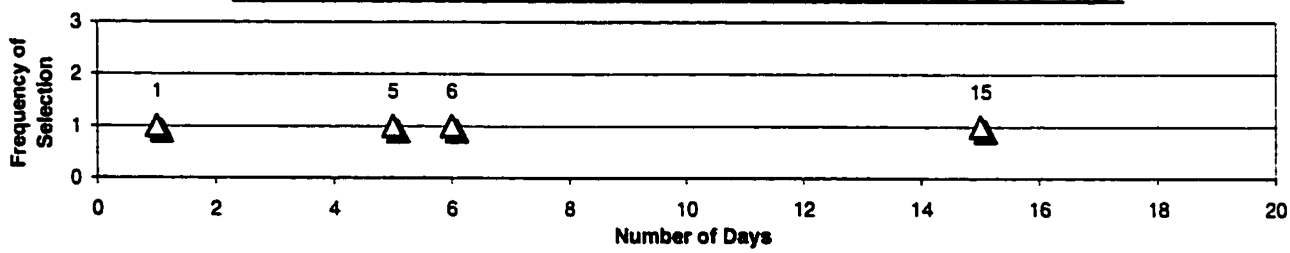
Input #10 - Overall Quality of Owner



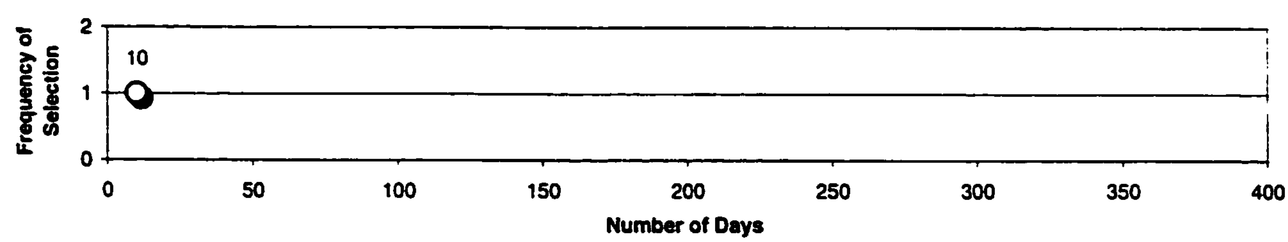
Input #10.1 - Length of Owner's Decisions, Variable 'Short'

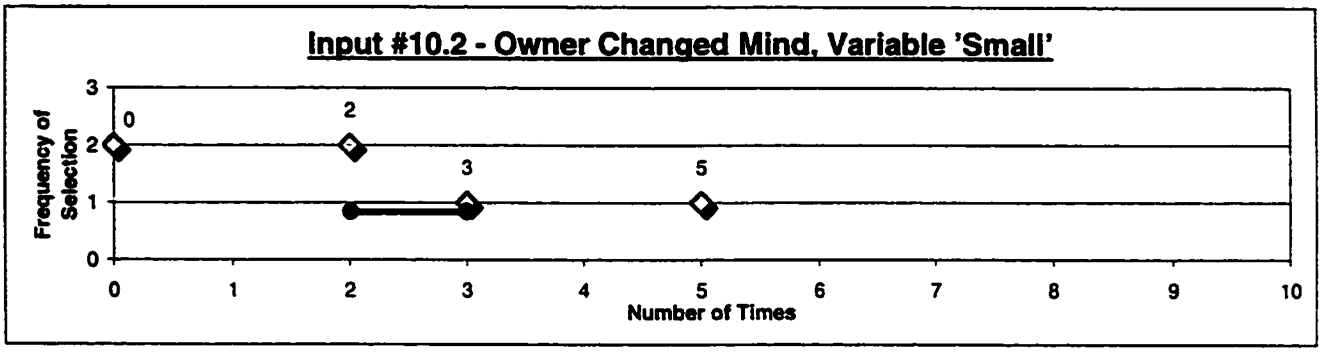


Input #10.1 - Length of Owner's Decisions, Variable 'Average'

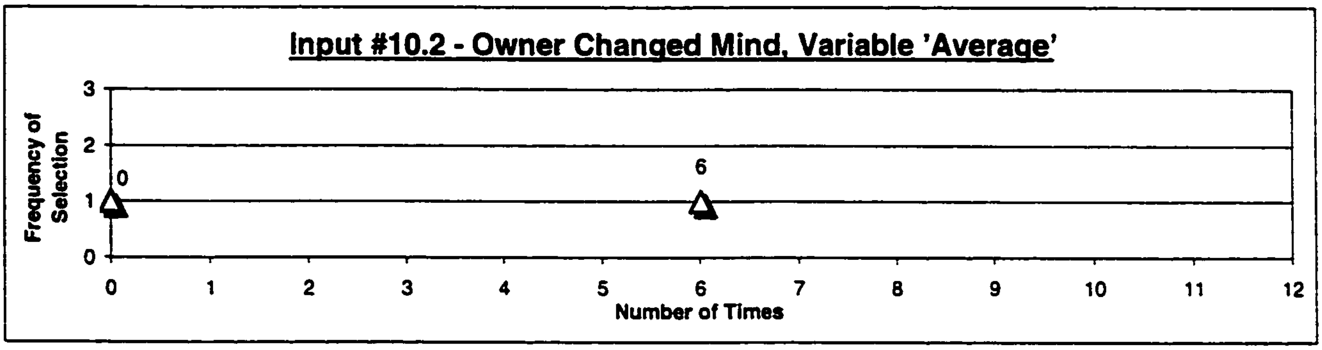


Input #10.1 - Length of Owner's Decisions, Variable 'Long'

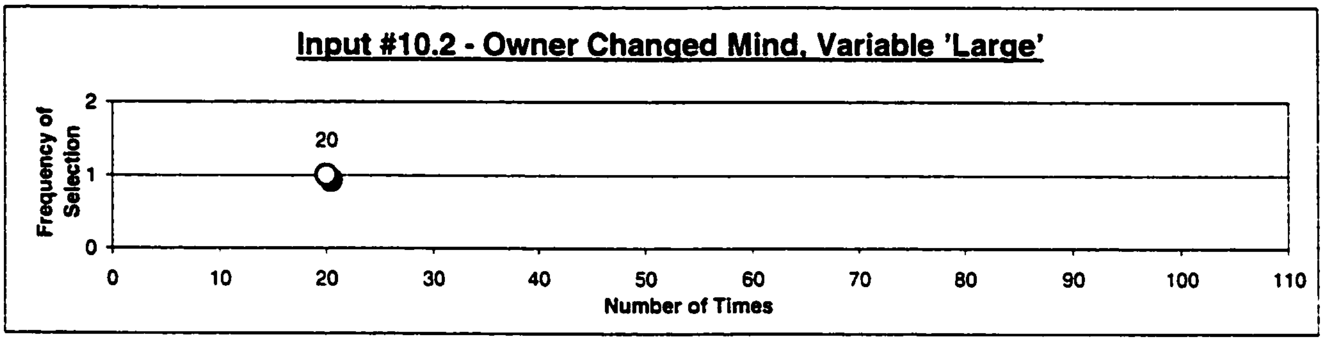




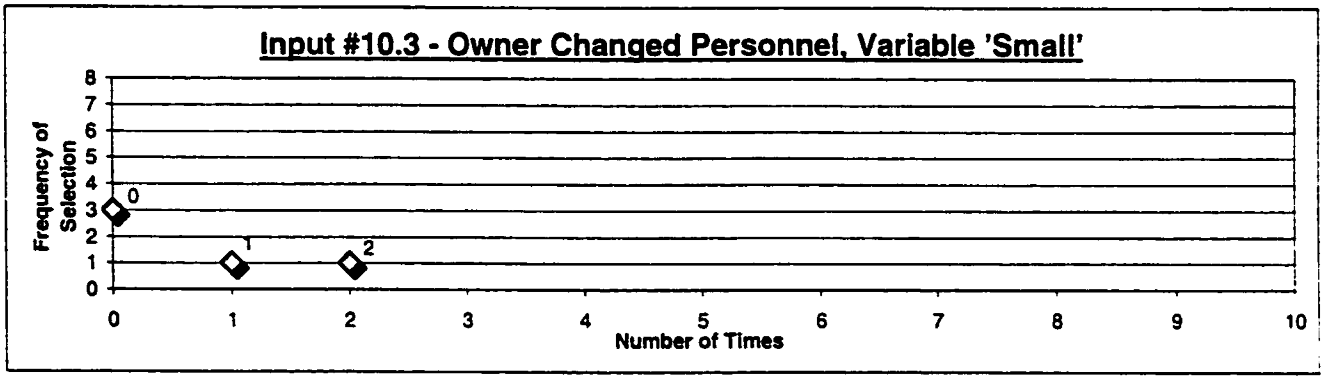
Input #10.2 - Owner Changed Mind, Variable 'Small'
Frequency of Selection
Number of Times



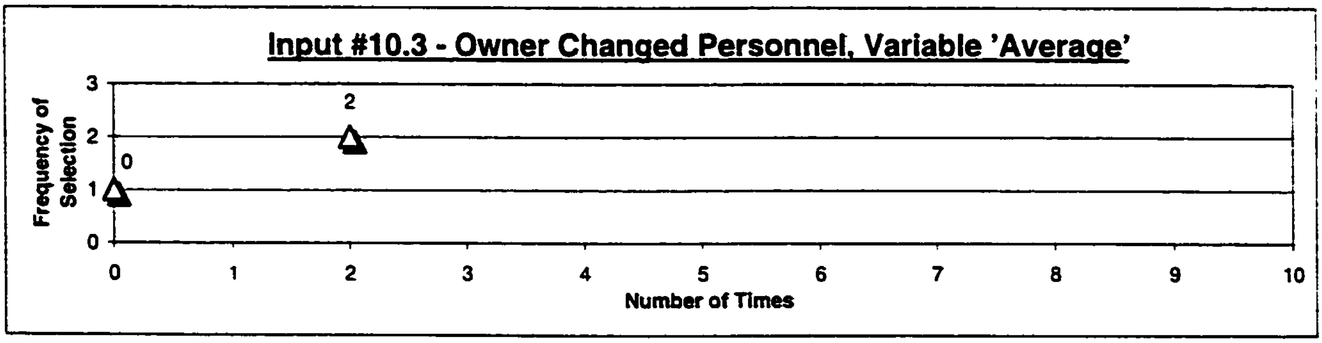
Input #10.2 - Owner Changed Mind, Variable 'Average'
Frequency of Selection
Number of Times



Input #10.2 - Owner Changed Mind, Variable 'Large'
Frequency of Selection
Number of Times

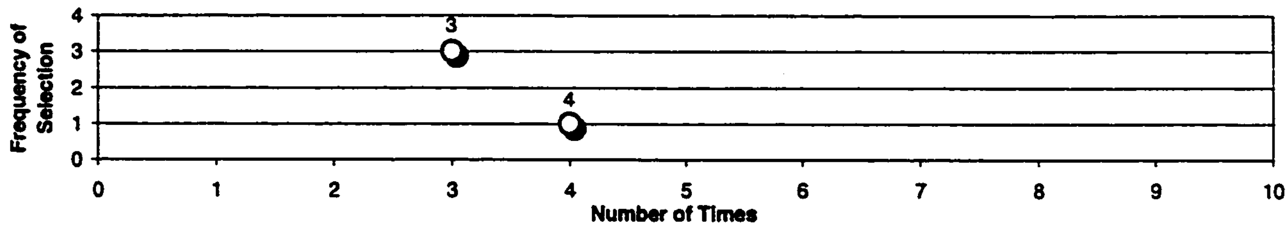


Input #10.3 - Owner Changed Personnel, Variable 'Small'
Frequency of Selection
Number of Times

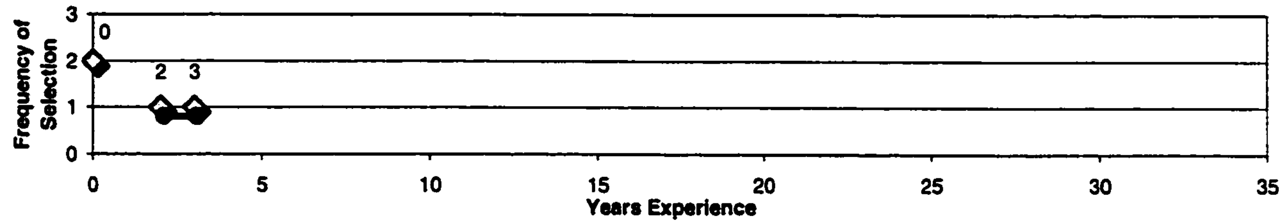


Input #10.3 - Owner Changed Personnel, Variable 'Average'
Frequency of Selection
Number of Times

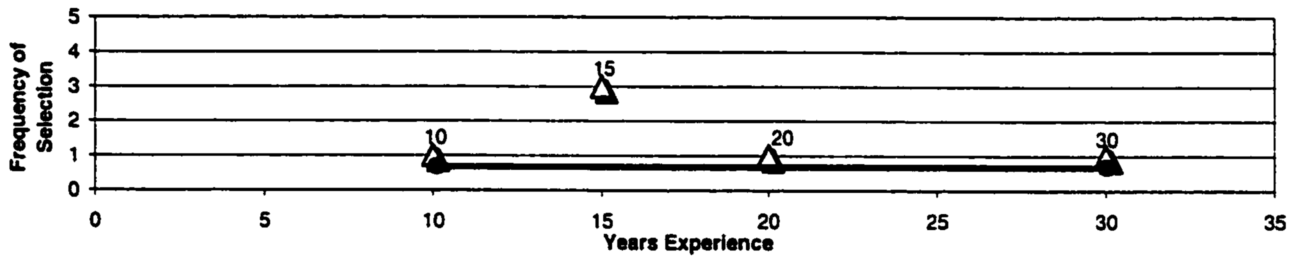
Input #10.3 - Owner Changed Personnel, Variable 'Large'



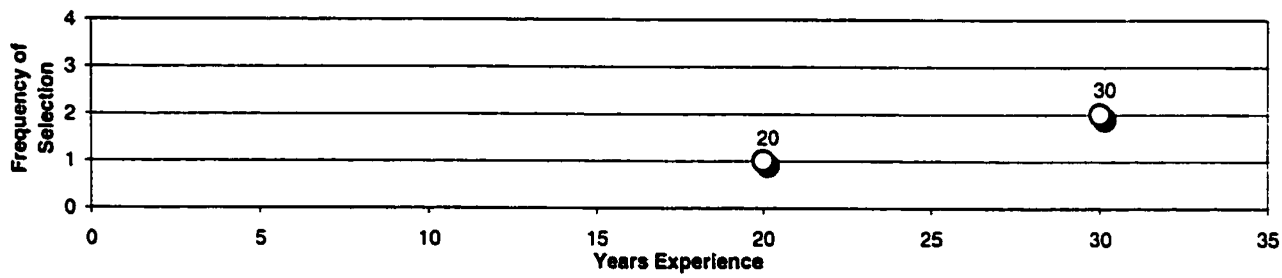
Input #10.4 - Experience of Owner's Rep., Variable 'Small'

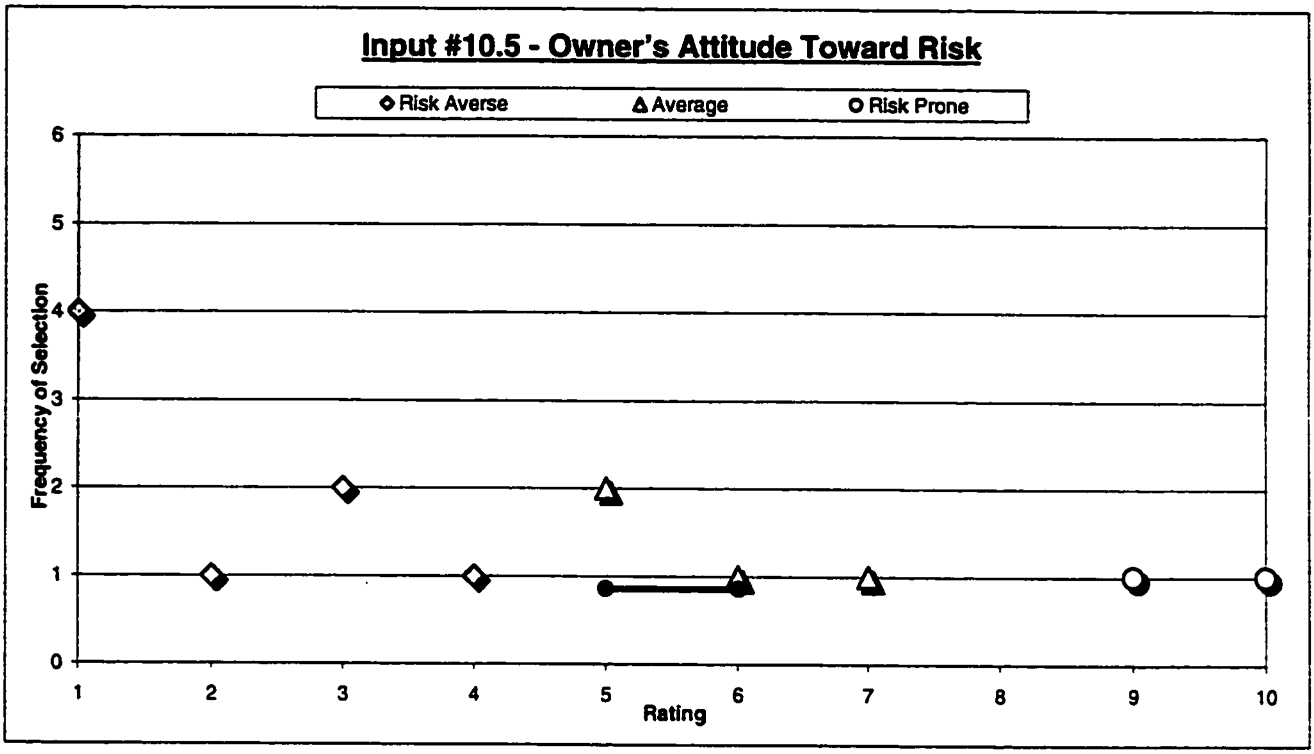


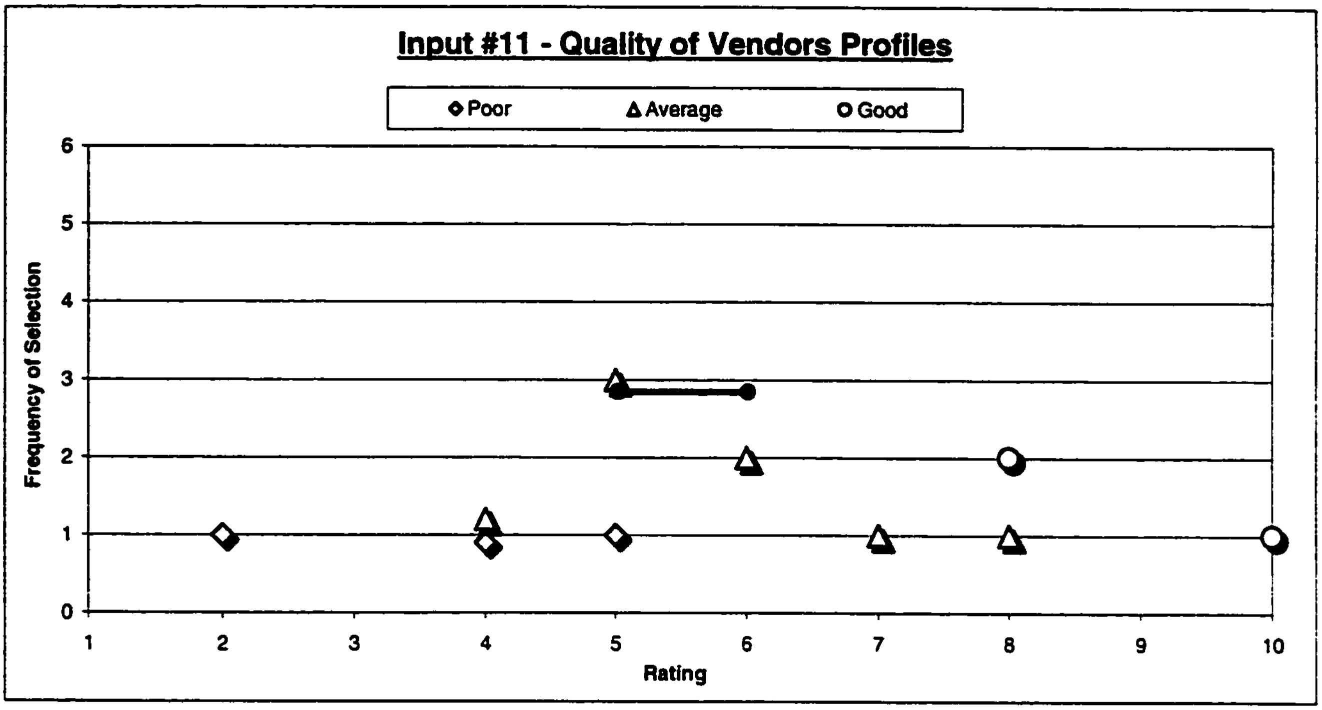
Input #10.4 - Experience of Owner's Rep., Variable 'Average'

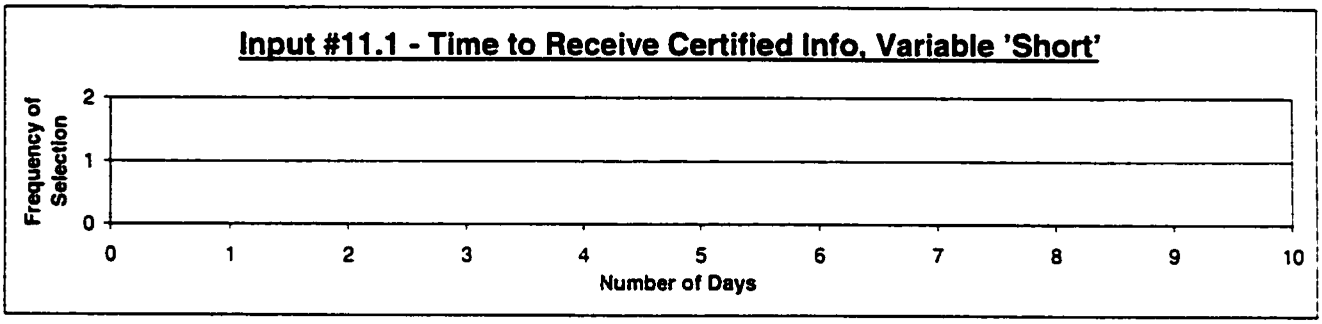


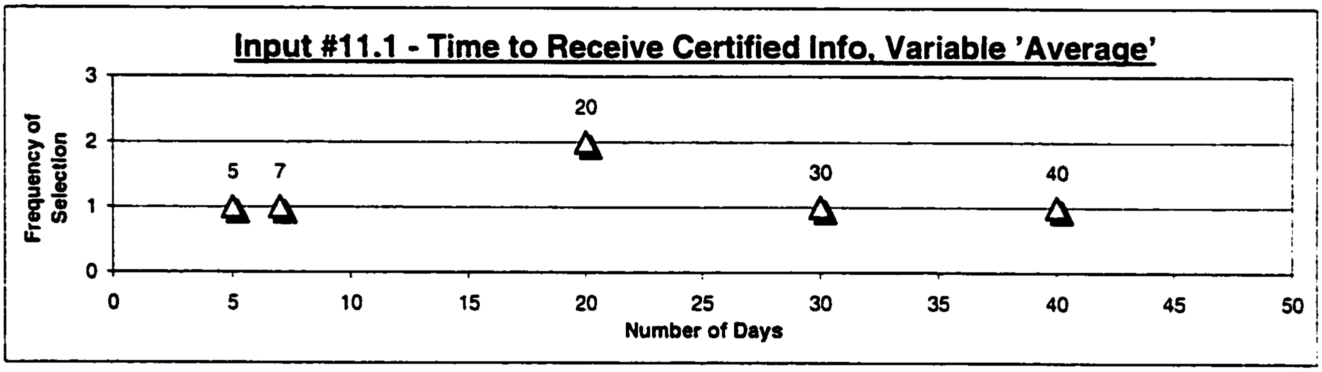
Input #10.4 - Experience of Owner's Rep., Variable 'Large'

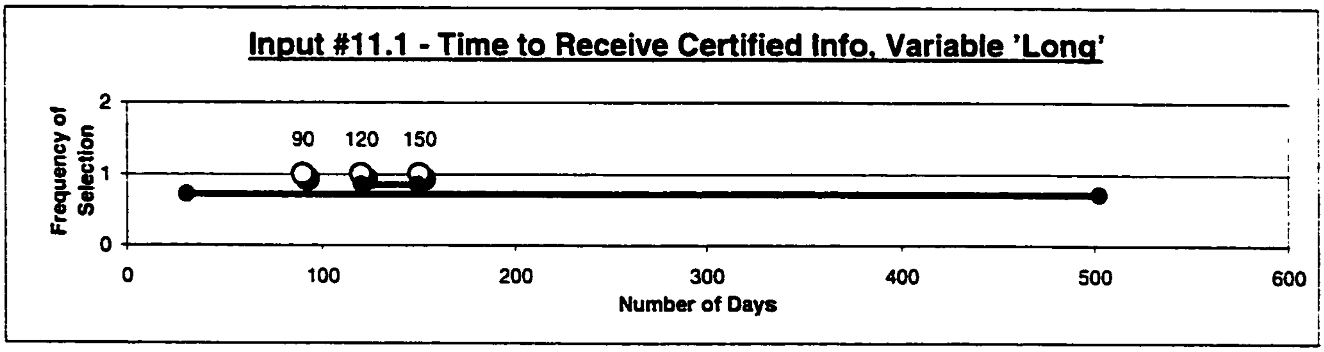


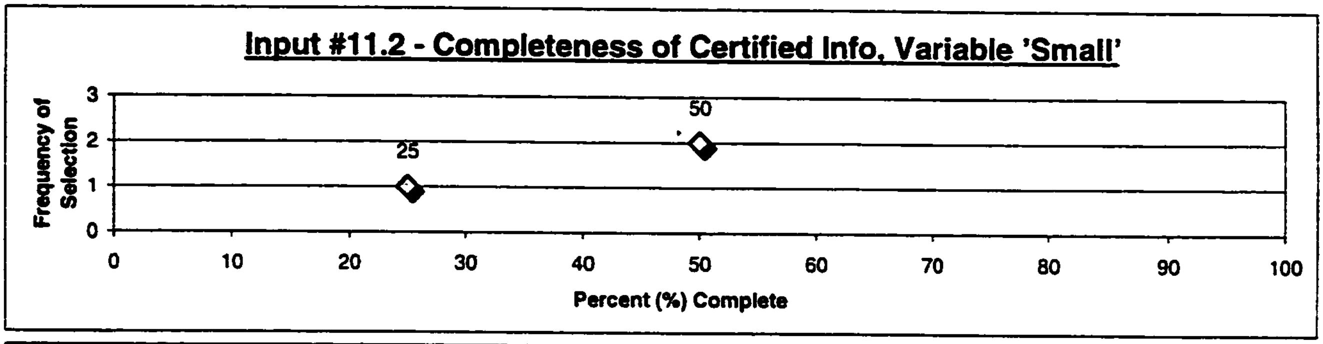

Input #10.5 - Owner's Attitude Toward Risk
Risk Averse
Average
Risk Prone
Frequency of Selection
Rating
0
1
2
3
4
5
6
1
2
3
4
5
6
7
8
9
10

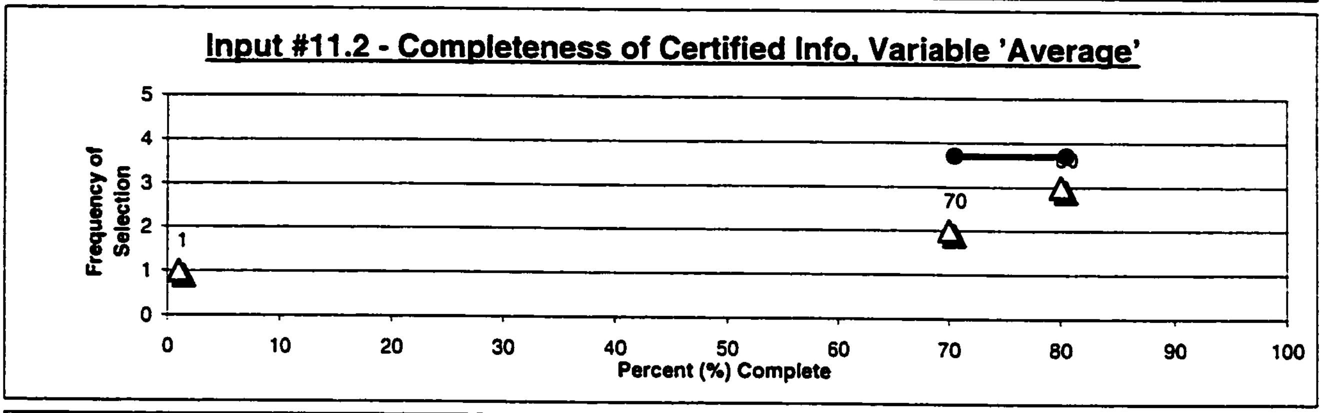

Input #11 - Quality of Vendors Profiles
Poor
Average
Good
Frequency of Selection
Rating
0
1
2
3
4
5
6
1
2
3
4
5
6
7
8
9
10

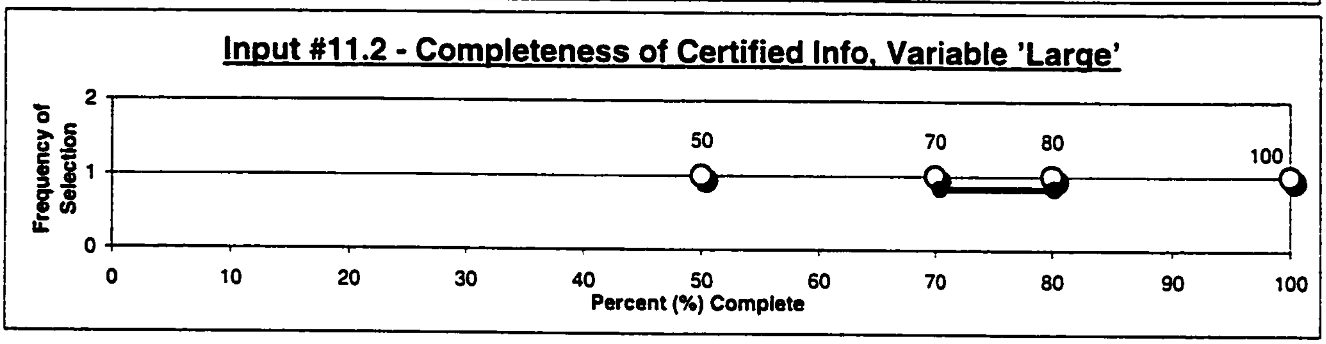

Input #11.1 - Time to Receive Certified Info, Variable 'Short'
Frequency of Selection
Number of Days
0
1
2
0
1
2
3
4
5
6
7
8
9
10

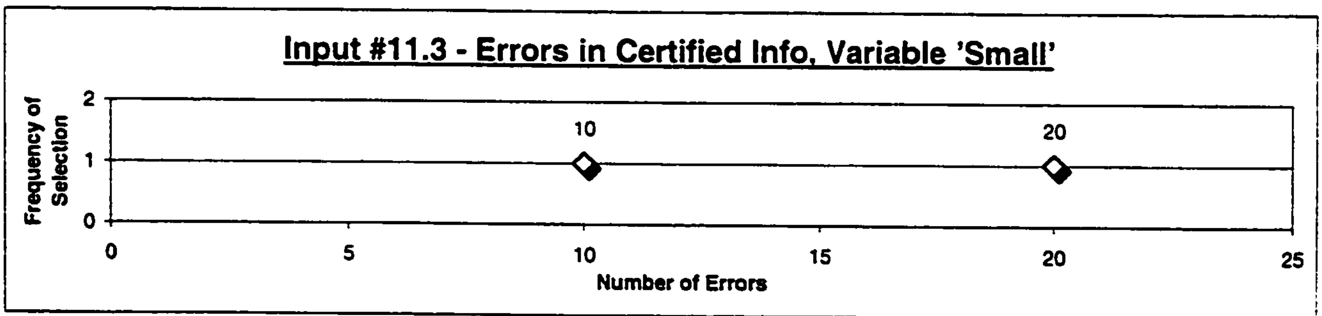

Input #11.1 - Time to Receive Certified Info, Variable 'Average'
Frequency of Selection
Number of Days
0
1
2
3
5
7
20
30
40
0
5
10
15
20
25
30
35
40
45
50

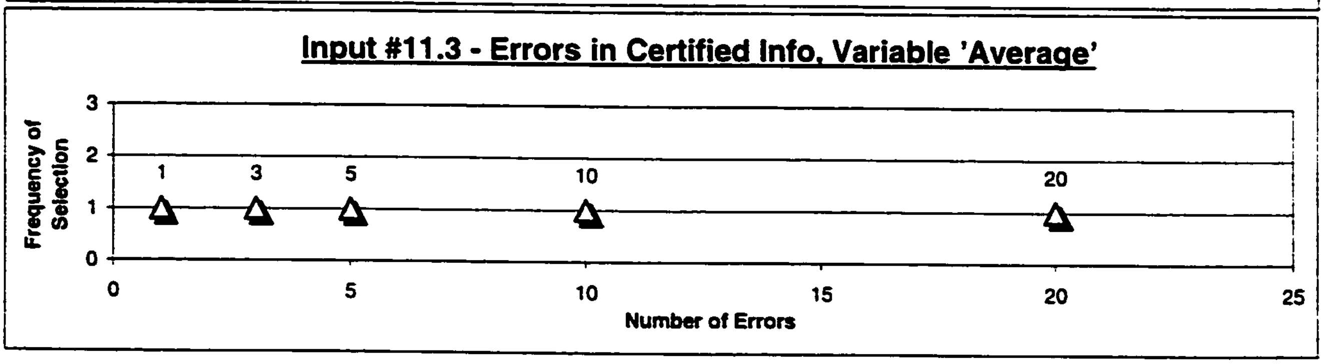

Input #11.1 - Time to Receive Certified Info, Variable 'Long'
Frequency of Selection
Number of Days
0
1
2
90
120
150
0
100
200
300
400
500
600

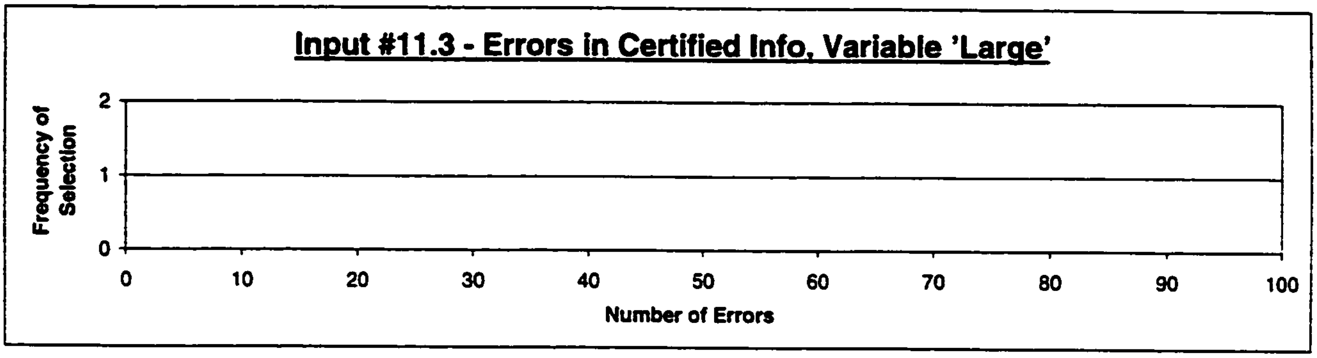

Input #11.2 - Completeness of Certified Info, Variable 'Small'
Frequency of Selection
Percent (%) Complete
25
50
0 10 20 30 40 50 60 70 80 90 100
0 1 2 3

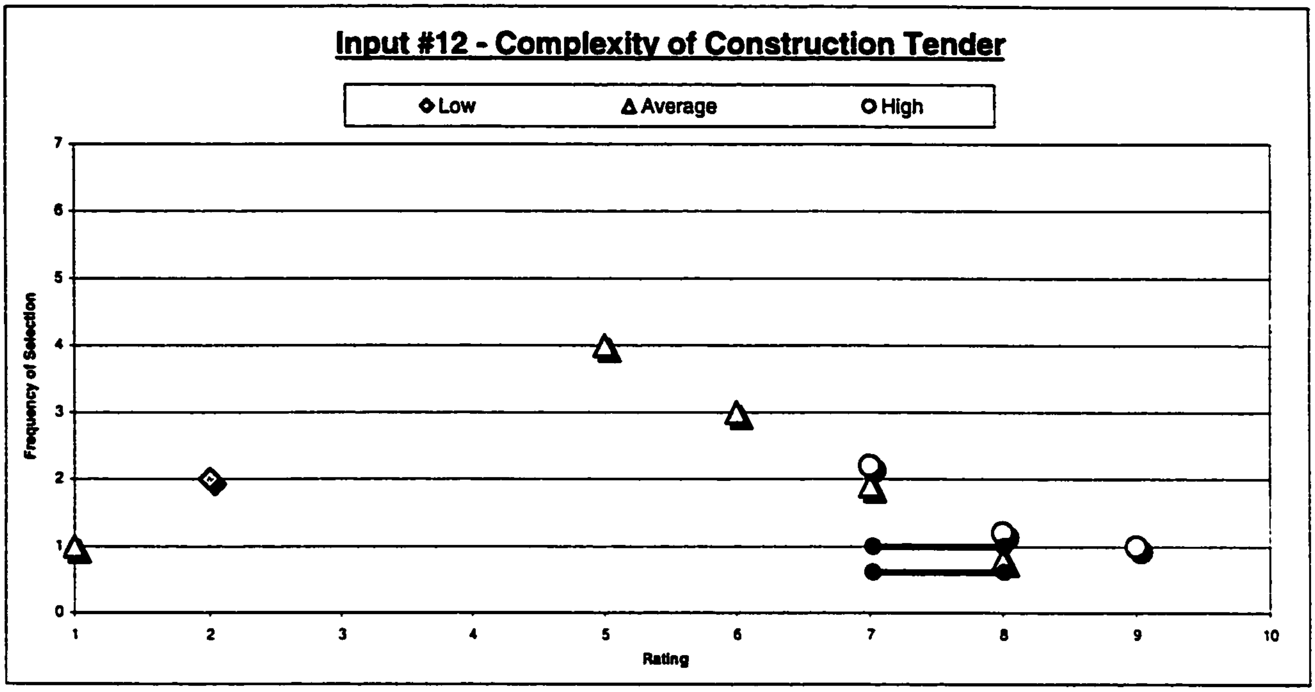

Input #11.2 - Completeness of Certified Info, Variable 'Average'
Frequency of Selection
Percent (%) Complete
1
70
0 10 20 30 40 50 60 70 80 90 100
0 1 2 3 4 5


Input #11.2 - Completeness of Certified Info, Variable 'Large'
Frequency of Selection
Percent (%) Complete
50
70
80
100
0 10 20 30 40 50 60 70 80 90 100
0 1 2

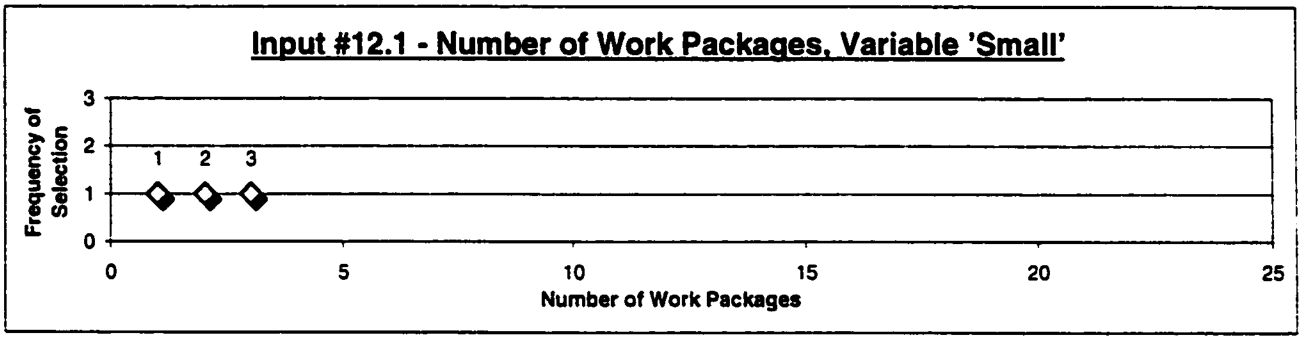

Input #11.3 - Errors in Certified Info, Variable 'Small'
Frequency of Selection
Number of Errors
10
20
0 5 10 15 20 25
0 1 2


Input #11.3 - Errors in Certified Info, Variable 'Average'
Frequency of Selection
Number of Errors
1
3
5
10
20
0 5 10 15 20 25
0 1 2 3

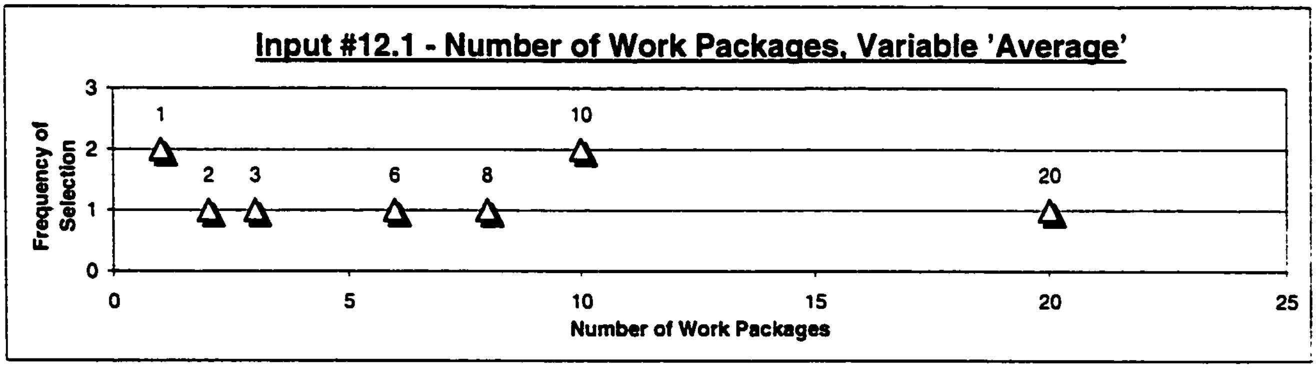

Input #11.3 - Errors in Certified Info, Variable 'Large'
Frequency of Selection
0
1
2
0
10
20
30
40
50
60
70
80
90
100
Number of Errors



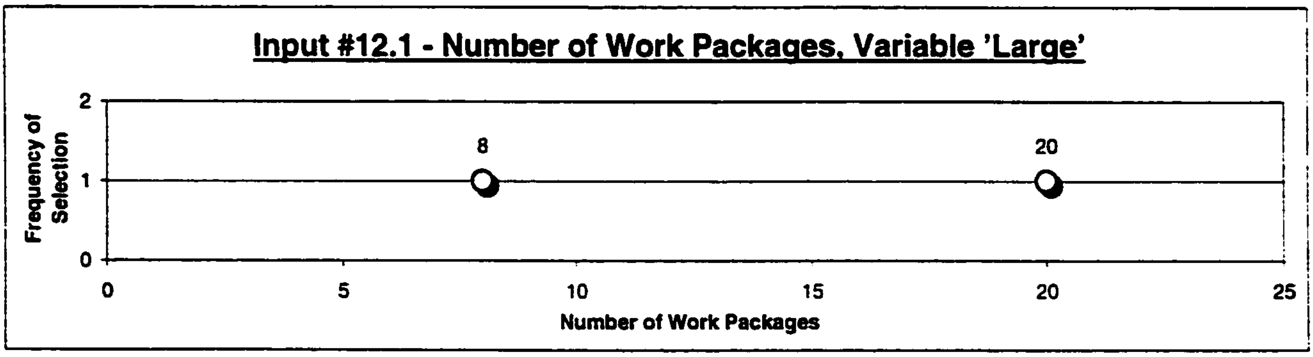
Input #12 - Complexity of Construction Tender
Low
Average
High
Frequency of Selection
Rating
0
1
2
3
4
5
6
7
1
2
3
4
5
6
7
8
9
10



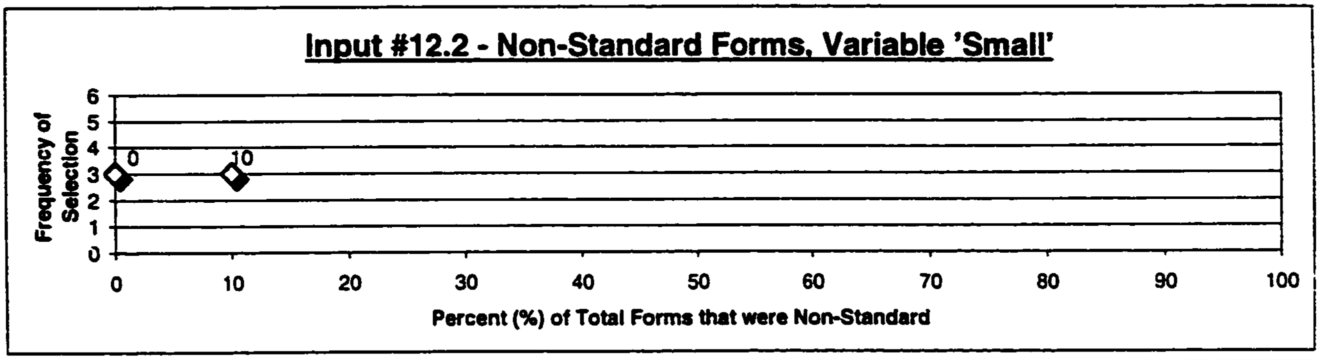
Input #12.1 - Number of Work Packages, Variable 'Small'
Frequency of Selection
Number of Work Packages
1
2
3
0
5
10
15
20
25

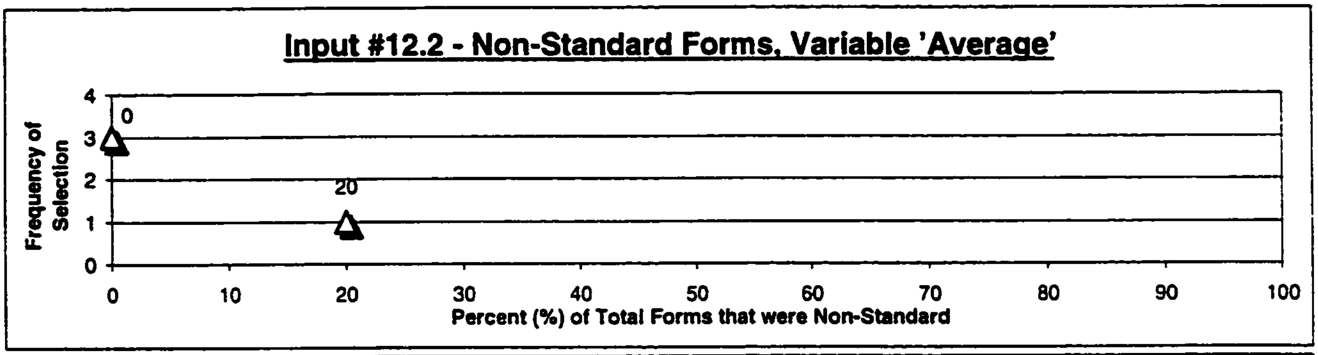


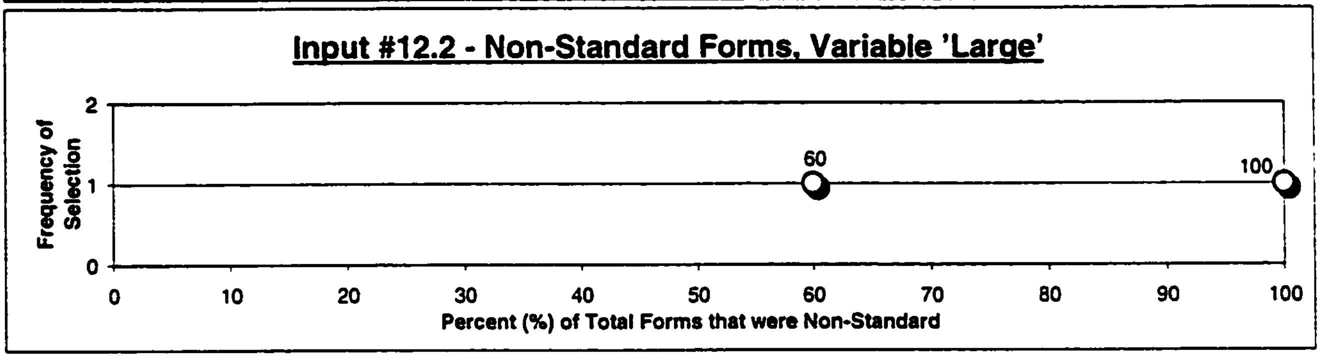
Input #12.1 - Number of Work Packages, Variable 'Average'
Frequency of Selection
Number of Work Packages
1
10
2
3
6
8
20
0
5
10
15
20
25

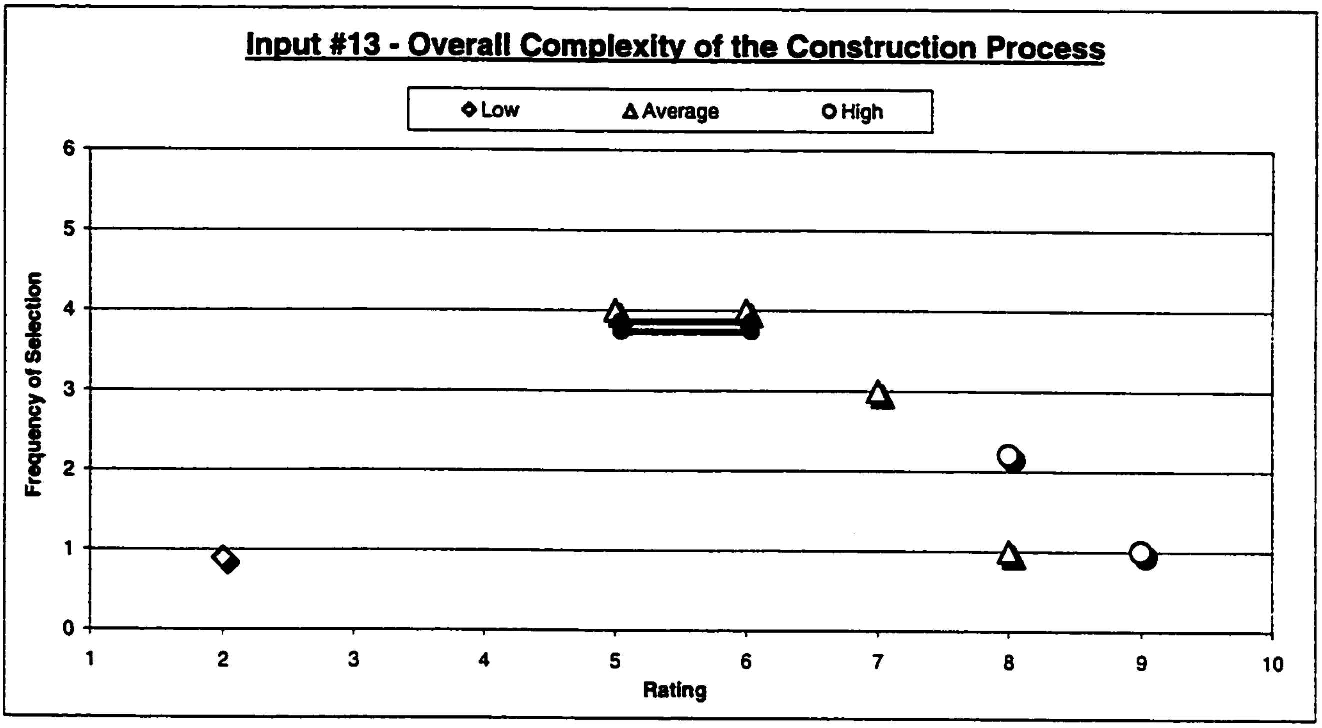


Input #12.1 - Number of Work Packages, Variable 'Large'
Frequency of Selection
Number of Work Packages
8
20
0
5
10
15
20
25

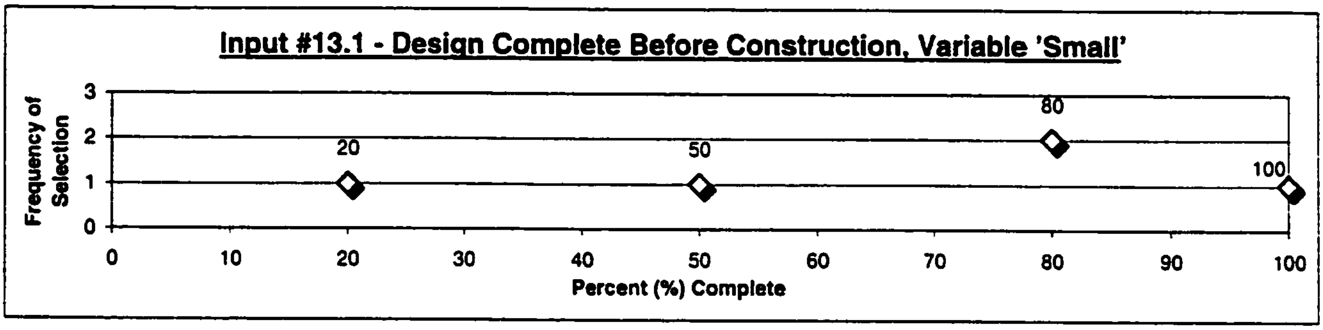

Input #12.2 - Non-Standard Forms, Variable 'Small'
Frequency of Selection
6
5
4
3
2
1
0
0
10
0
10
20
30
40
50
60
70
80
90
100
Percent (%) of Total Forms that were Non-Standard


Input #12.2 - Non-Standard Forms, Variable 'Average'
Frequency of Selection
4
3
2
1
0
0
20
0
10
20
30
40
50
60
70
80
90
100
Percent (%) of Total Forms that were Non-Standard

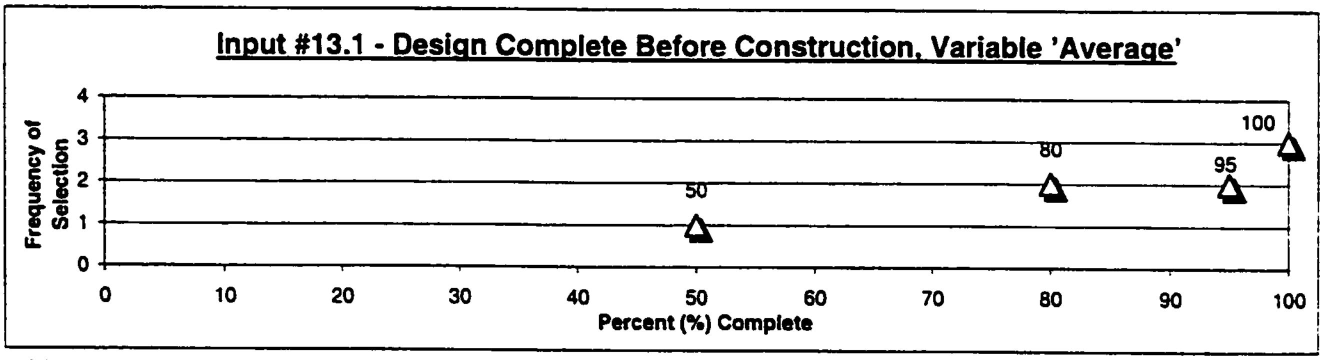

Input #12.2 - Non-Standard Forms, Variable 'Large'
Frequency of Selection
2
1
0
60
100
0
10
20
30
40
50
60
70
80
90
100
Percent (%) of Total Forms that were Non-Standard



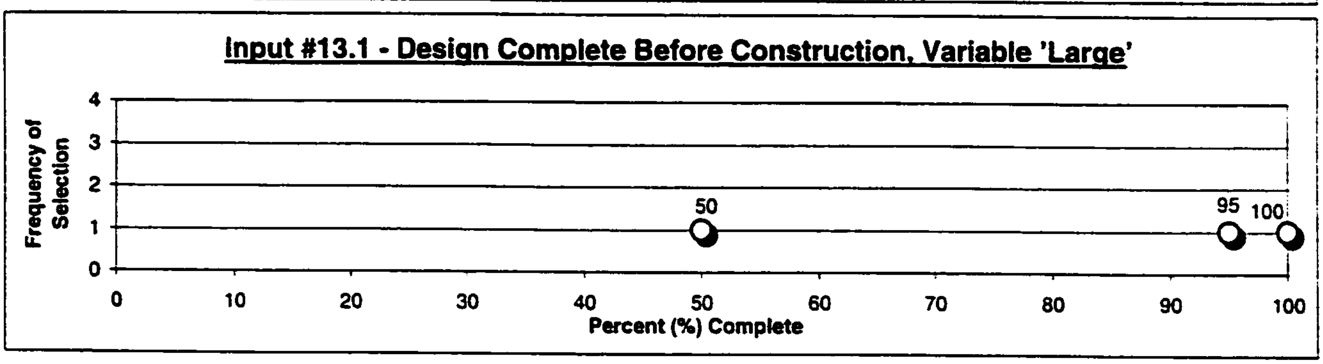
Input #13 - Overall Complexity of the Construction Process
Low
Average
High
Frequency of Selection
Rating



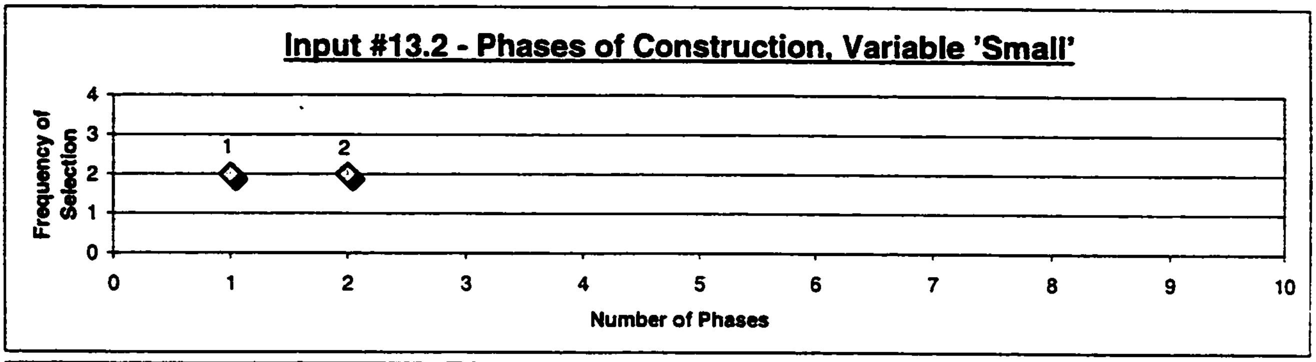
Input #13.1 - Design Complete Before Construction, Variable 'Small'
Frequency of Selection
20
50
80
100
Percent (%) Complete

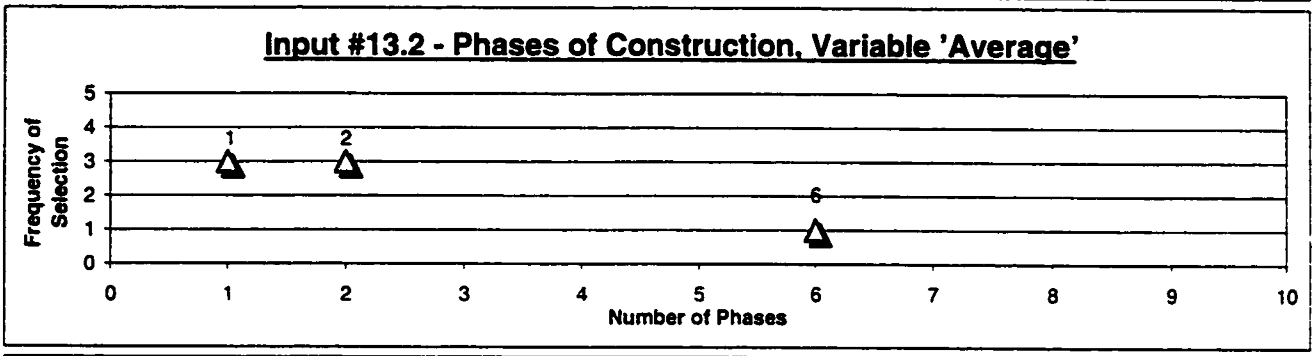


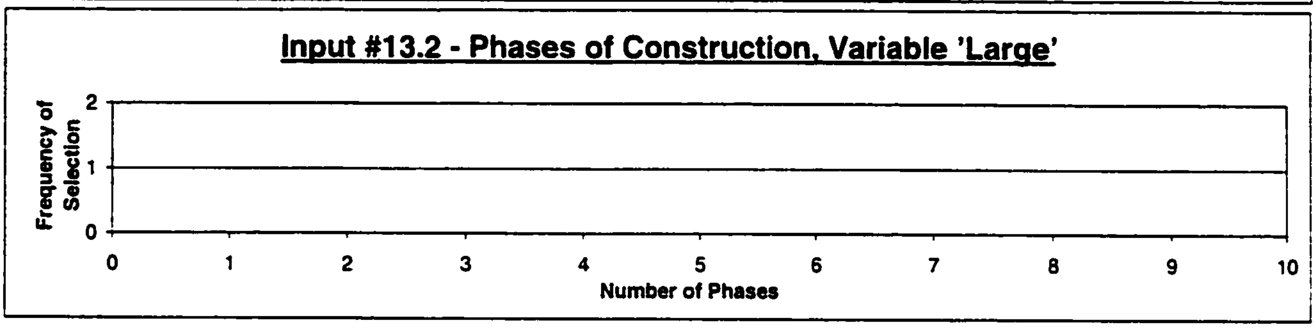
Input #13.1 - Design Complete Before Construction, Variable 'Average'
Frequency of Selection
50
80
95
100
Percent (%) Complete

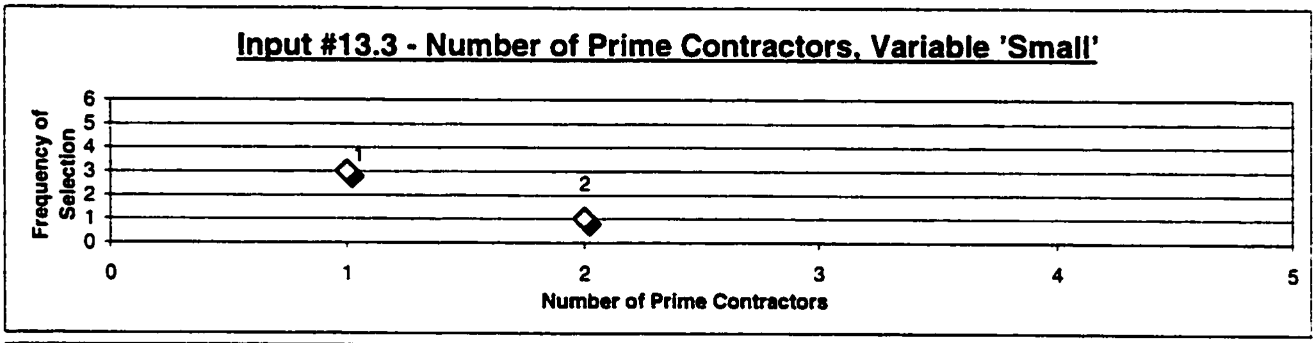


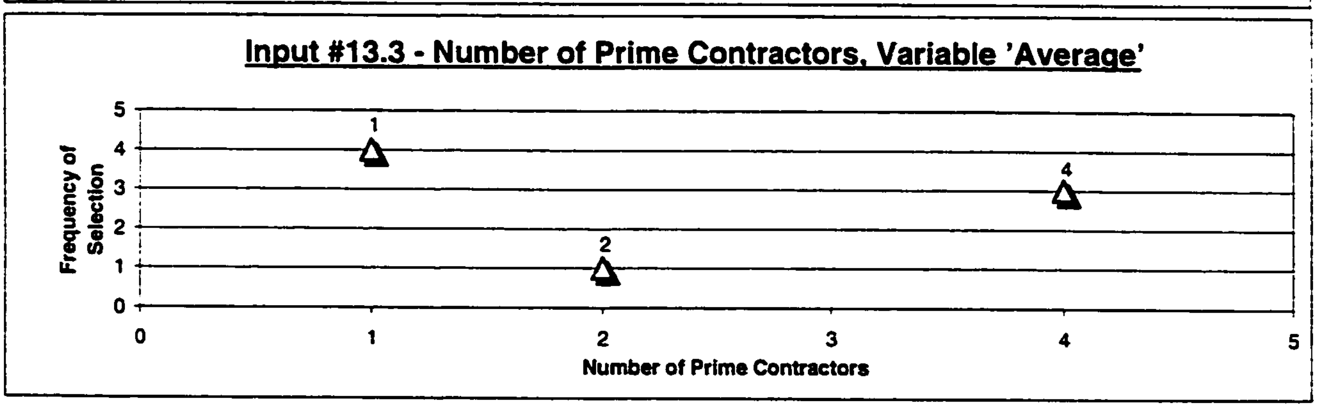
Input #13.1 - Design Complete Before Construction, Variable 'Large'
Frequency of Selection
50
95
100
Percent (%) Complete


Input #13.2 - Phases of Construction, Variable 'Small'
Frequency of Selection
Number of Phases

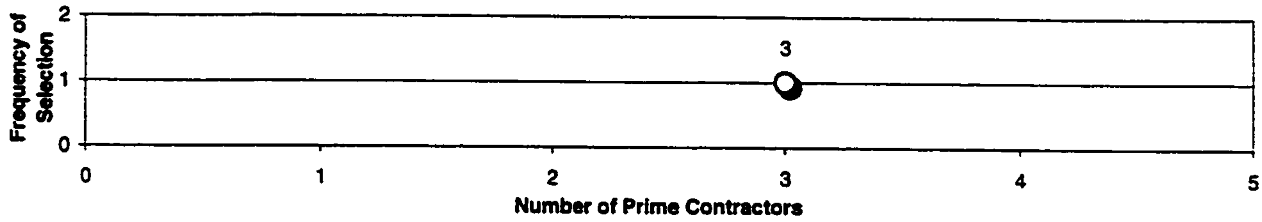

Input #13.2 - Phases of Construction, Variable 'Average'
Frequency of Selection
Number of Phases


Input #13.2 - Phases of Construction, Variable 'Large'
Frequency of Selection
Number of Phases

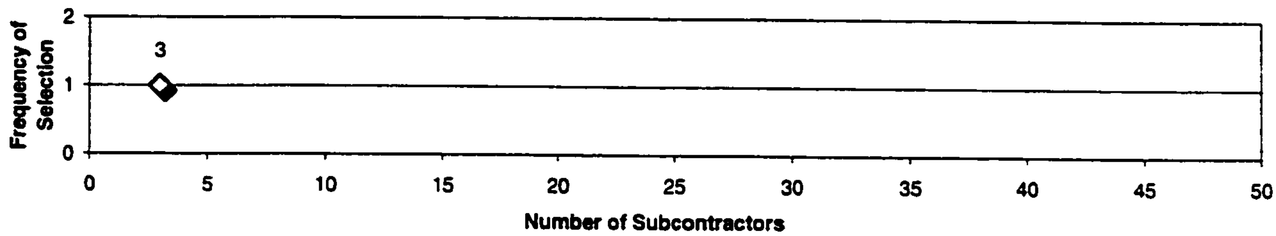

Input #13.3 - Number of Prime Contractors, Variable 'Small'
Frequency of Selection
Number of Prime Contractors


Input #13.3 - Number of Prime Contractors, Variable 'Average'
Frequency of Selection
Number of Prime Contractors

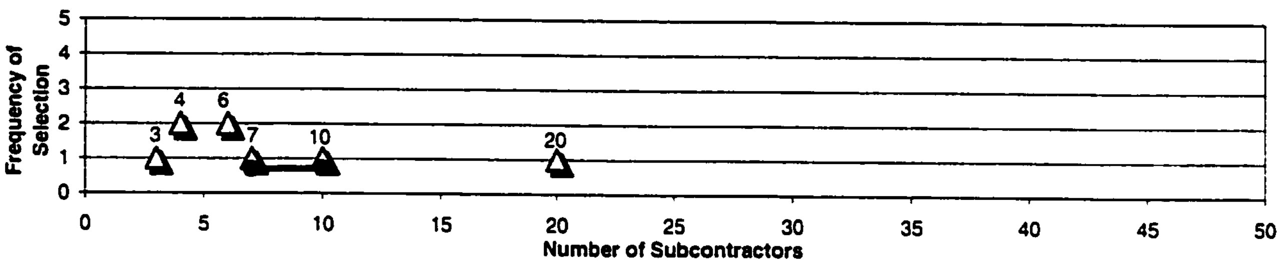
Input #13.3 - Number of Prime Contractors, Variable 'Large'



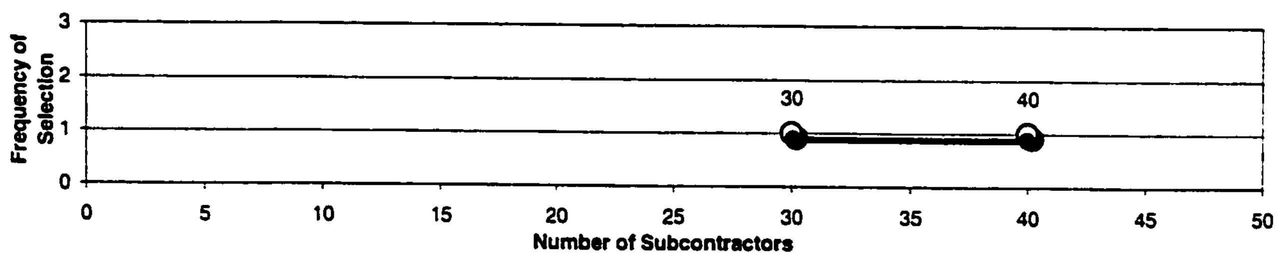
Input #13.4 - Number of Subcontractors, Variable 'Small'



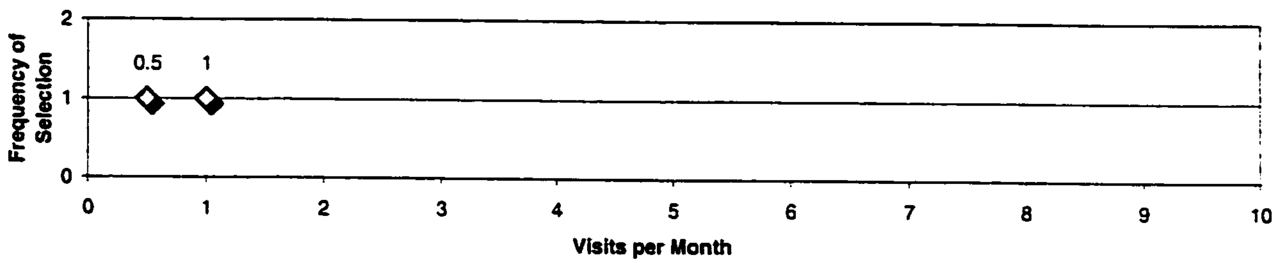
Input #13.4 - Number of Subcontractors, Variable 'Average'

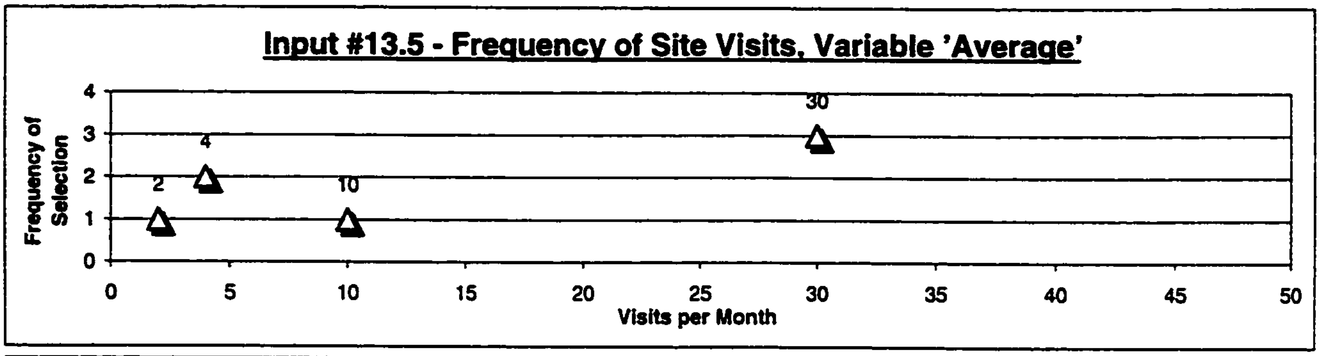


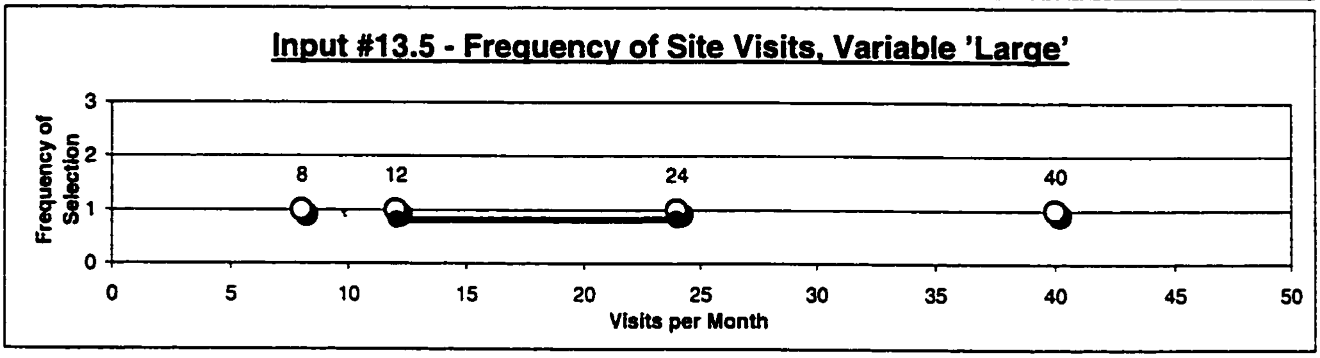
Input #13.4 - Number of Subcontractors, Variable 'Large'

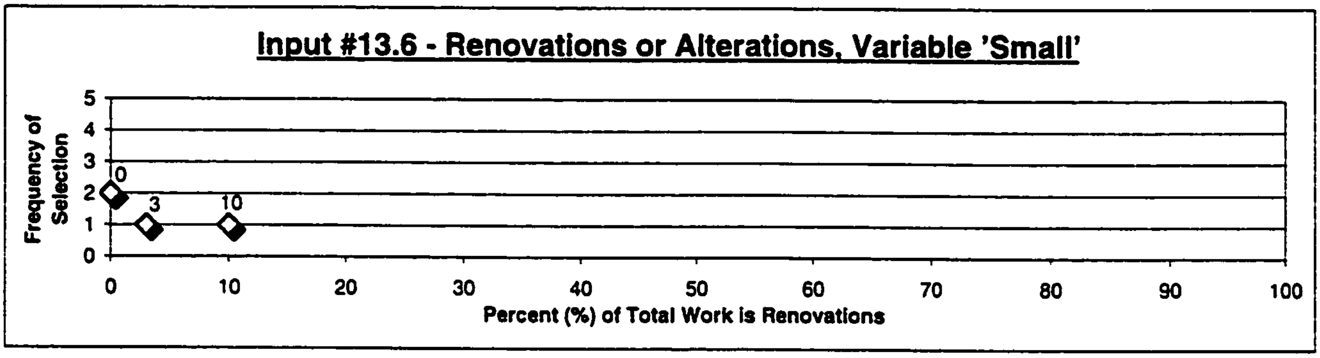


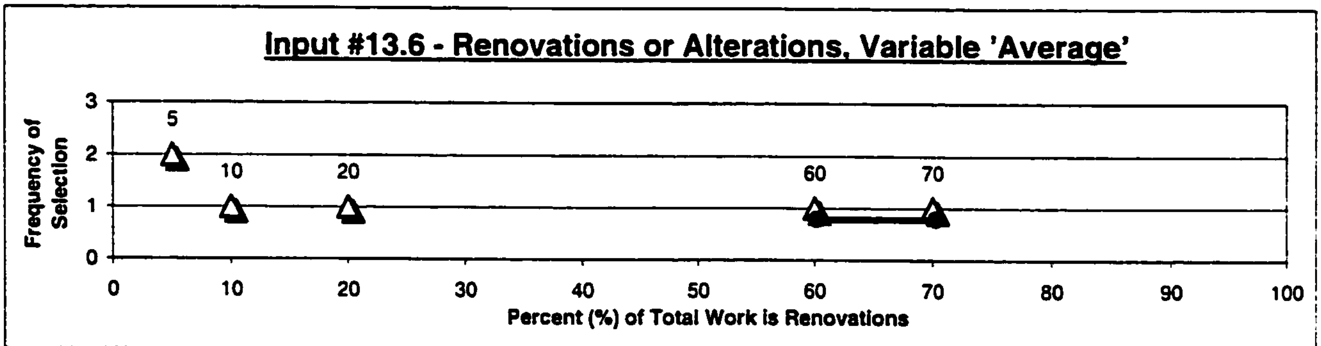
Input #13.5 - Frequency of Site Visits, Variable 'Small'

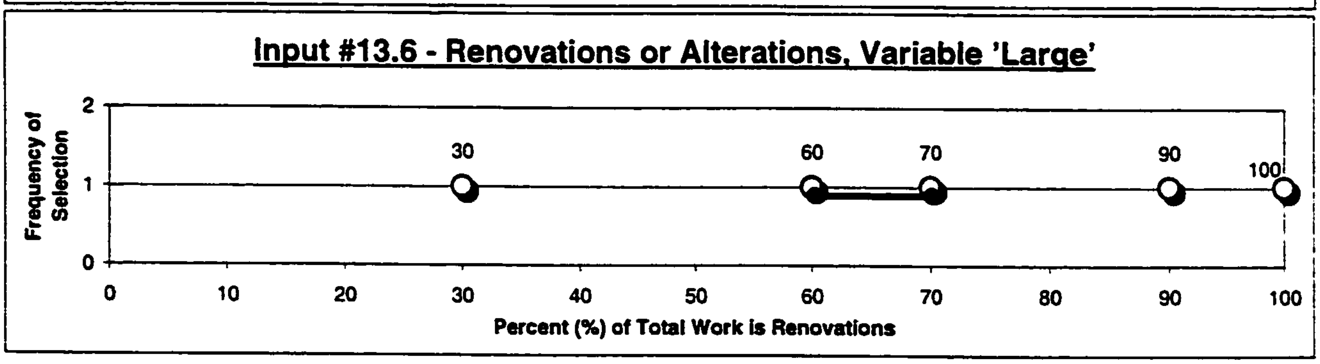


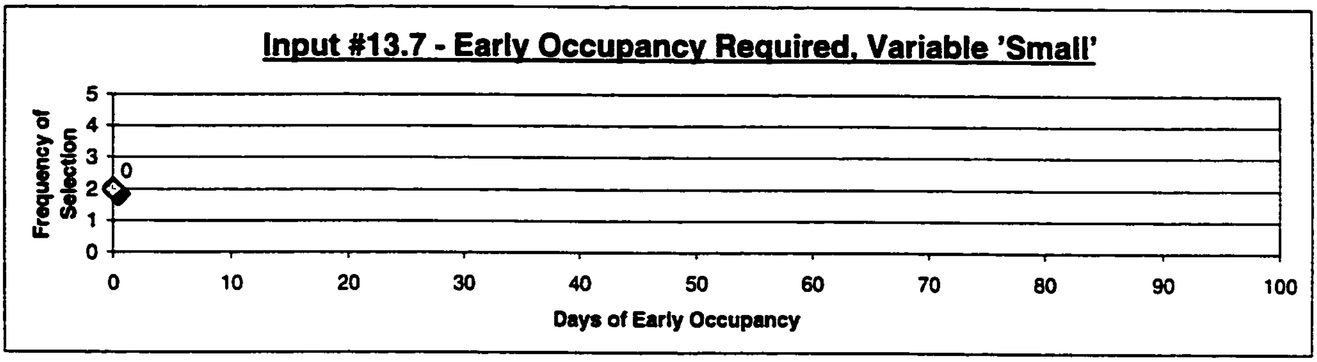

Input #13.5 - Frequency of Site Visits, Variable 'Average'
Frequency of Selection
Visits per Month
2
4
10
30

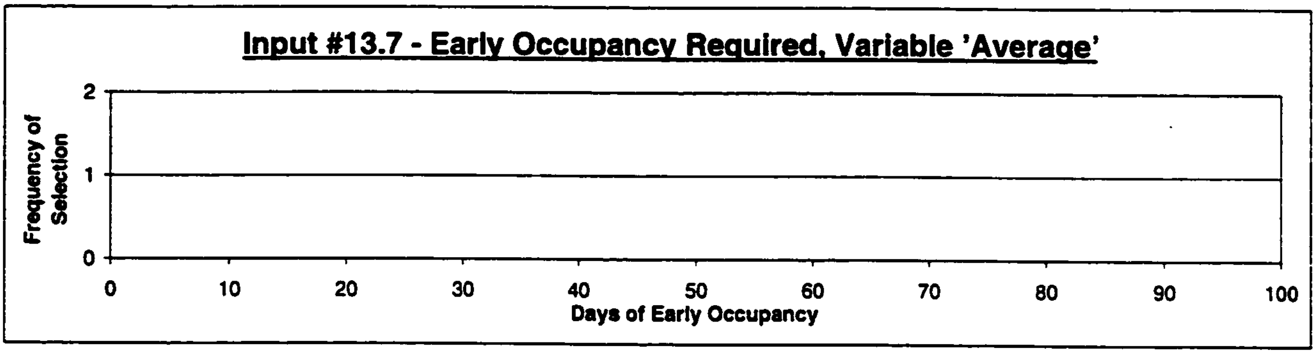

Input #13.5 - Frequency of Site Visits, Variable 'Large'
Frequency of Selection
Visits per Month
8
12
24
40

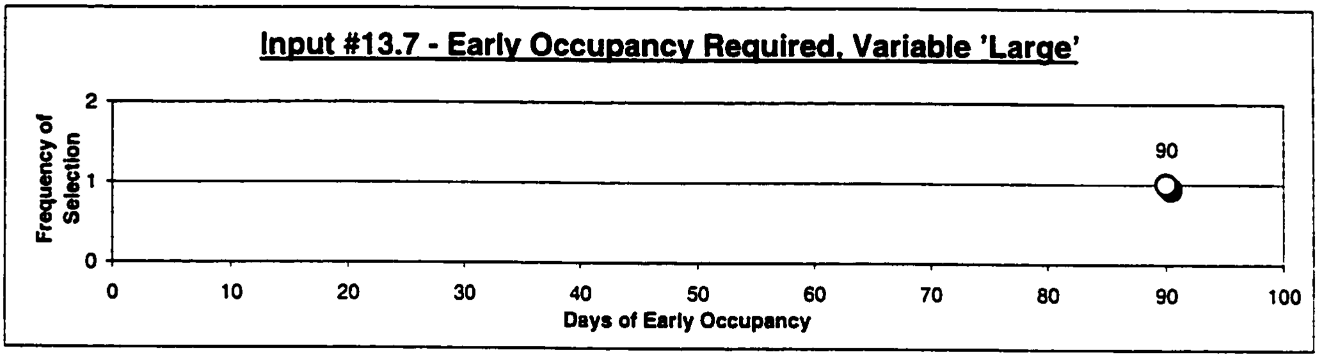

Input #13.6 - Renovations or Alterations, Variable 'Small'
Frequency of Selection
Percent (%) of Total Work is Renovations
0
3
10

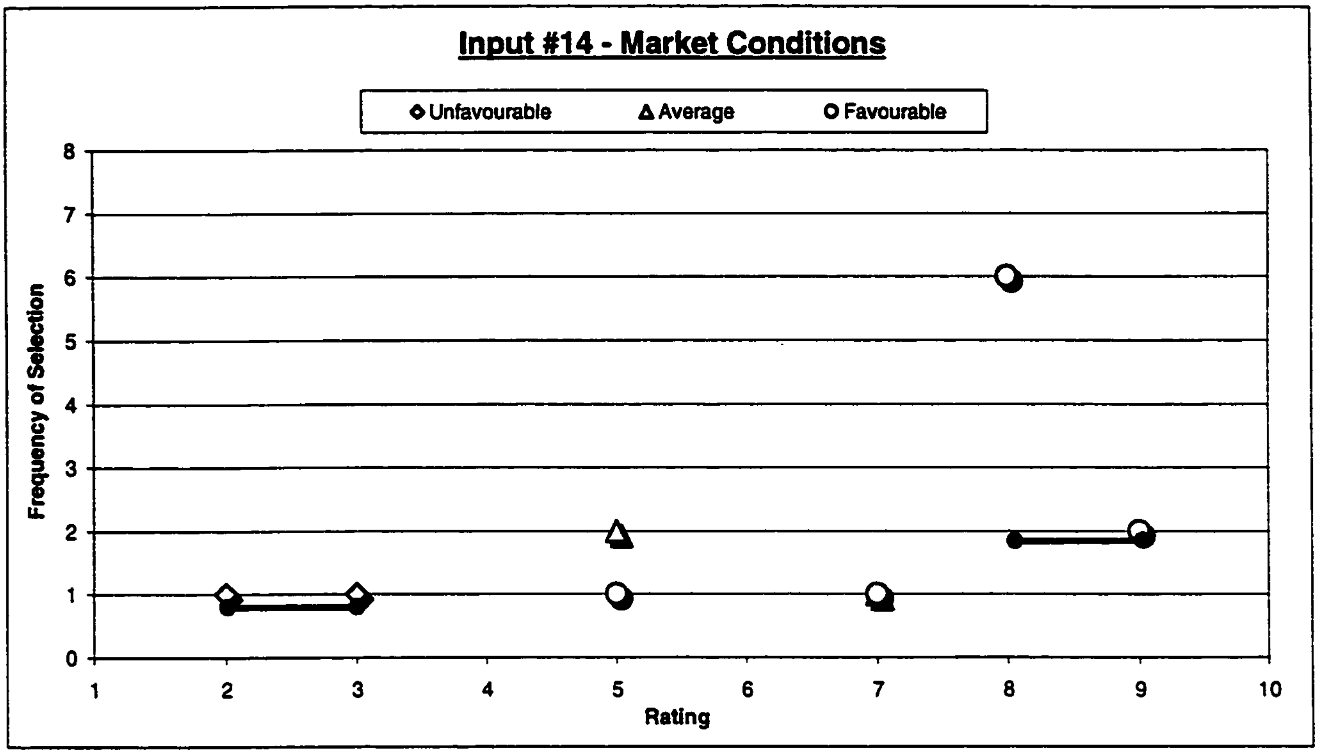

Input #13.6 - Renovations or Alterations, Variable 'Average'
Frequency of Selection
Percent (%) of Total Work is Renovations
5
10
20
60
70

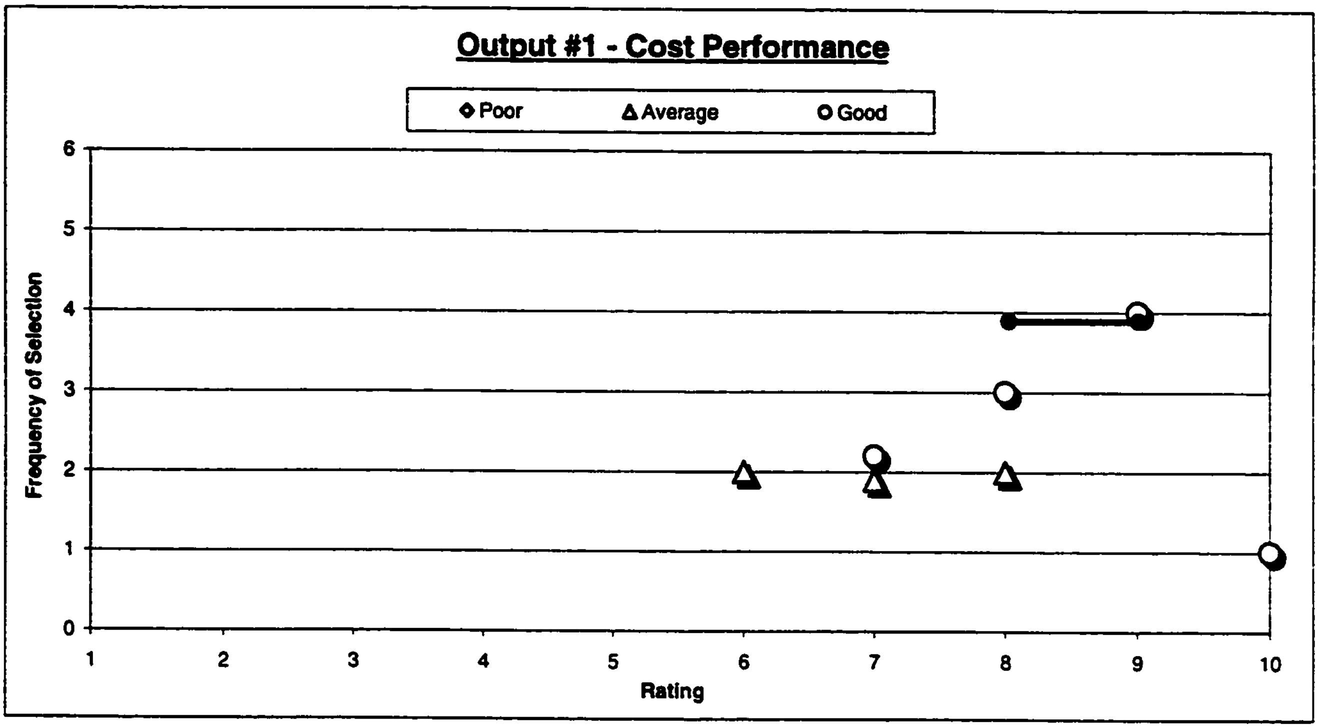

Input #13.6 - Renovations or Alterations, Variable 'Large'
Frequency of Selection
Percent (%) of Total Work is Renovations
30
60
70
90
100


Input #13.7 - Early Occupancy Required, Variable 'Small'
Frequency of Selection
Days of Early Occupancy
0

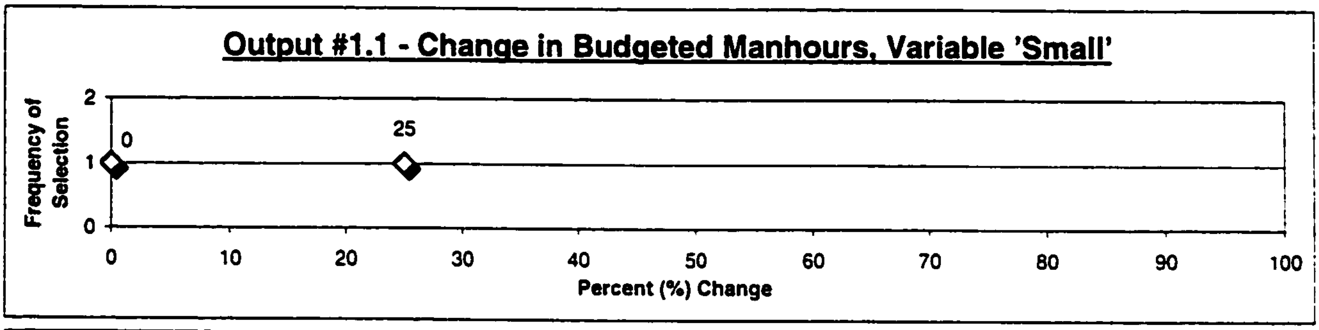

Input #13.7 - Early Occupancy Required, Variable 'Average'
Frequency of Selection
Days of Early Occupancy


Input #13.7 - Early Occupancy Required, Variable 'Large'
Frequency of Selection
Days of Early Occupancy
90

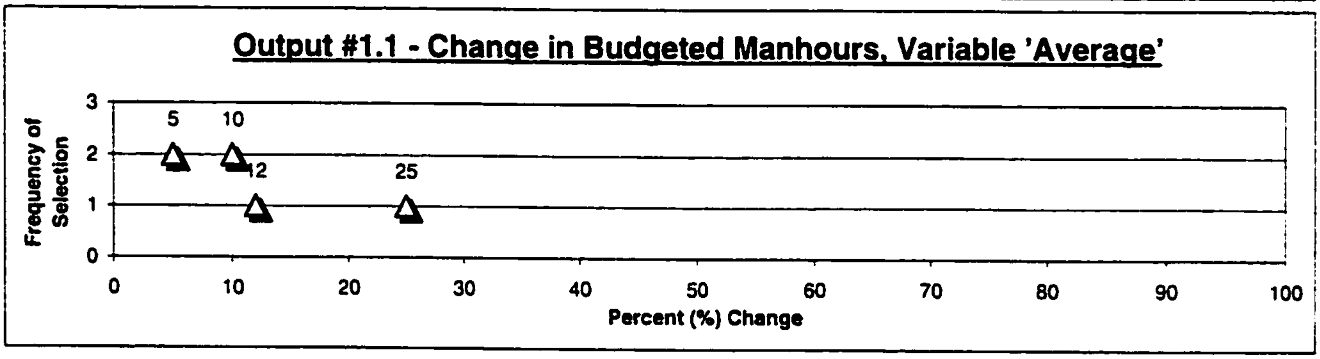

Input #14 - Market Conditions
Unfavourable
Average
Favourable
Frequency of Selection
Rating
0
1
2
3
4
5
6
7
8
1
2
3
4
5
6
7
8
9
10



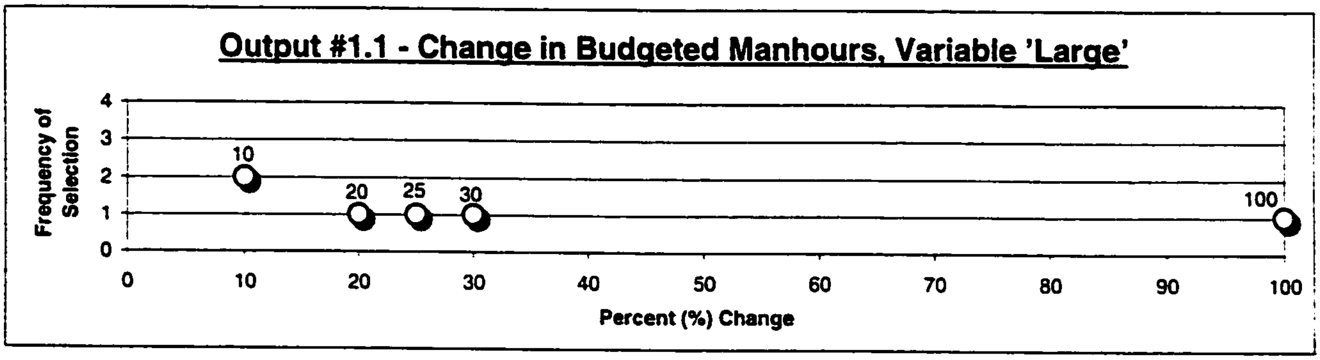
Output #1 - Cost Performance
Poor
Average
Good
Frequency of Selection
Rating



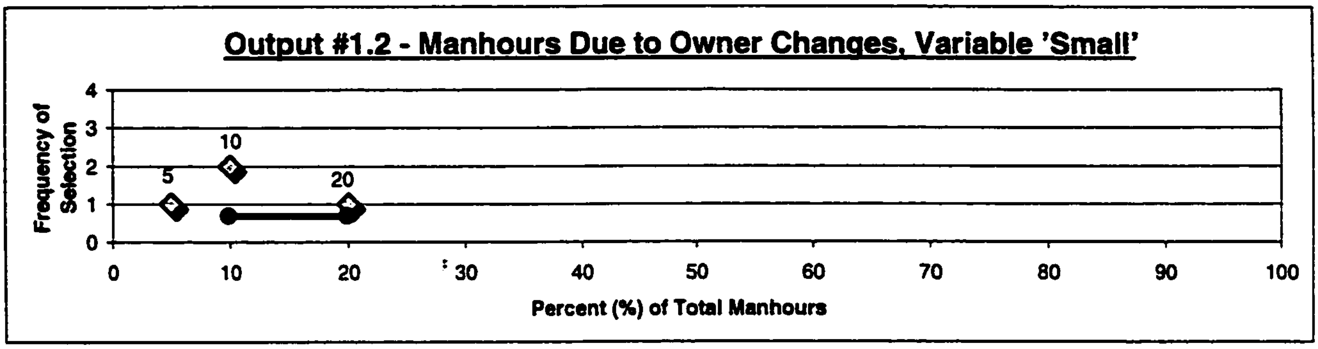
Output #1.1 - Change in Budgeted Manhours, Variable 'Small'
Frequency of Selection
Percent (%) Change

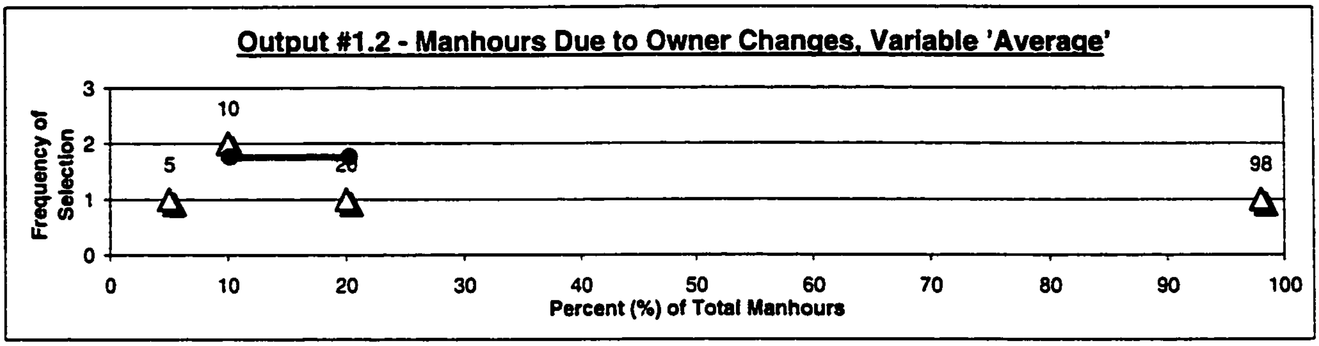


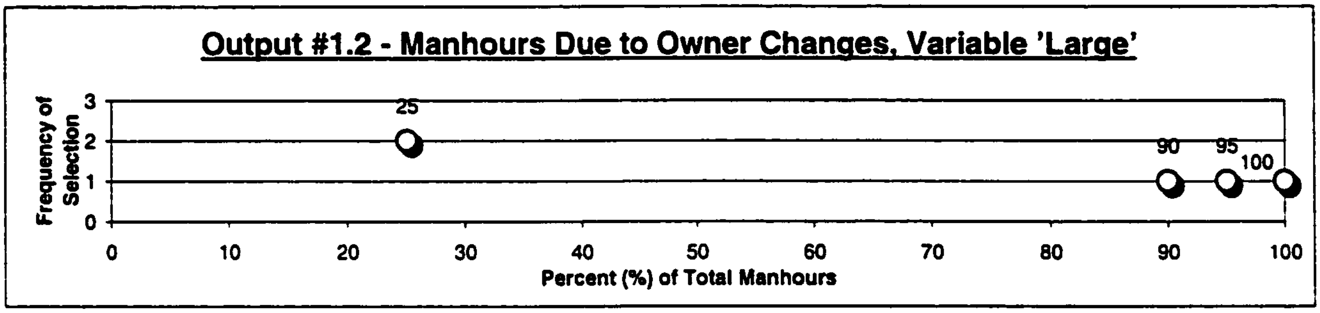
Output #1.1 - Change in Budgeted Manhours, Variable 'Average'
Frequency of Selection
Percent (%) Change

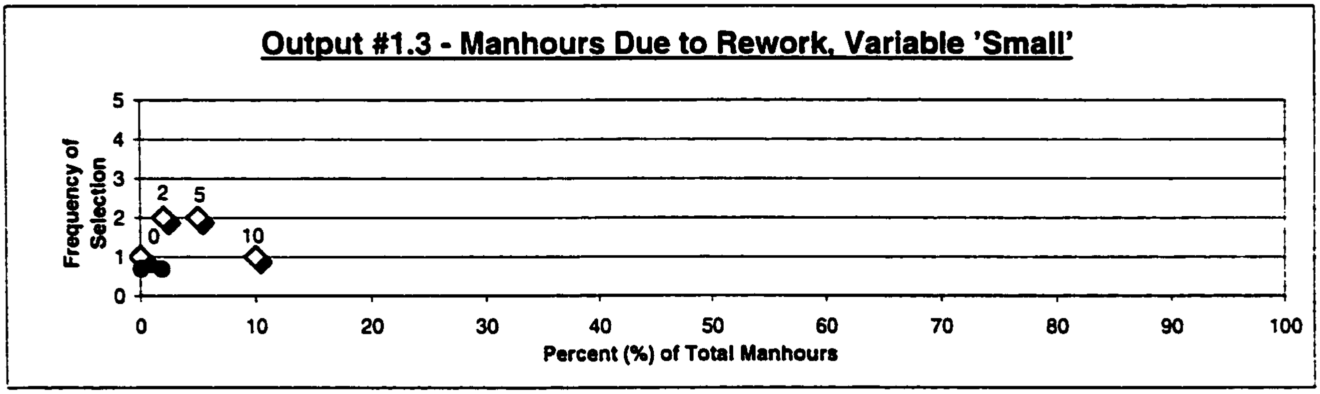


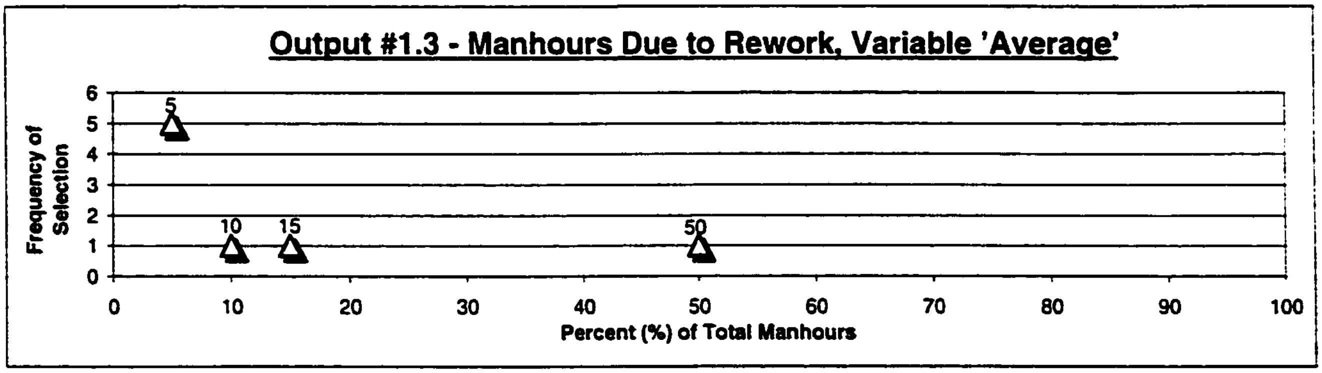
Output #1.1 - Change in Budgeted Manhours, Variable 'Large'
Frequency of Selection
Percent (%) Change

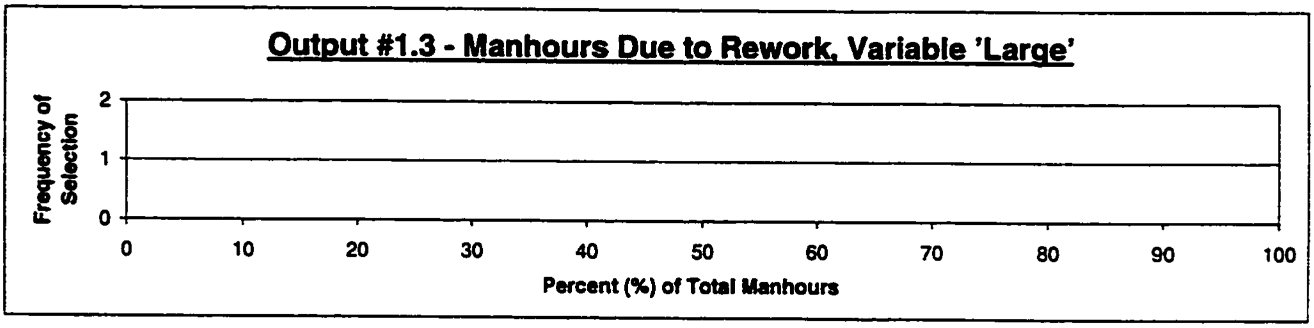

Output #1.2 - Manhours Due to Owner Changes, Variable 'Small'
Frequency of Selection
Percent (%) of Total Manhours
5
10
20

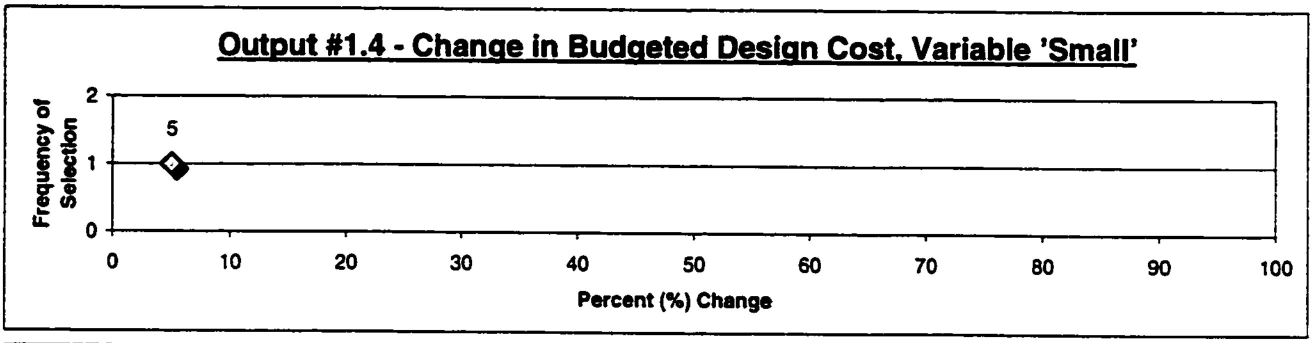

Output #1.2 - Manhours Due to Owner Changes, Variable 'Average'
Frequency of Selection
Percent (%) of Total Manhours
5
10
20
98

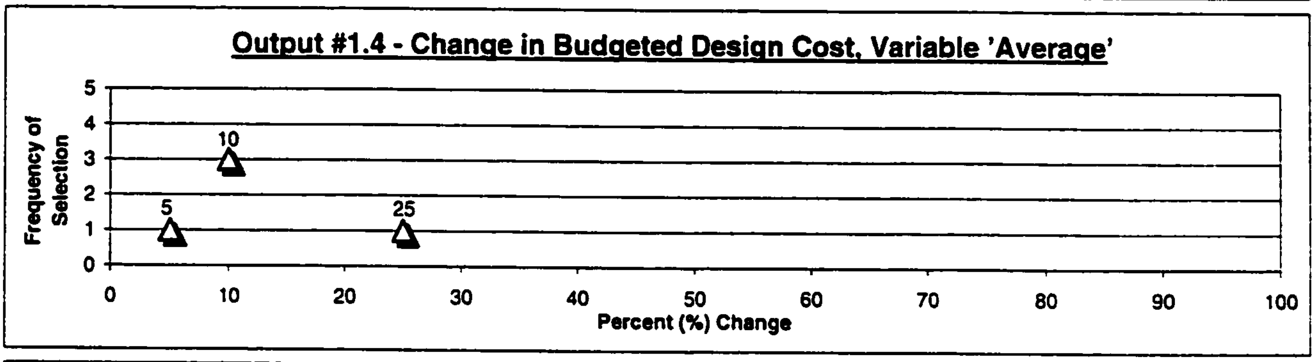

Output #1.2 - Manhours Due to Owner Changes, Variable 'Large'
Frequency of Selection
Percent (%) of Total Manhours
25
90
95
100

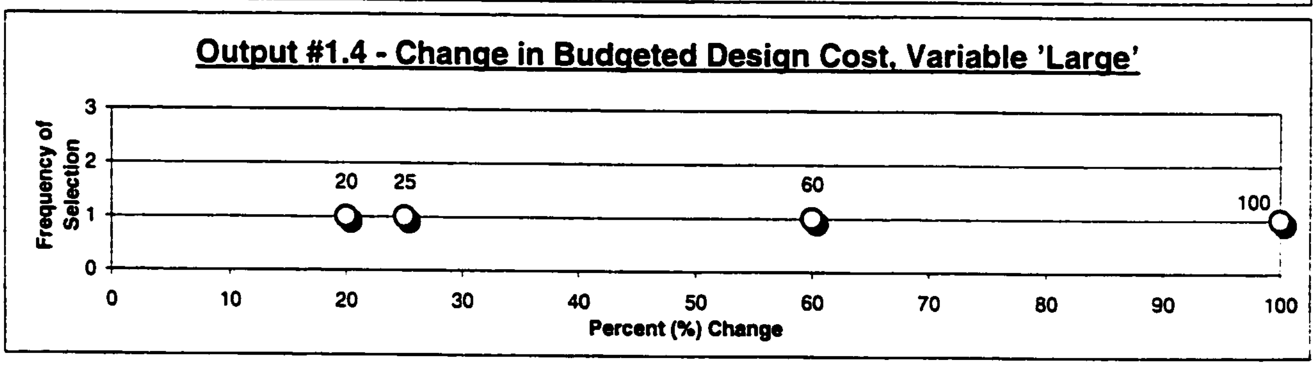

Output #1.3 - Manhours Due to Rework, Variable 'Small'
Frequency of Selection
Percent (%) of Total Manhours
0
2
5
10

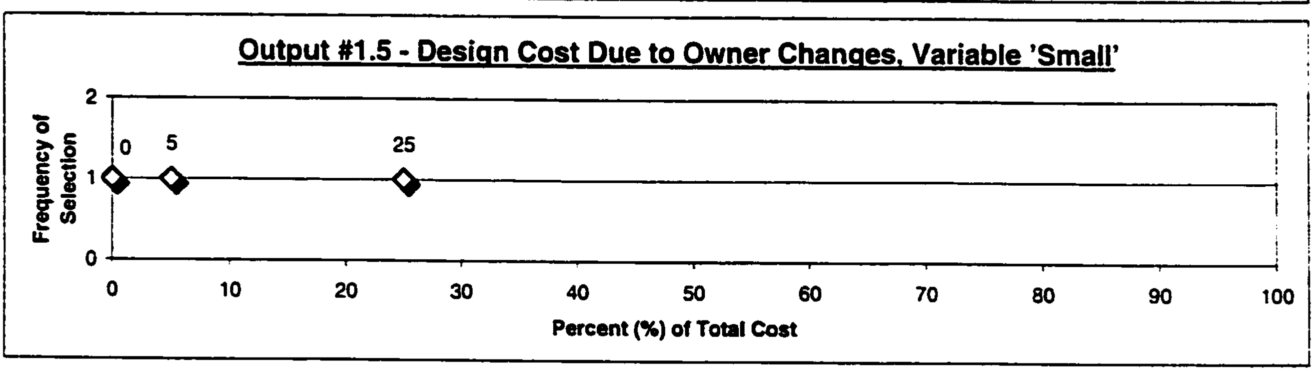

Output #1.3 - Manhours Due to Rework, Variable 'Average'
Frequency of Selection
Percent (%) of Total Manhours
5
10
15
50


Output #1.3 - Manhours Due to Rework, Variable 'Large'
Frequency of Selection
Percent (%) of Total Manhours

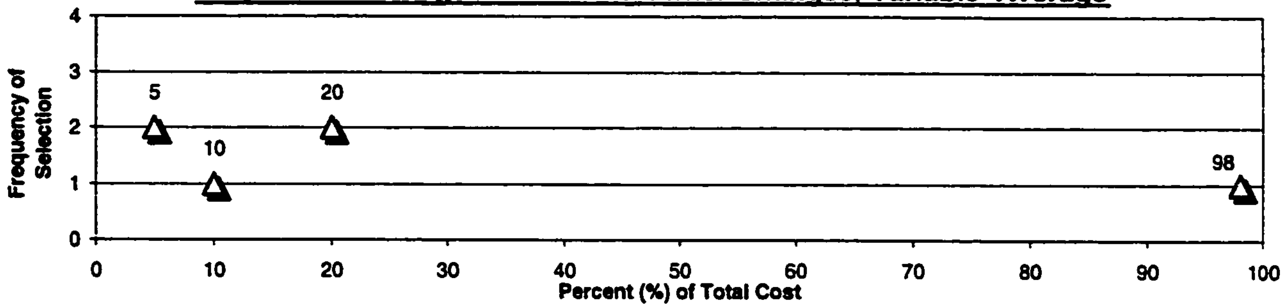

Output #1.4 - Change in Budgeted Design Cost, Variable 'Small'
Frequency of Selection
5
Percent (%) Change


Output #1.4 - Change in Budgeted Design Cost, Variable 'Average'
Frequency of Selection
10
5
25
Percent (%) Change

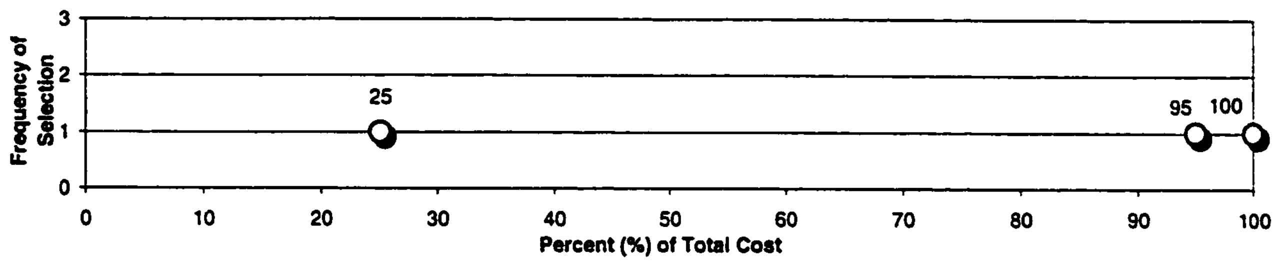

Output #1.4 - Change in Budgeted Design Cost, Variable 'Large'
Frequency of Selection
20
25
60
100
Percent (%) Change


Output #1.5 - Design Cost Due to Owner Changes, Variable 'Small'
Frequency of Selection
0
5
25
Percent (%) of Total Cost

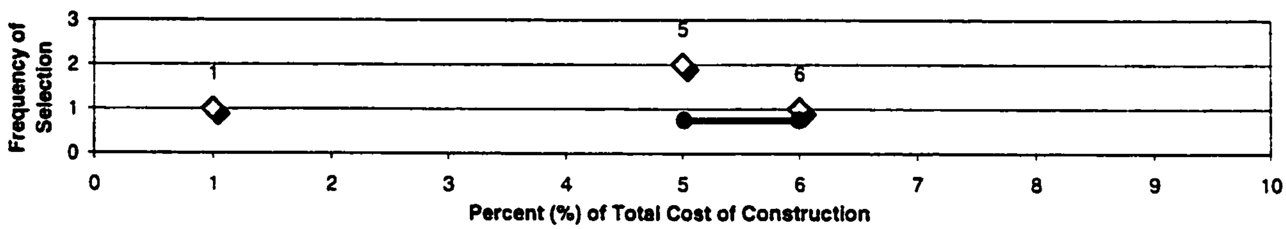
Output #1.5 - Design Cost Due to Owner Changes, Variable 'Average'



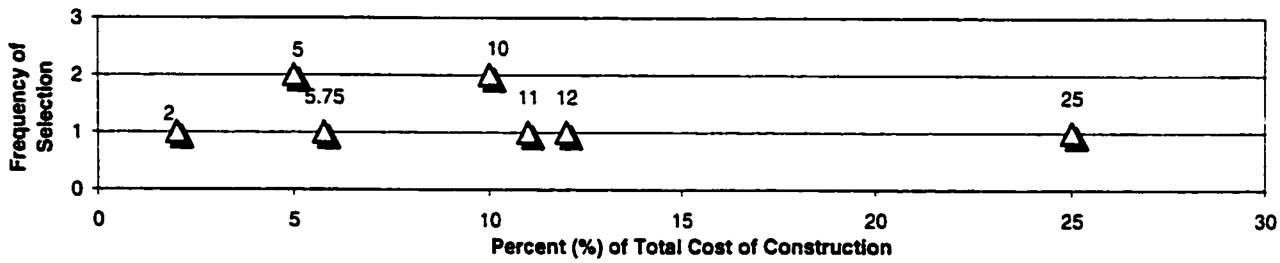
Output #1.5 - Design Cost Due to Owner Changes, Variable 'Large'



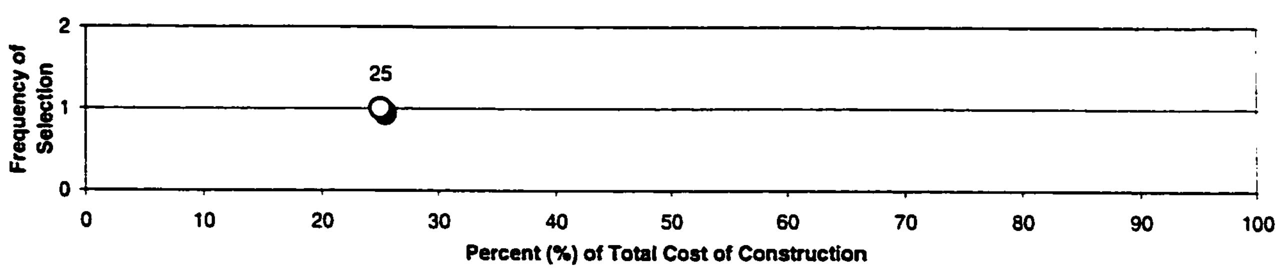
Output #1.6 - Proportion of Design Cost in Total Cost, Variable 'Small'

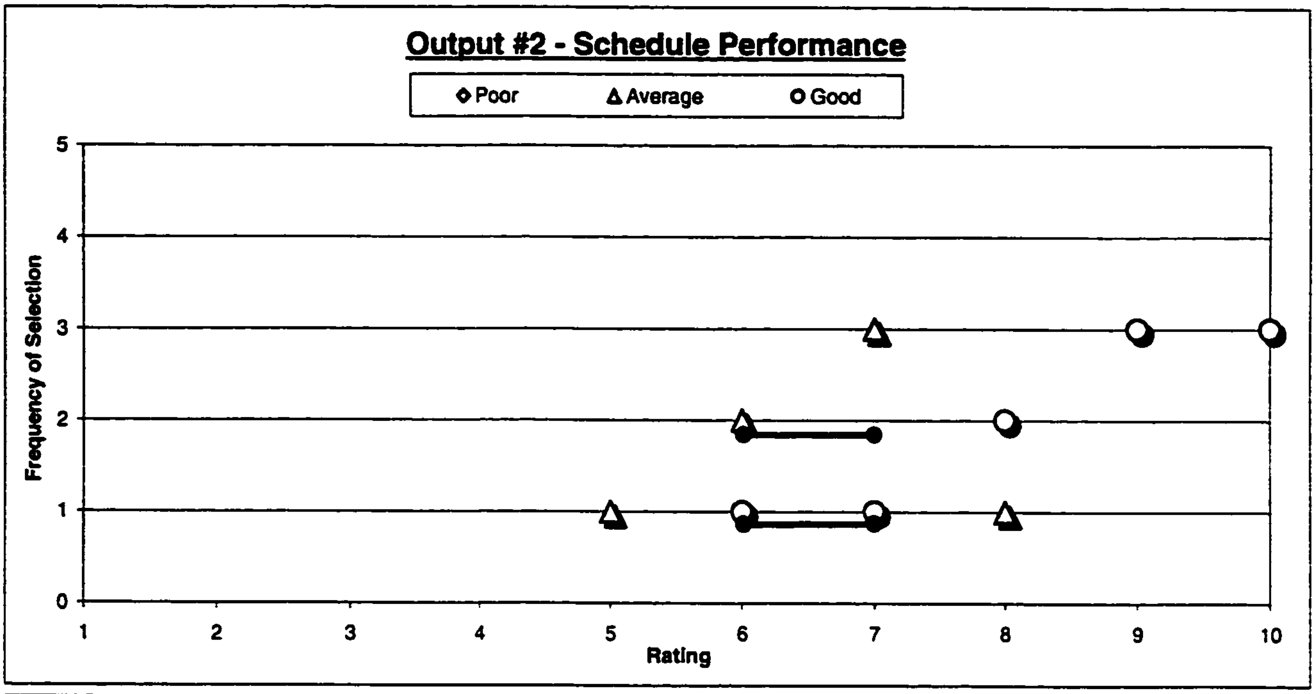


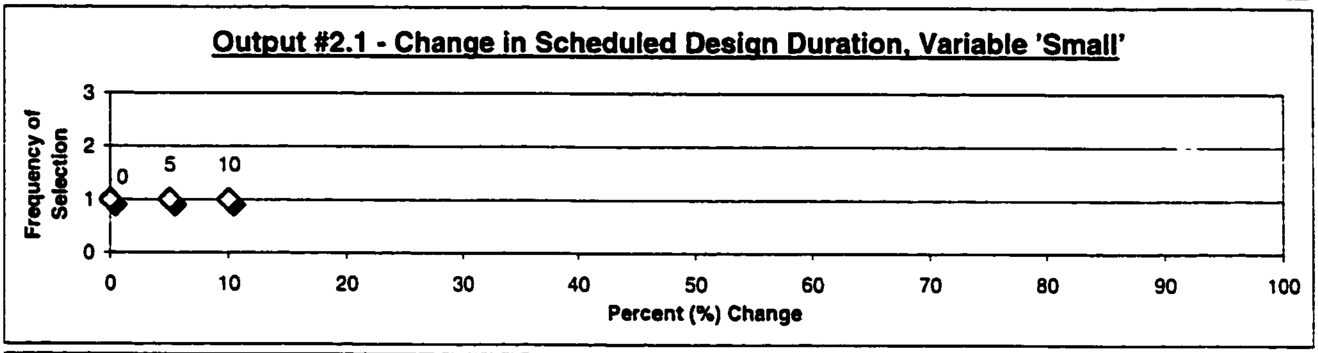
Output #1.6 - Proportion Design Cost in Total Cost, Variable 'Average'

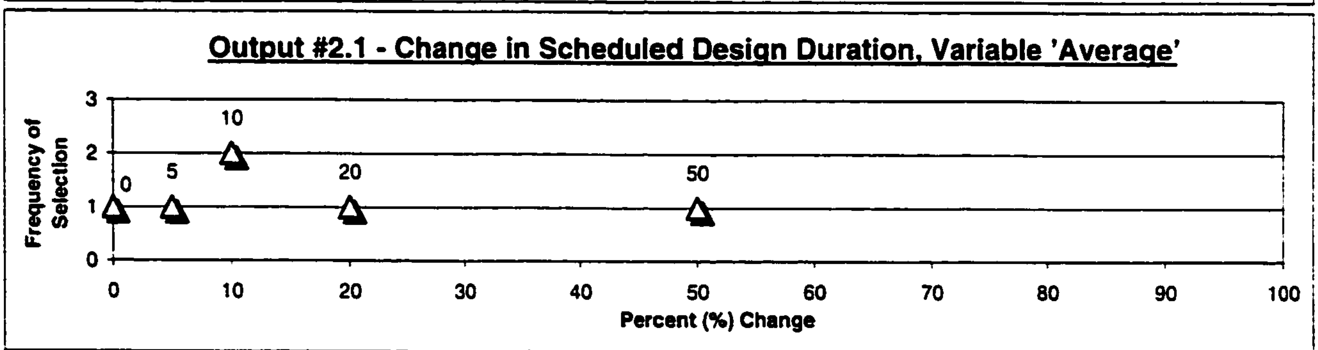


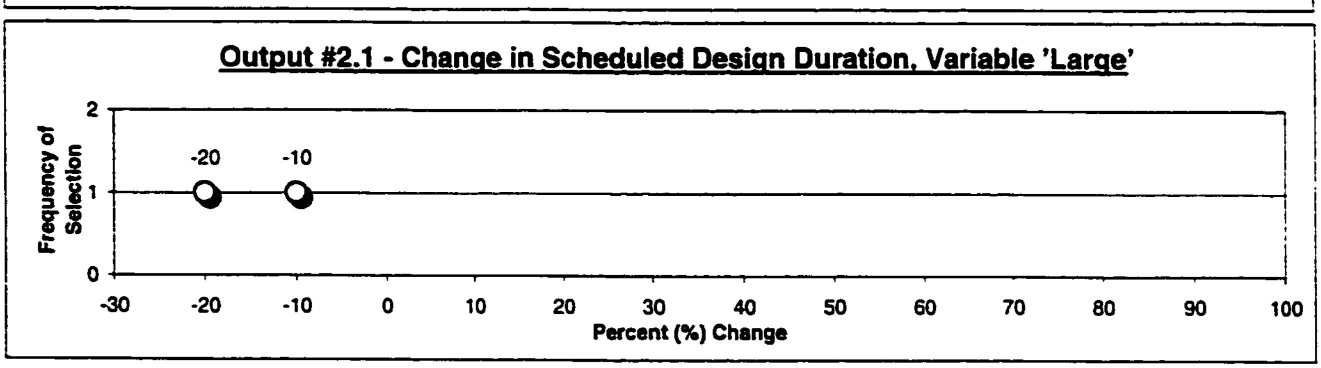
Output #1.6 - Proportion of Design Cost in Total Cost, Variable 'Large'

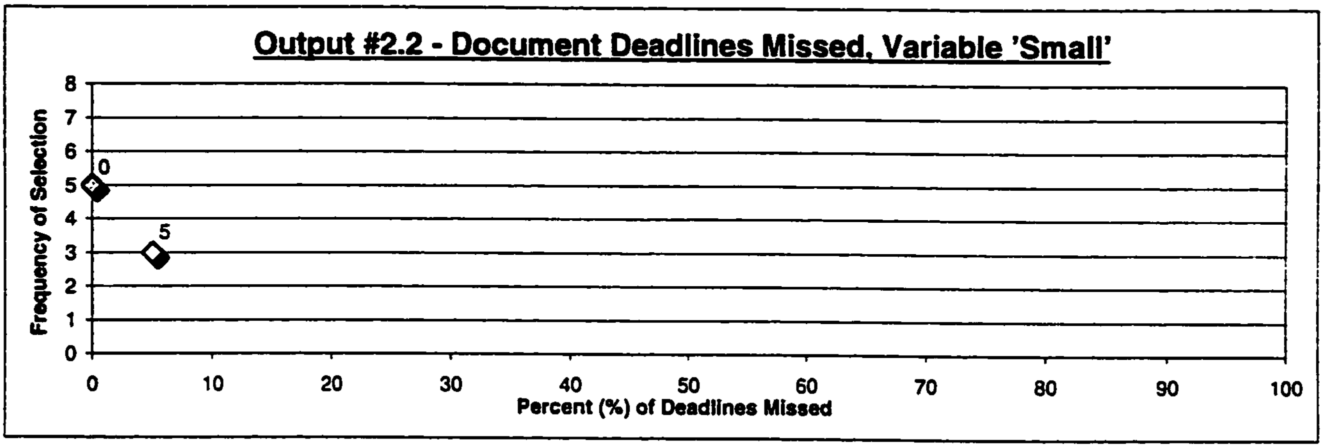


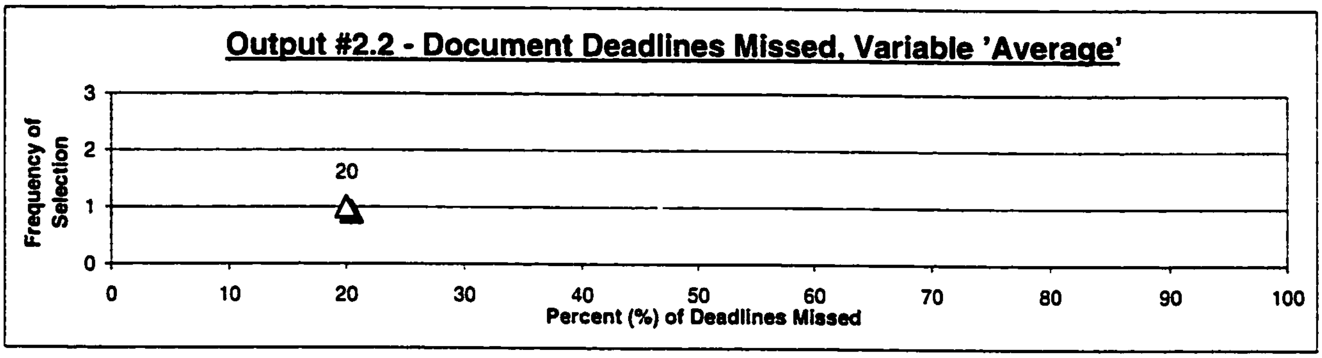

Output #2 - Schedule Performance
Poor
Average
Good
Frequency of Selection
Rating
0
1
2
3
4
5
1
2
3
4
5
6
7
8
9
10

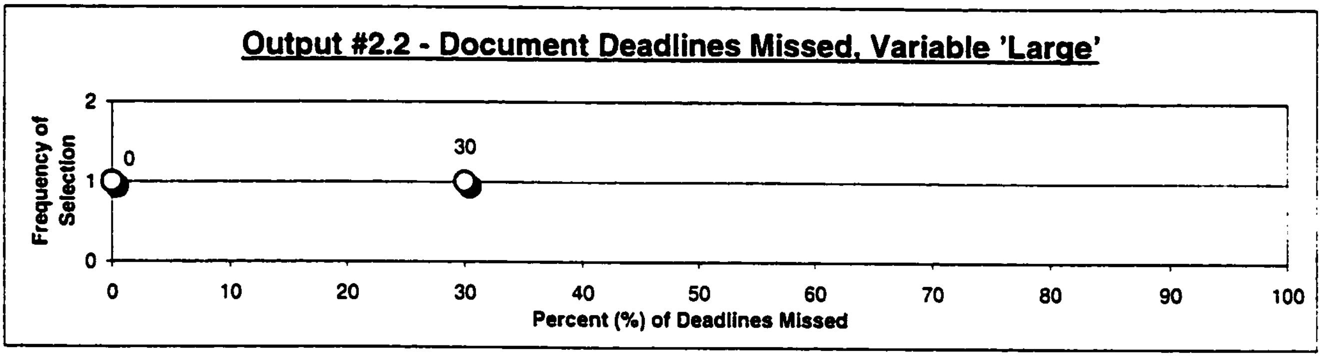

Output #2.1 - Change in Scheduled Design Duration, Variable 'Small'
Frequency of Selection
Percent (%) Change
0
5
10

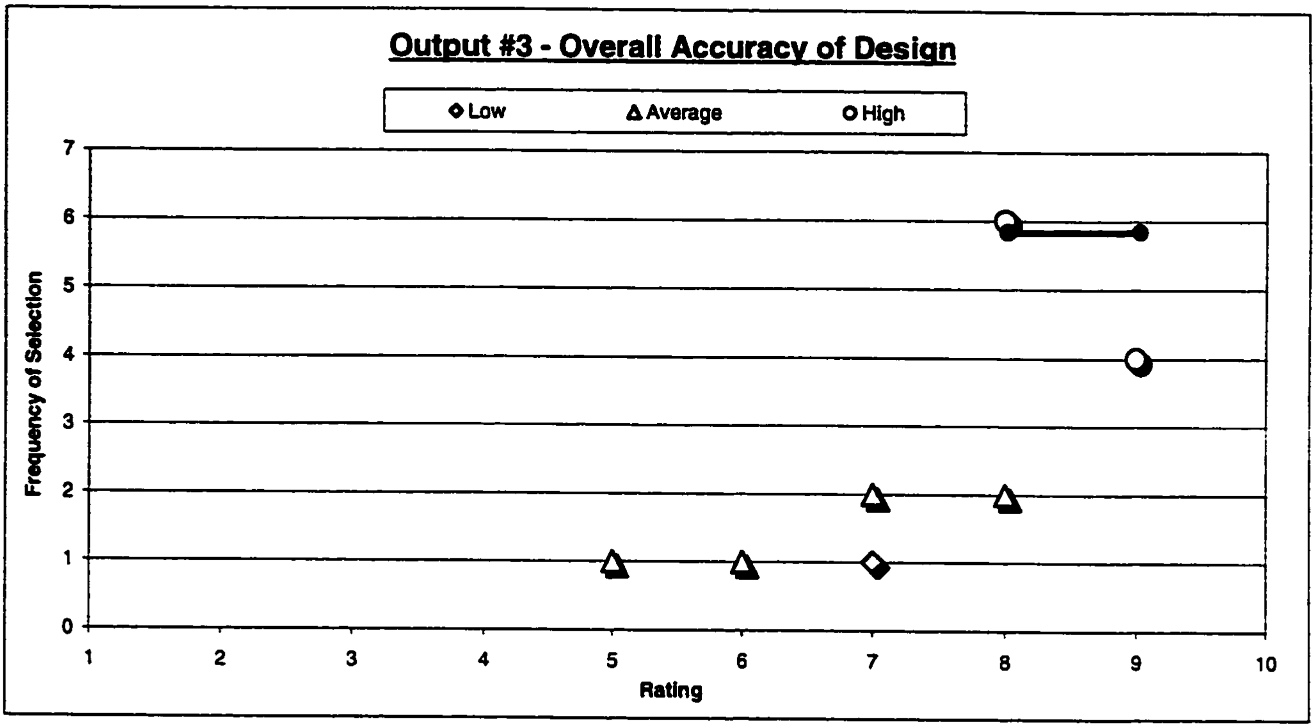

Output #2.1 - Change in Scheduled Design Duration, Variable 'Average'
Frequency of Selection
Percent (%) Change
0
5
10
20
50

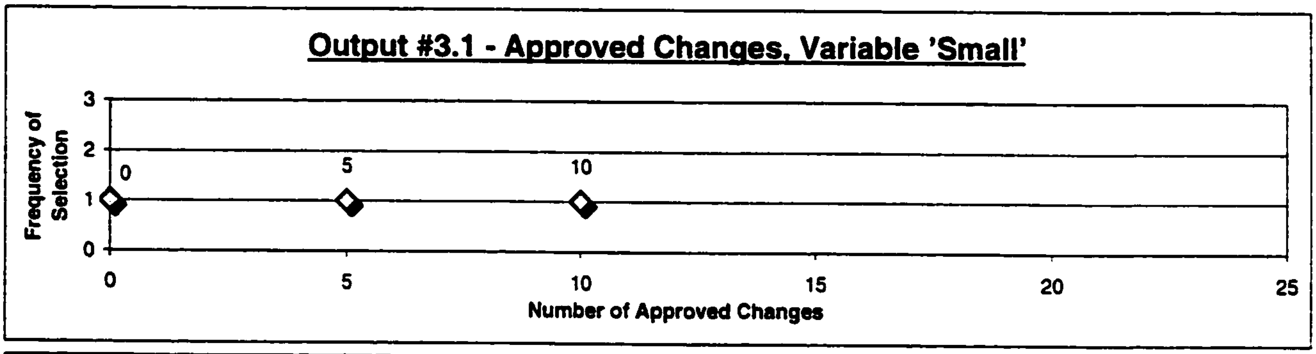

Output #2.1 - Change in Scheduled Design Duration, Variable 'Large'
Frequency of Selection
Percent (%) Change
-20
-10

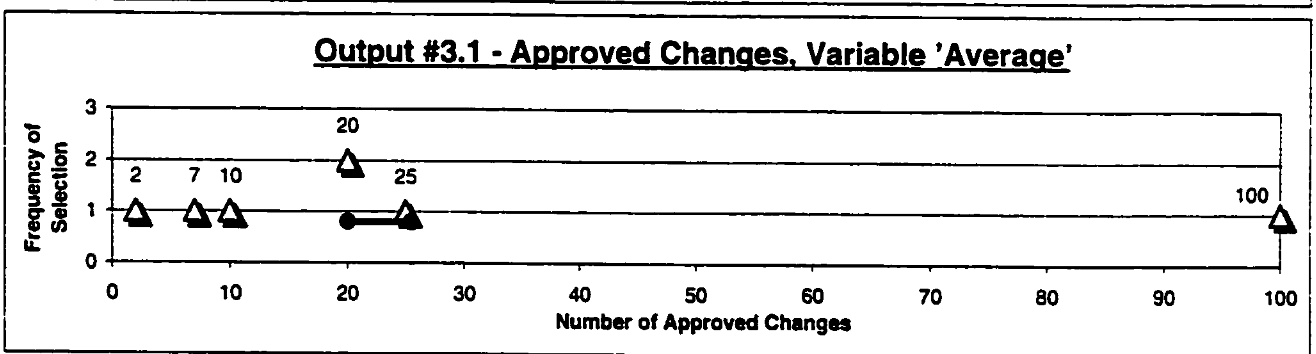

Output #2.2 - Document Deadlines Missed, Variable 'Small'
Frequency of Selection
Percent (%) of Deadlines Missed
0
5

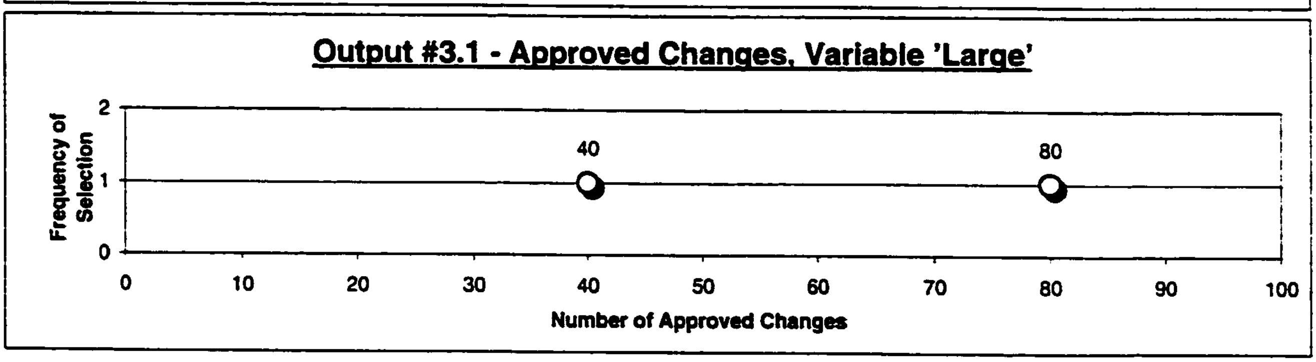

Output #2.2 - Document Deadlines Missed, Variable 'Average'
Frequency of Selection
Percent (%) of Deadlines Missed
20

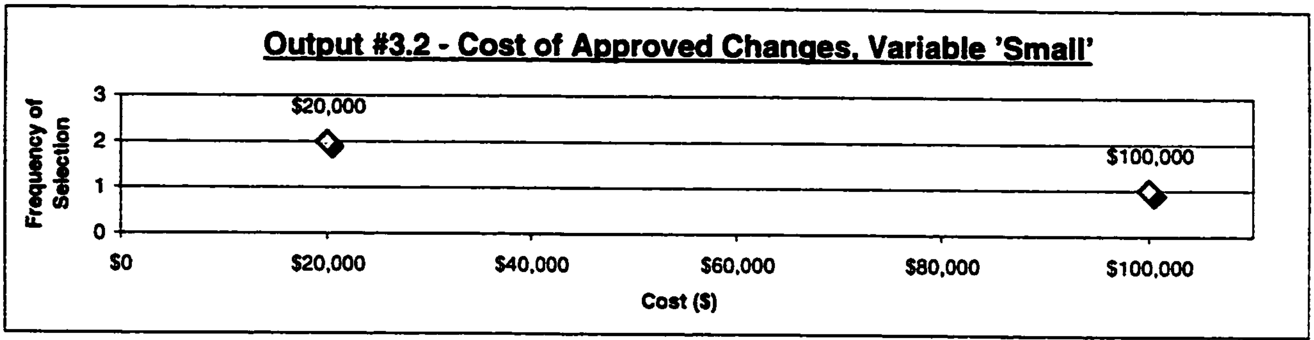

Output #2.2 - Document Deadlines Missed, Variable 'Large'
Frequency of Selection
Percent (%) of Deadlines Missed
0
30


Output #3 - Overall Accuracy of Design
Low
Average
High
Frequency of Selection
Rating

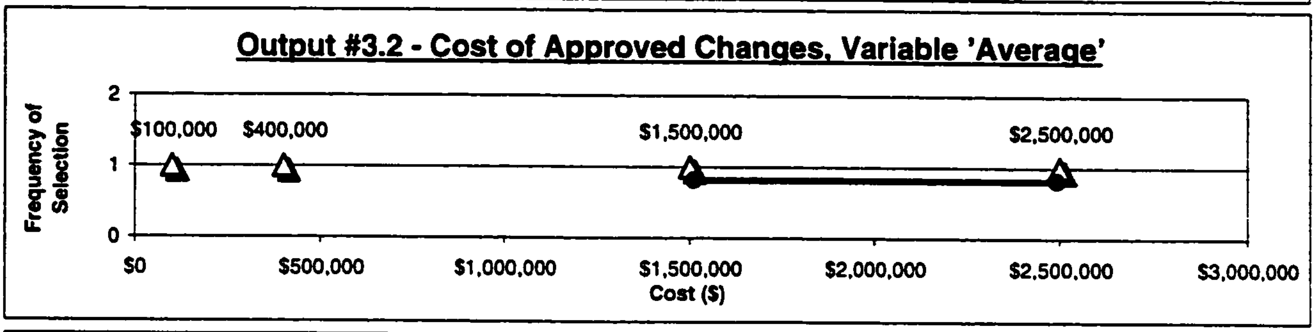

Output #3.1 - Approved Changes, Variable 'Small'
Frequency of Selection
Number of Approved Changes


Output #3.1 - Approved Changes, Variable 'Average'
Frequency of Selection
Number of Approved Changes

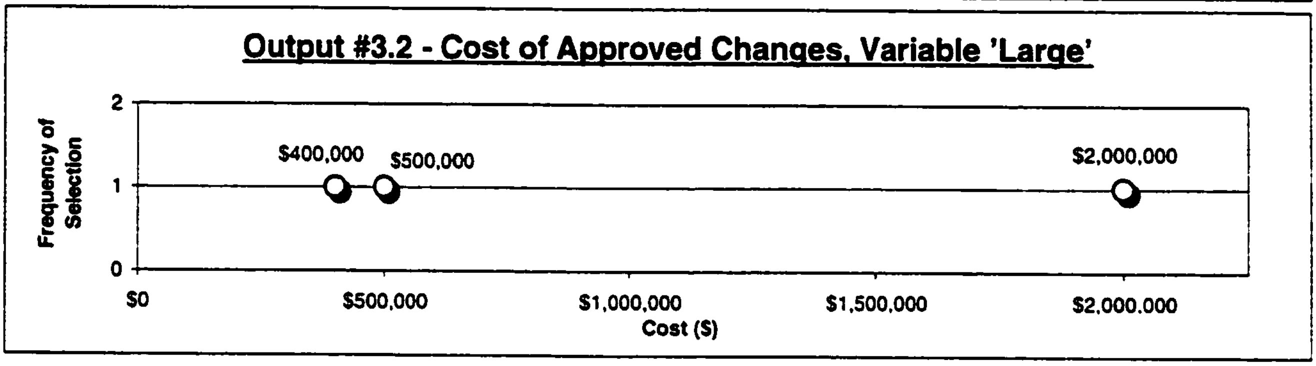

Output #3.1 - Approved Changes, Variable 'Large'
Frequency of Selection
Number of Approved Changes



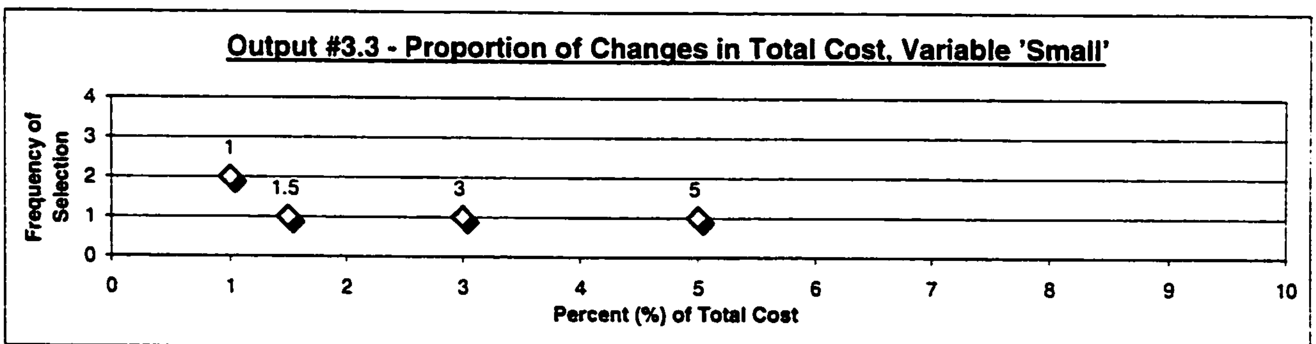
Output #3.2 - Cost of Approved Changes, Variable 'Small'
Frequency of Selection
$20,000
$100,000
$0
$20,000
$40,000
$60,000
$80,000
$100,000
Cost ($)



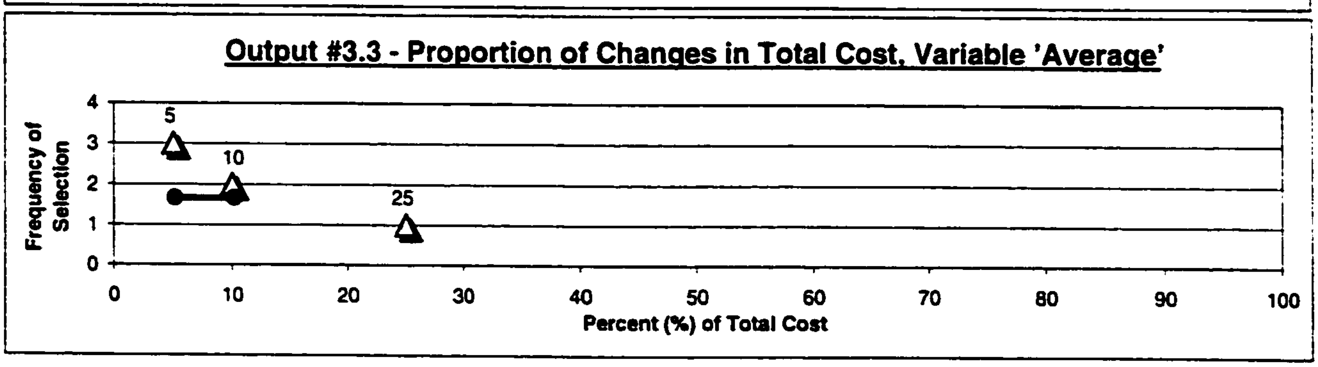
Output #3.2 - Cost of Approved Changes, Variable 'Average'
Frequency of Selection
$100,000
$400,000
$1,500,000
$2,500,000
$0
$500,000
$1,000,000
$1,500,000
$2,000,000
$2,500,000
$3,000,000
Cost ($)



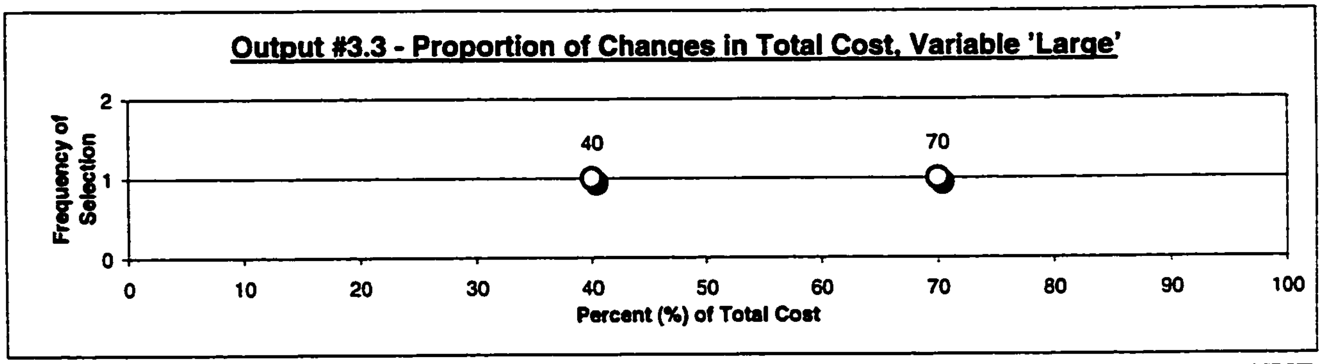
Output #3.2 - Cost of Approved Changes, Variable 'Large'
Frequency of Selection
$400,000
$500,000
$2,000,000
$0
$500,000
$1,000,000
$1,500,000
$2,000,000
Cost ($)



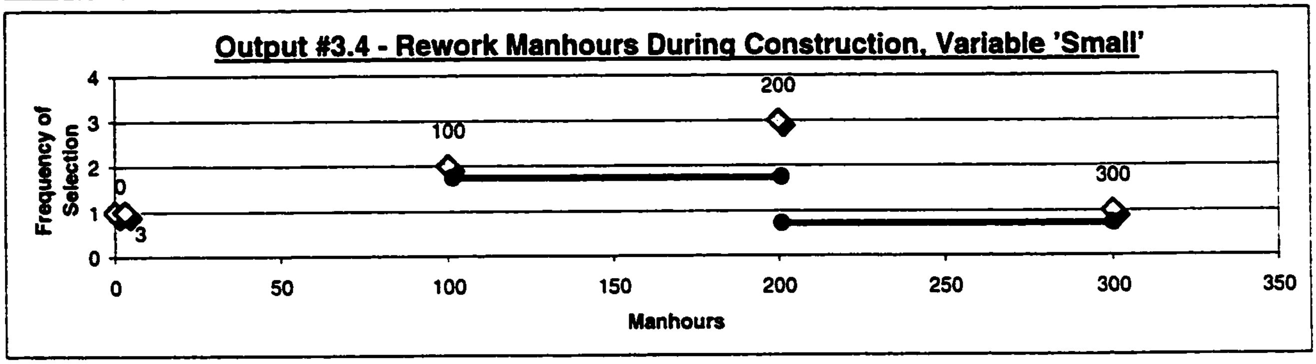
Output #3.3 - Proportion of Changes in Total Cost, Variable 'Small'
Frequency of Selection
1
1.5
3
5
Percent (%) of Total Cost



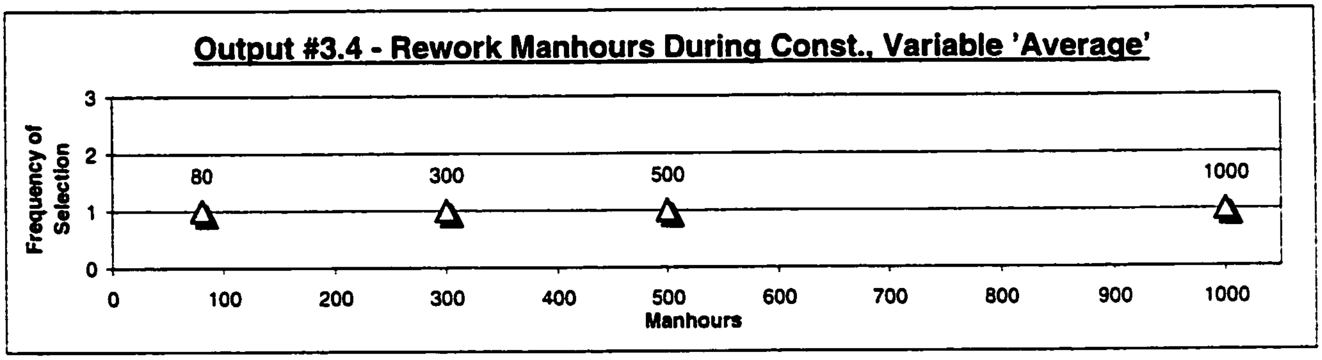
Output #3.3 - Proportion of Changes in Total Cost, Variable 'Average'
Frequency of Selection
5
10
25
Percent (%) of Total Cost



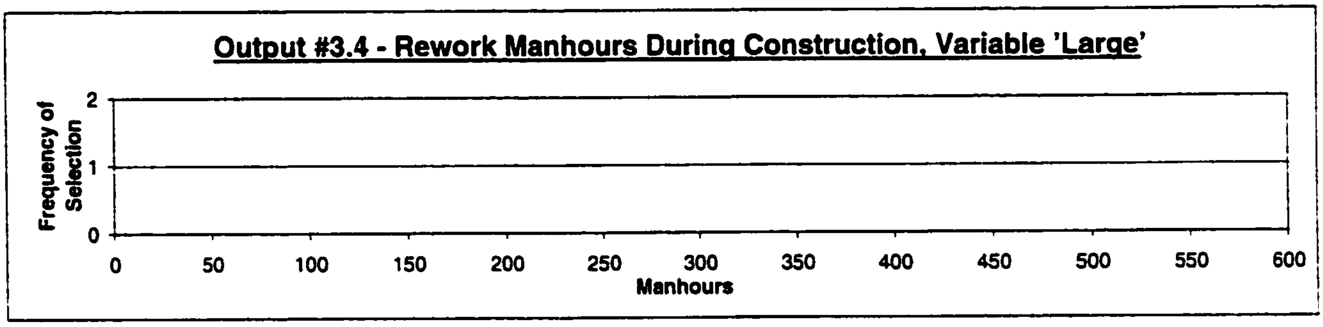
Output #3.3 - Proportion of Changes in Total Cost, Variable 'Large'
Frequency of Selection
Percent (%) of Total Cost
40
70



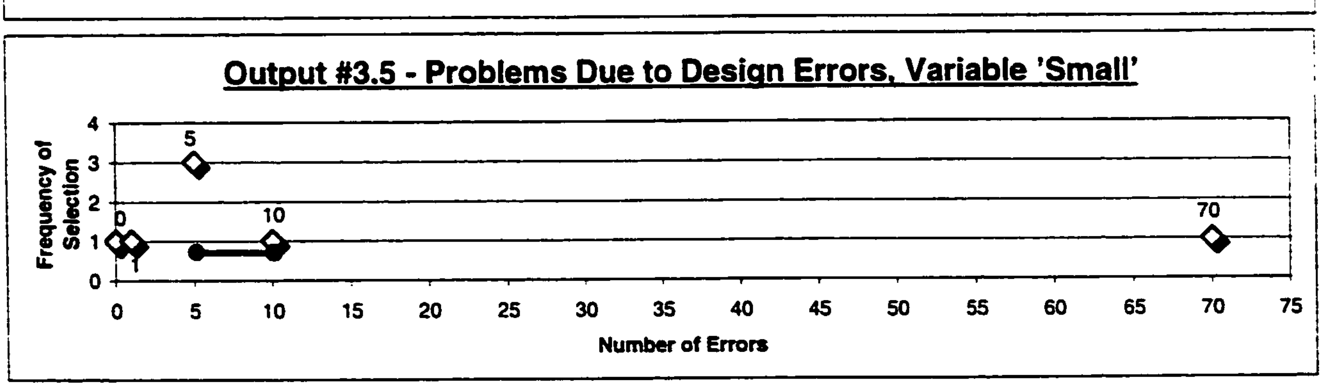
Output #3.4 - Rework Manhours During Construction, Variable 'Small'
Frequency of Selection
Manhours
0
3
100
200
300



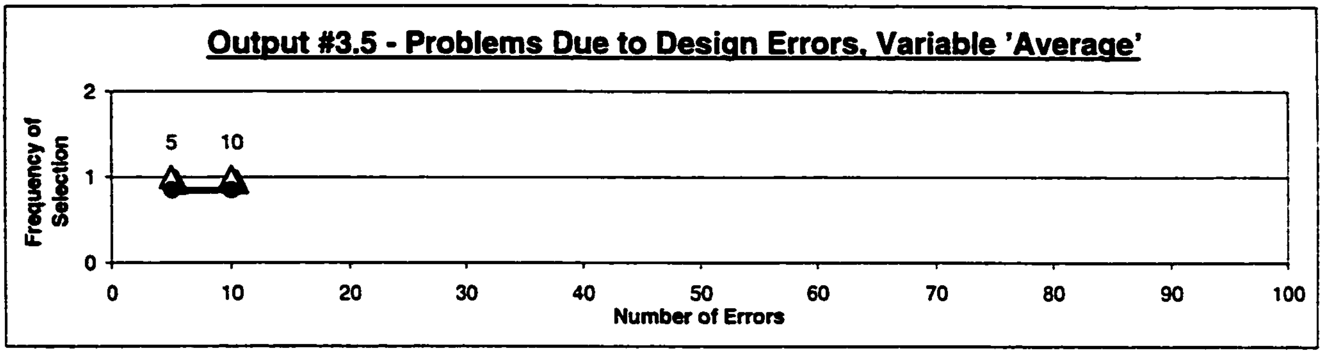
Output #3.4 - Rework Manhours During Const., Variable 'Average'
Frequency of Selection
Manhours
80
300
500
1000



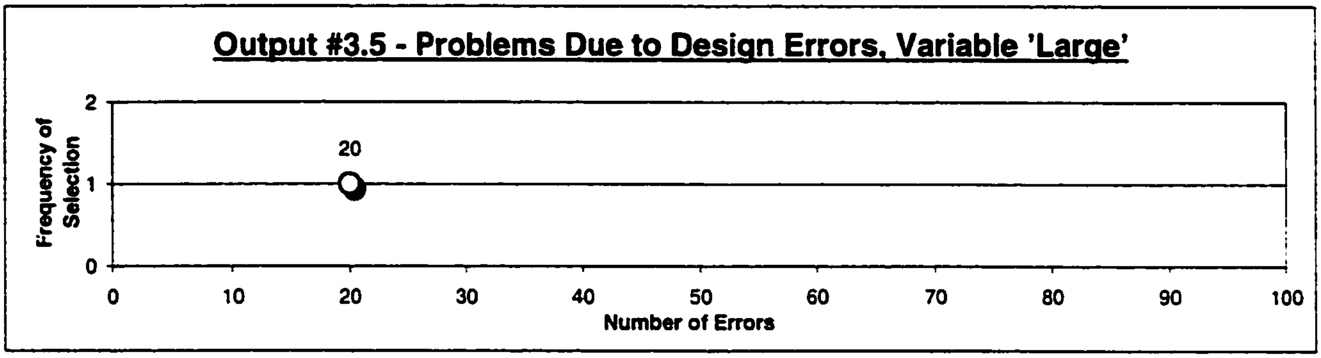
Output #3.4 - Rework Manhours During Construction, Variable 'Large'
Frequency of Selection
Manhours



Output #3.5 - Problems Due to Design Errors, Variable 'Small'
Frequency of Selection
Number of Errors
0
1
5
10
70

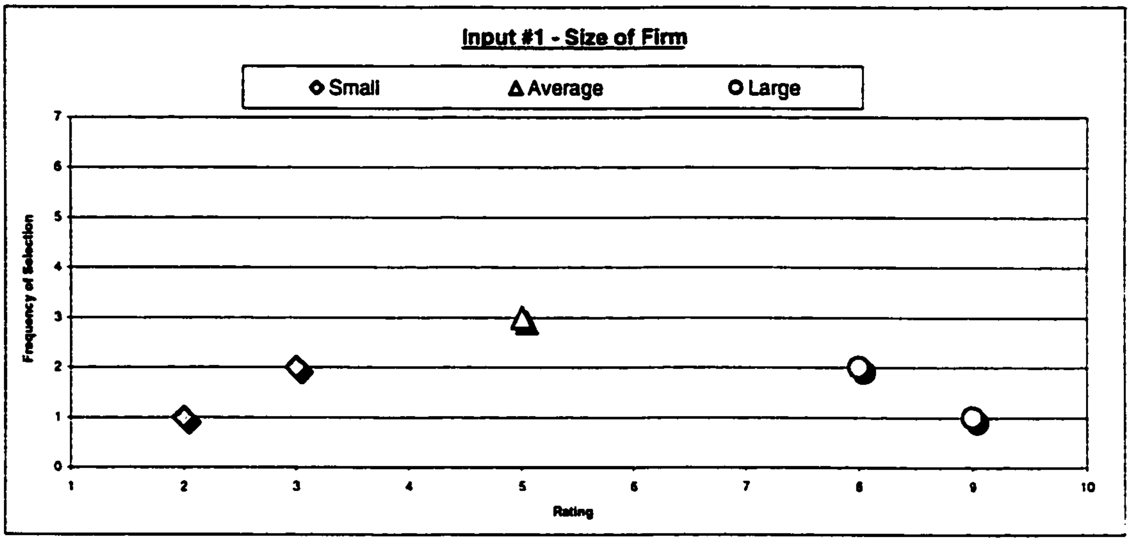


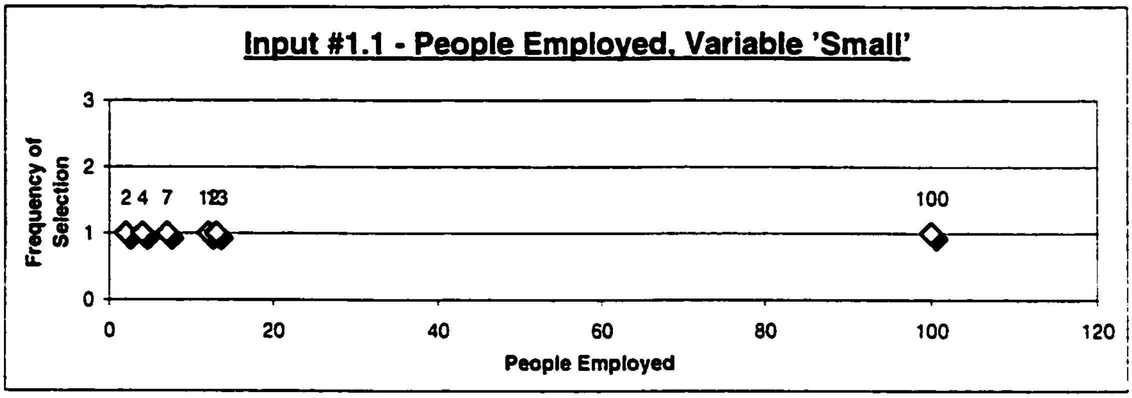
Output #3.5 - Problems Due to Design Errors, Variable 'Average'
Frequency of Selection
Number of Errors
5
10
0
1
2
0
10
20
30
40
50
60
70
80
90
100

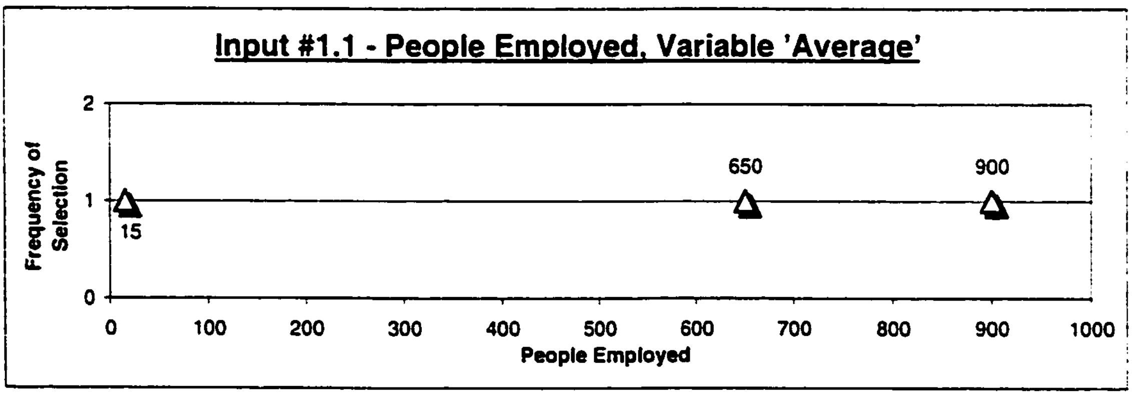


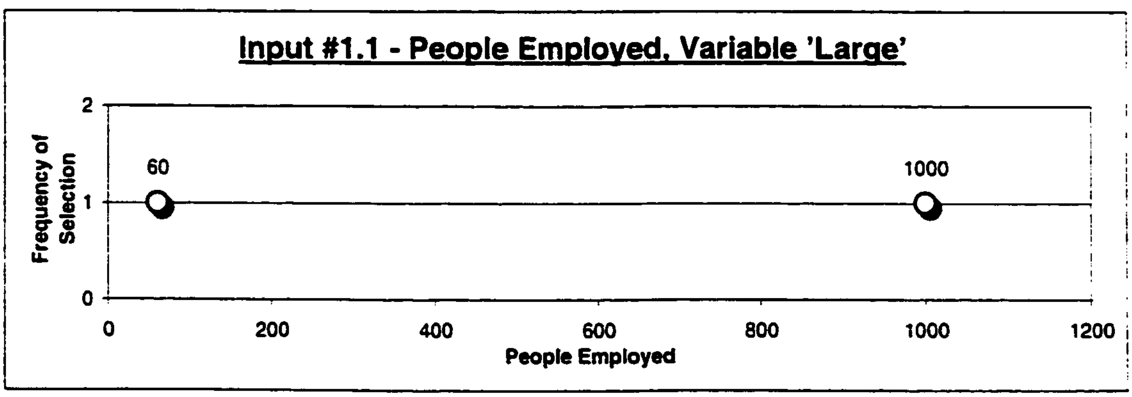
Output #3.5 - Problems Due to Design Errors, Variable 'Large'
Frequency of Selection
Number of Errors
20
0
1
2
0
10
20
30
40
50
60
70
80
90
100

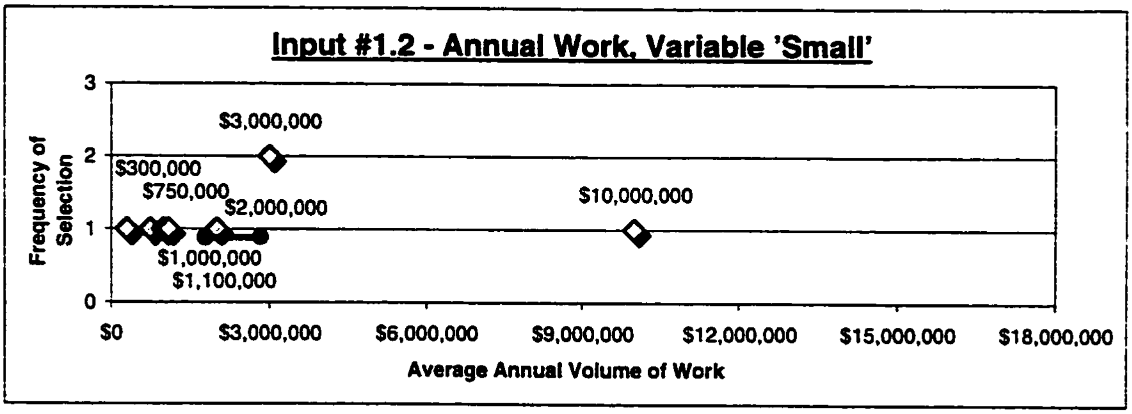
Appendix 3: Data Processing Results (Trial 2)

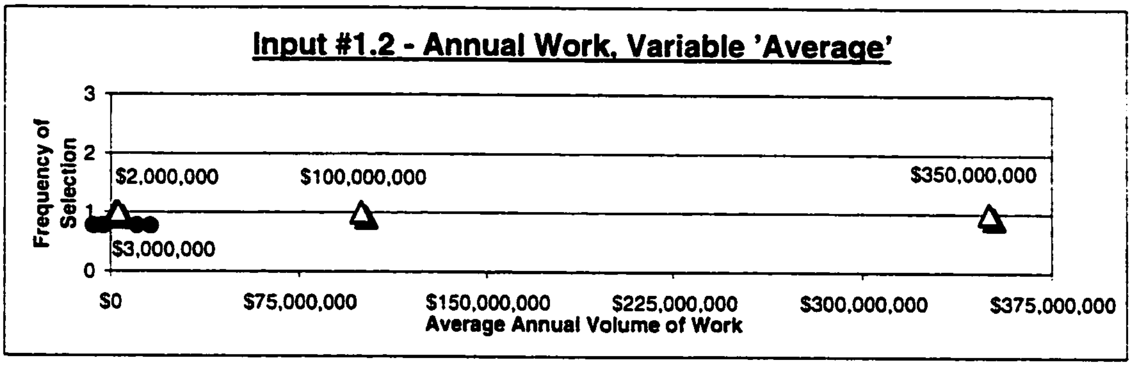


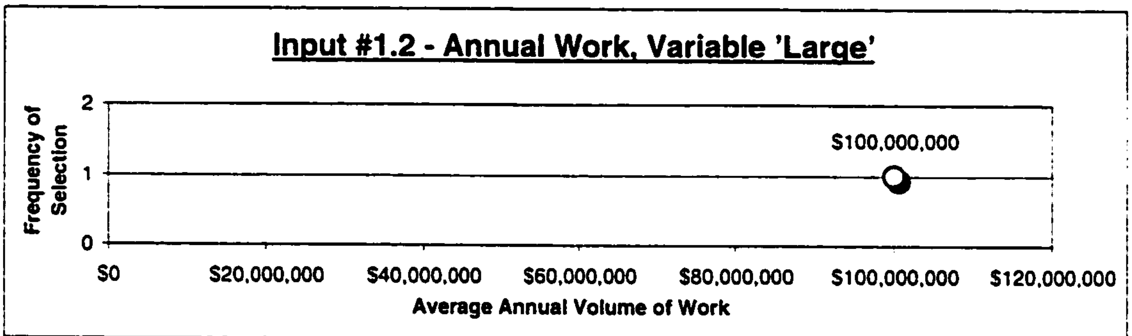


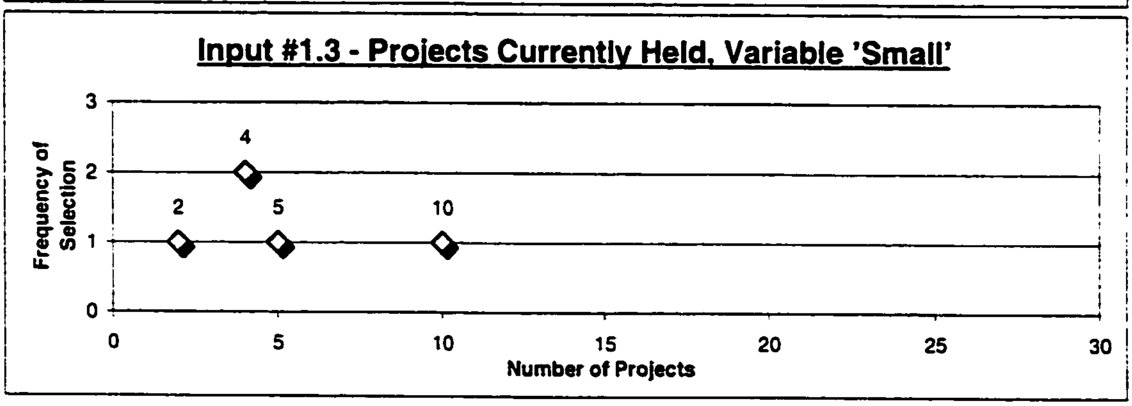


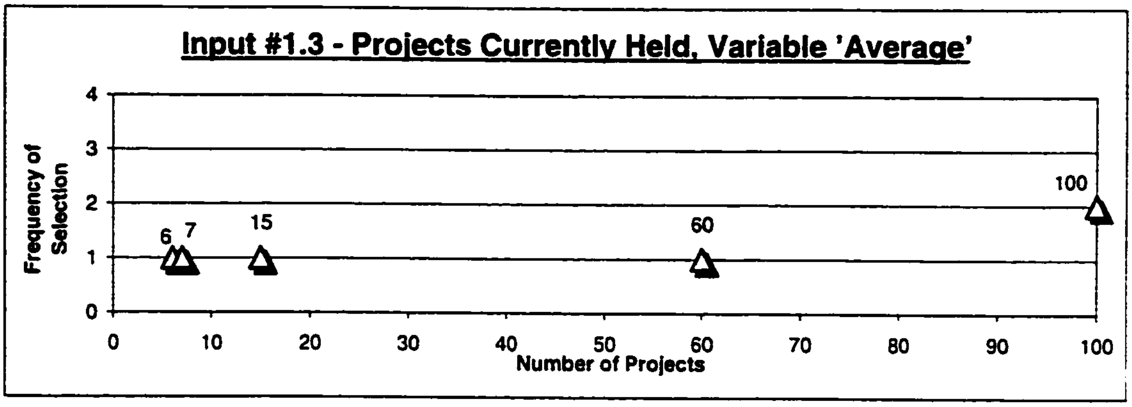


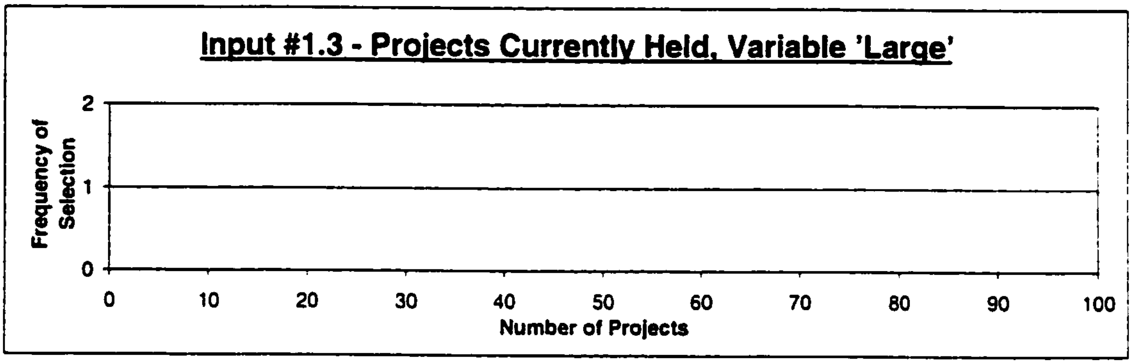

Input #1.2 - Annual Work, Variable 'Small'
Frequency of Selection
$3,000,000
$300,000
$750,000
$2,000,000
$10,000,000
$1,000,000
$1,100,000
$0
$3,000,000
$6,000,000
$9,000,000
$12,000,000
$15,000,000
$18,000,000
Average Annual Volume of Work

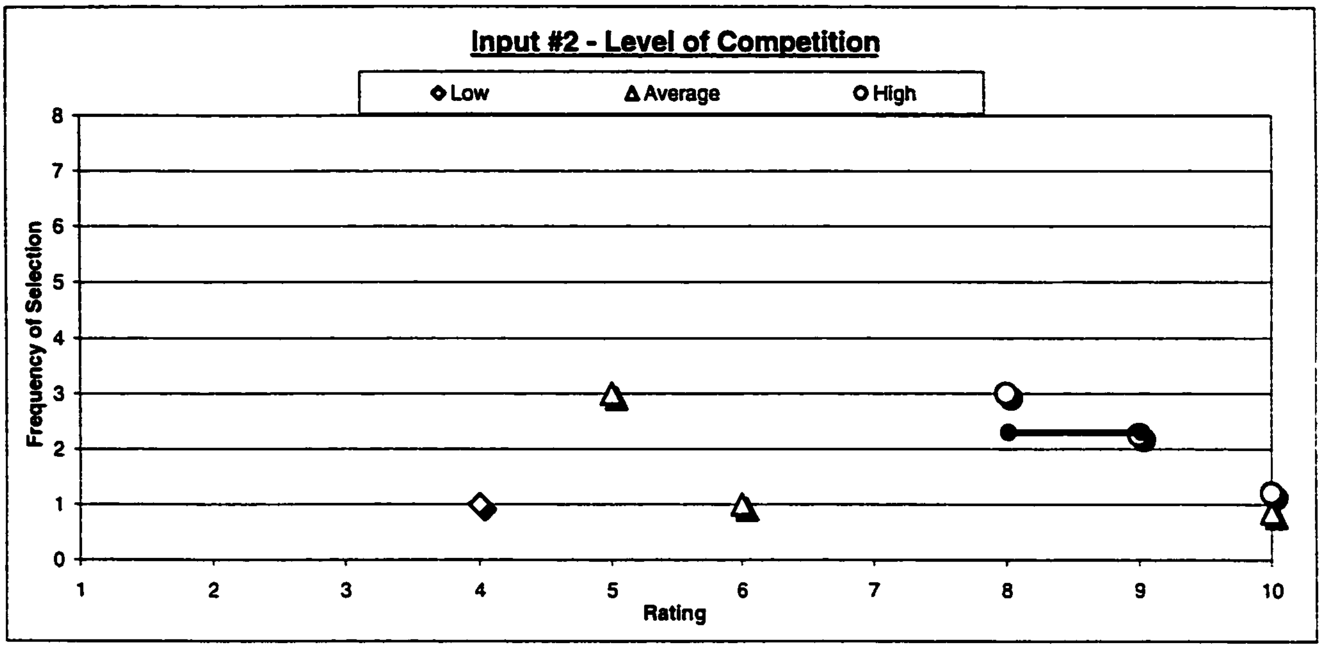

Input #1.2 - Annual Work, Variable 'Average'
Frequency of Selection
$2,000,000
$100,000,000
$350,000,000
$3,000,000
$0
$75,000,000
$150,000,000
$225,000,000
$300,000,000
$375,000,000
Average Annual Volume of Work


Input #1.2 - Annual Work, Variable 'Large'
Frequency of Selection
$100,000,000
$0
$20,000,000
$40,000,000
$60,000,000
$80,000,000
$100,000,000
$120,000,000
Average Annual Volume of Work

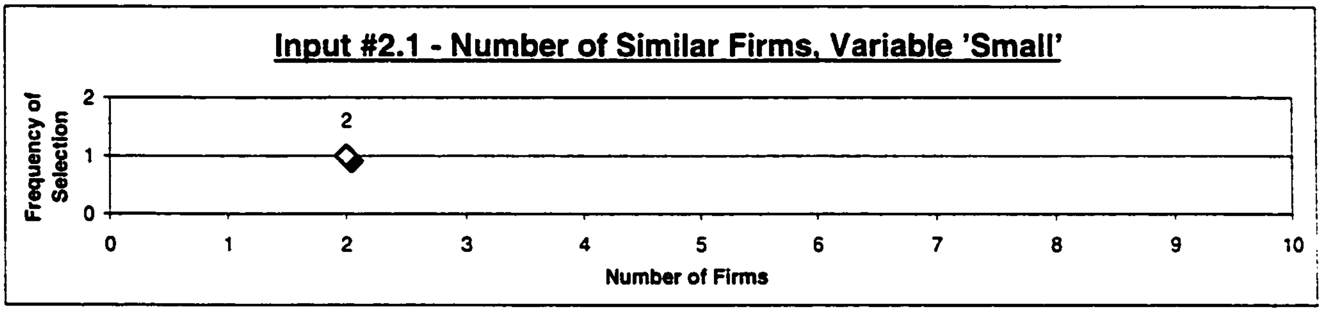

Input #1.3 - Projects Currently Held, Variable 'Small'
Frequency of Selection
4
2
5
10
Number of Projects


Input #1.3 - Projects Currently Held, Variable 'Average'
Frequency of Selection
Number of Projects
0
1
2
3
4
0
10
20
30
40
50
60
70
80
90
100
6
7
15
60
100

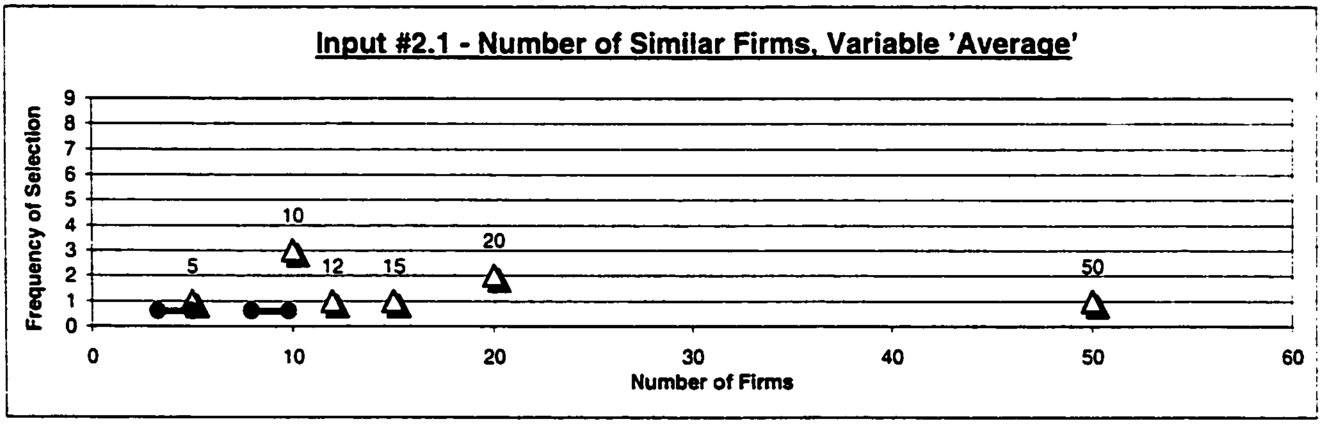

Input #1.3 - Projects Currently Held, Variable 'Large'
Frequency of Selection
Number of Projects
0
1
2
0
10
20
30
40
50
60
70
80
90
100



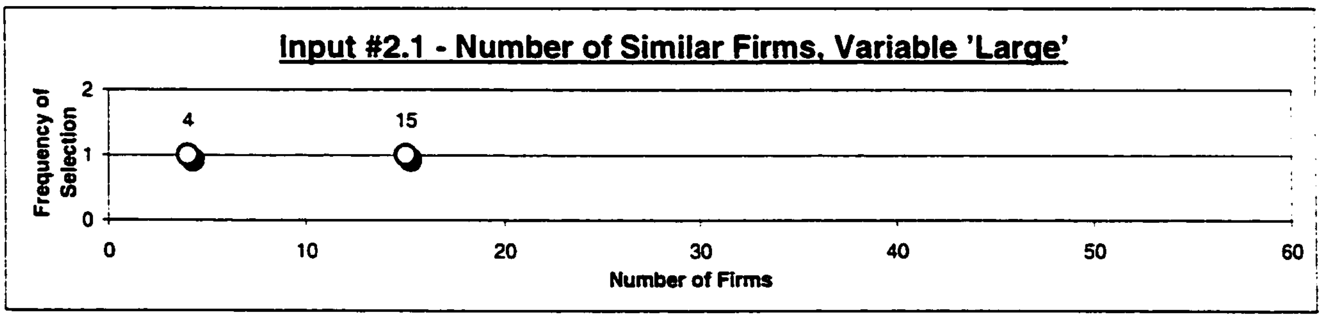
Input #2 - Level of Competition
Low
Average
High
Frequency of Selection
Rating
0
1
2
3
4
5
6
7
8
1
2
3
4
5
6
7
8
9
10



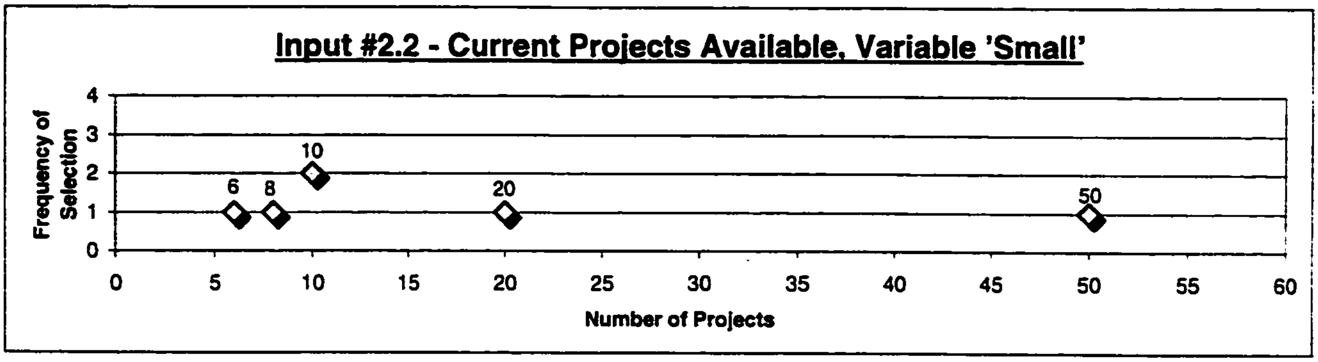
Input #2.1 - Number of Similar Firms, Variable 'Small'
Frequency of Selection
Number of Firms
2
0
1
2
0
1
2
3
4
5
6
7
8
9
10



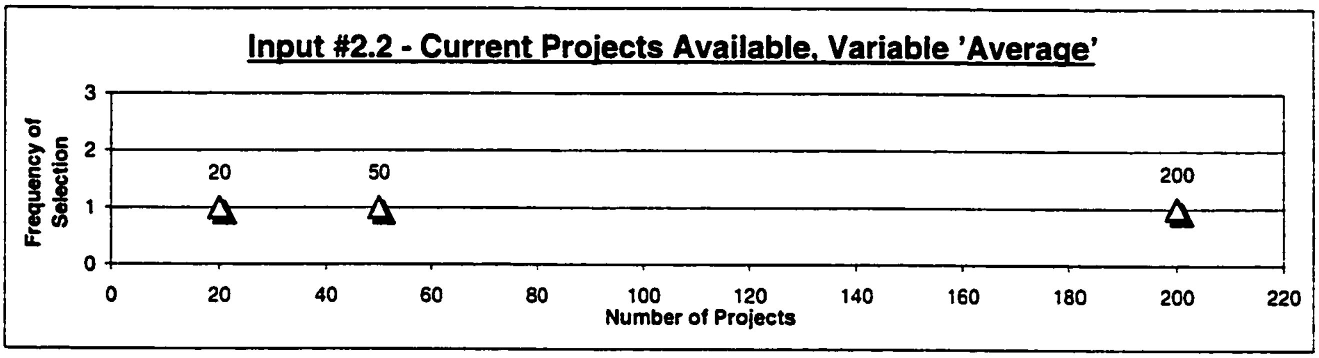
Input #2.1 - Number of Similar Firms, Variable 'Average'
Frequency of Selection
Number of Firms
5
10
12
15
20
50
0
1
2
3
4
5
6
7
8
9
0
10
20
30
40
50
60



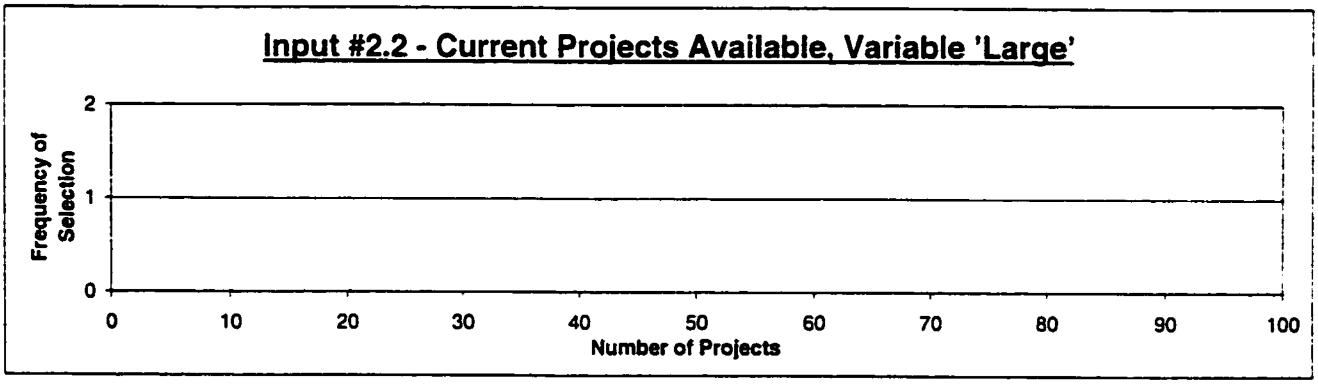
Input #2.1 - Number of Similar Firms, Variable 'Large'
Frequency of Selection
Number of Firms
4
15
0
1
2
0
10
20
30
40
50
60



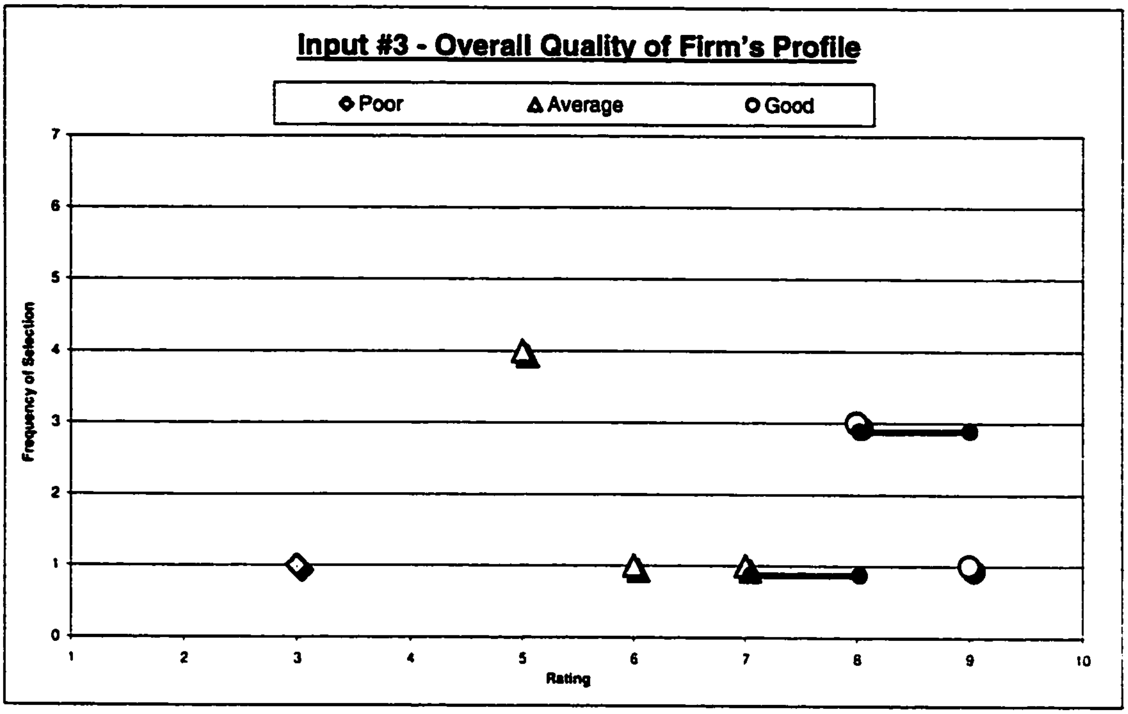
Input #2.2 - Current Projects Available, Variable 'Small'
Frequency of Selection
Number of Projects
6
8
10
20
50



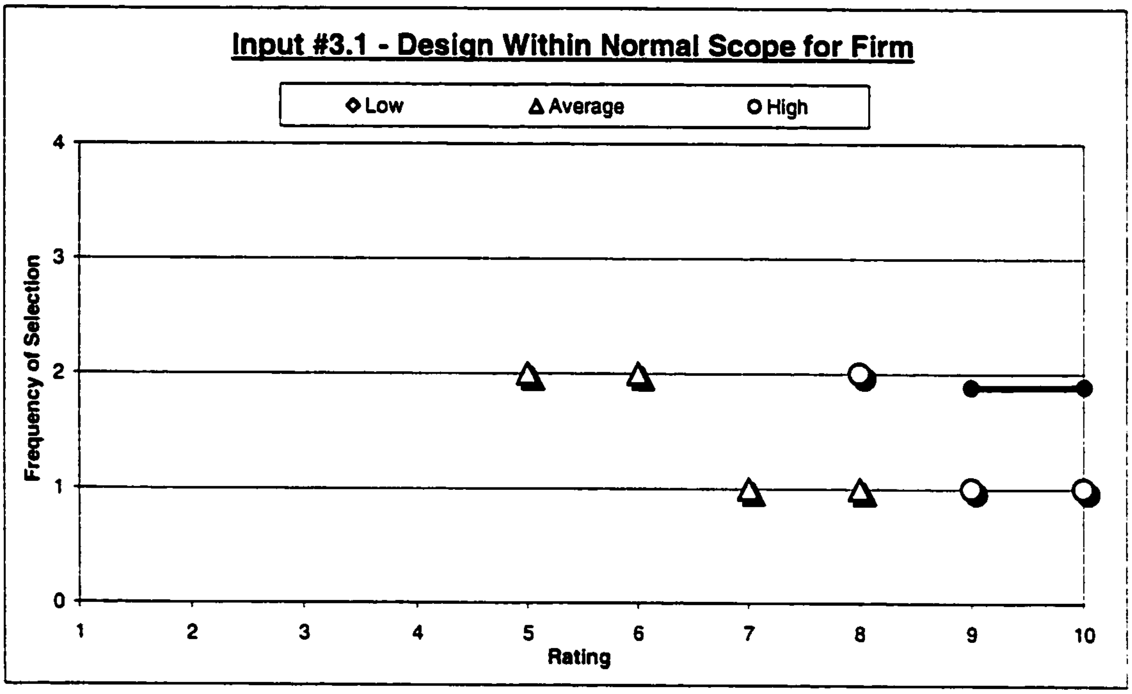
Input #2.2 - Current Projects Available, Variable 'Average'
Frequency of Selection
Number of Projects
20
50
200



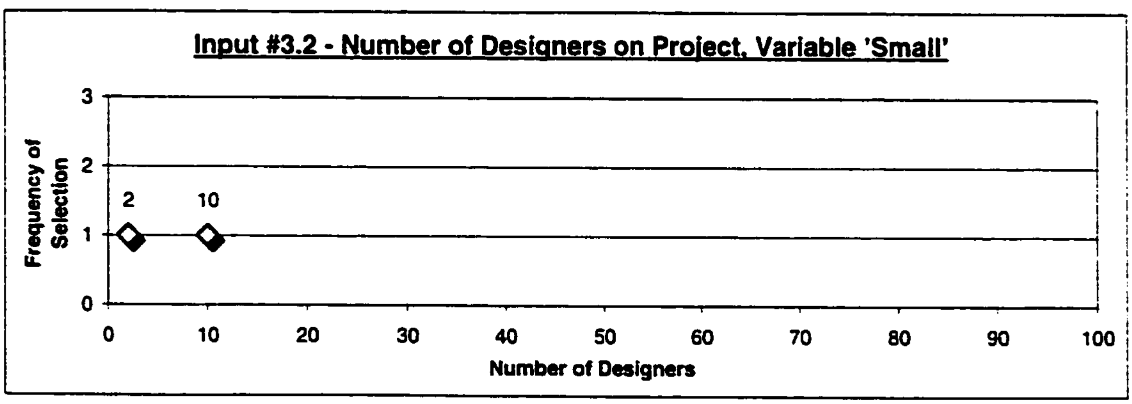
Input #2.2 - Current Projects Available, Variable 'Large'
Frequency of Selection
Number of Projects



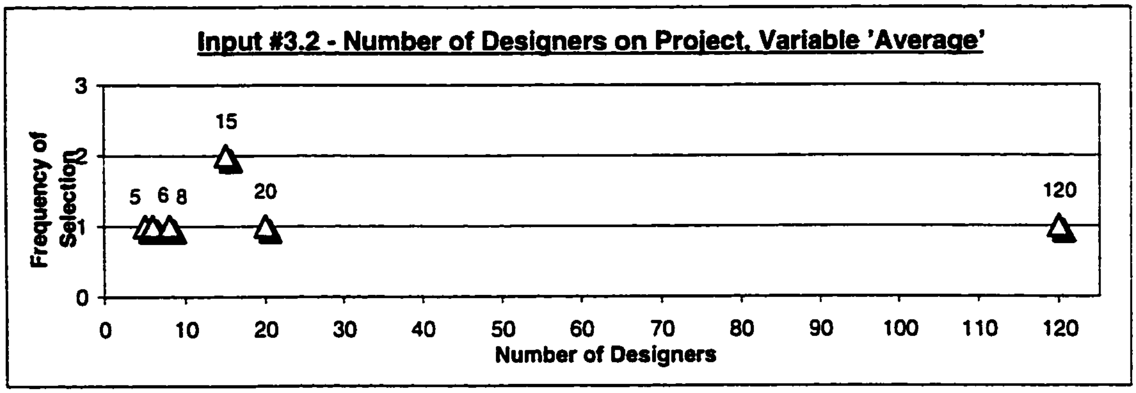
Input #3 - Overall Quality of Firm's Profile
Poor
Average
Good
Frequency of Selection
Rating
0
1
2
3
4
5
6
7
8
9
10

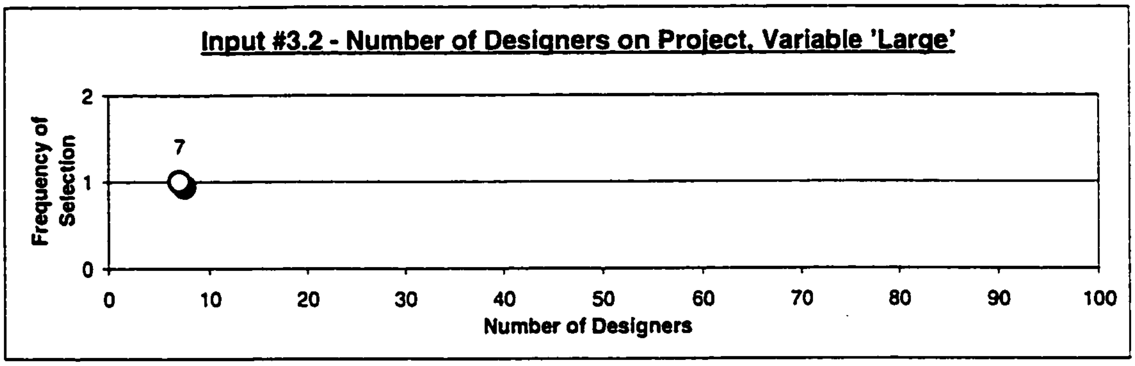


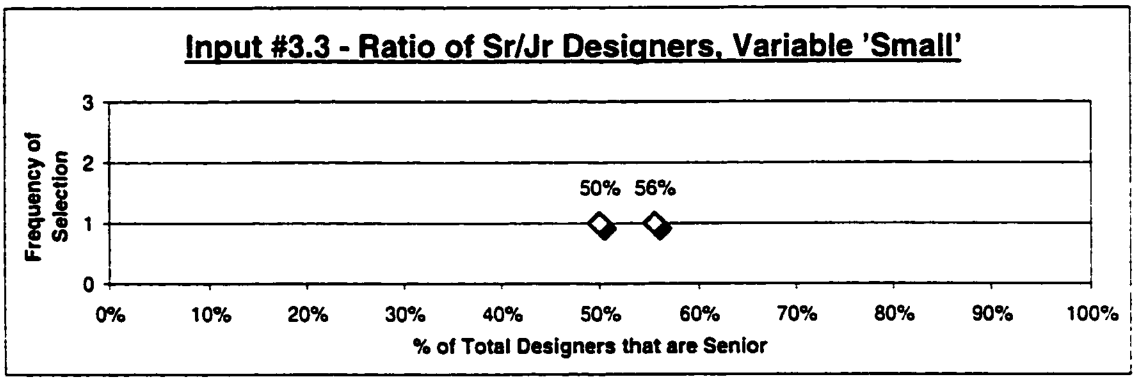
Input #3.1 - Design Within Normal Scope for Firm
Low
Average
High
Frequency of Selection
Rating
0
1
2
3
4
5
6
7
8
9
10

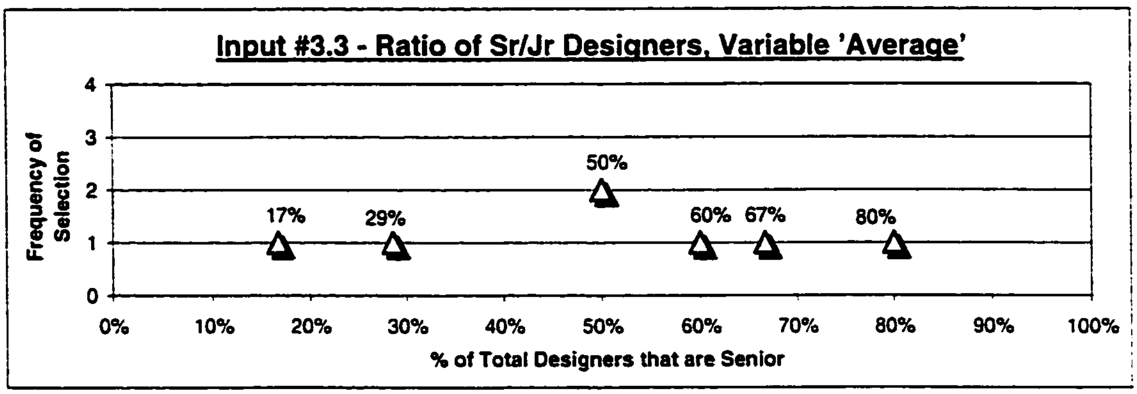


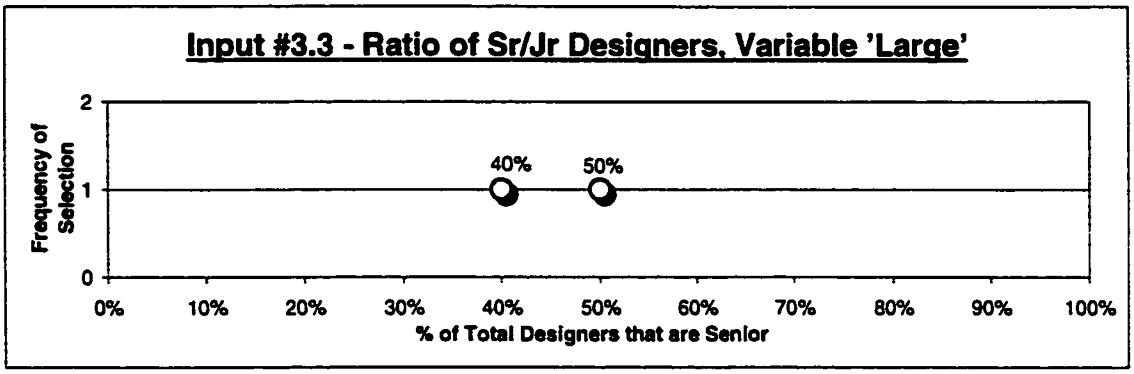
Input #3.2 - Number of Designers on Project, Variable 'Small'
Frequency of Selection
Number of Designers
2
10
0
1
2
3
10
20
30
40
50
60
70
80
90
100

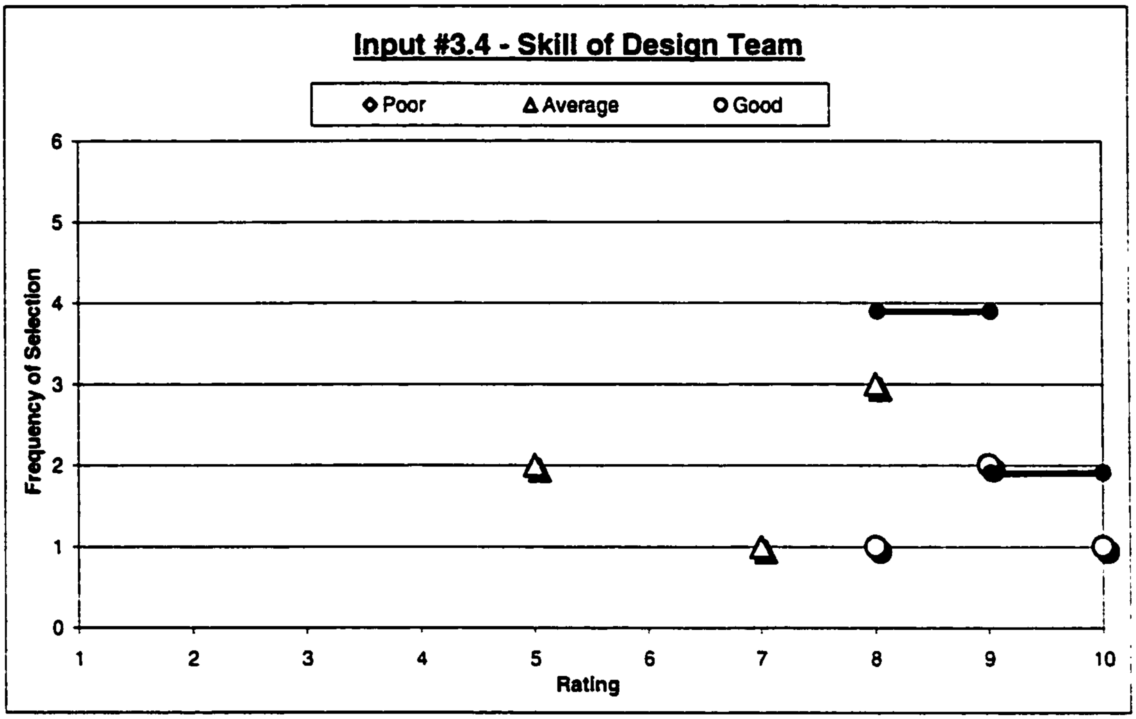

Input #3.2 - Number of Designers on Project, Variable 'Average'
Frequency of Selection
Number of Designers
15
5 6 8
20
120
0 10 20 30 40 50 60 70 80 90 100 110 120
0 1 2 3

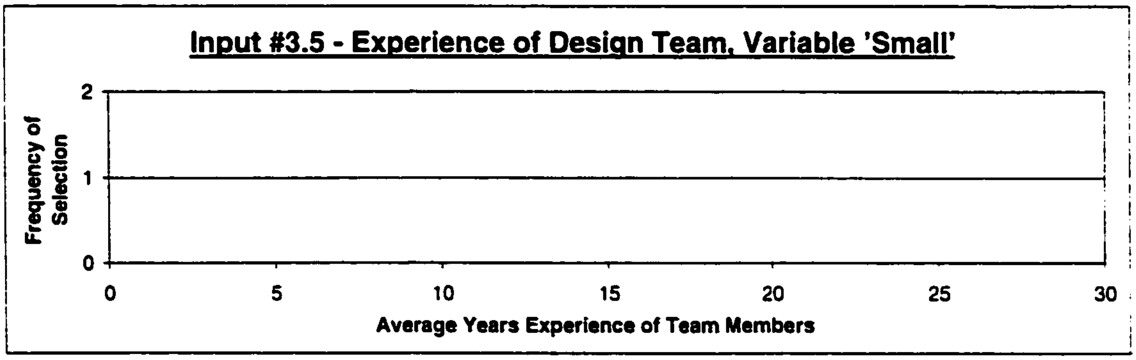

Input #3.2 - Number of Designers on Project, Variable 'Large'
Frequency of Selection
Number of Designers
7
0 10 20 30 40 50 60 70 80 90 100
0 1 2

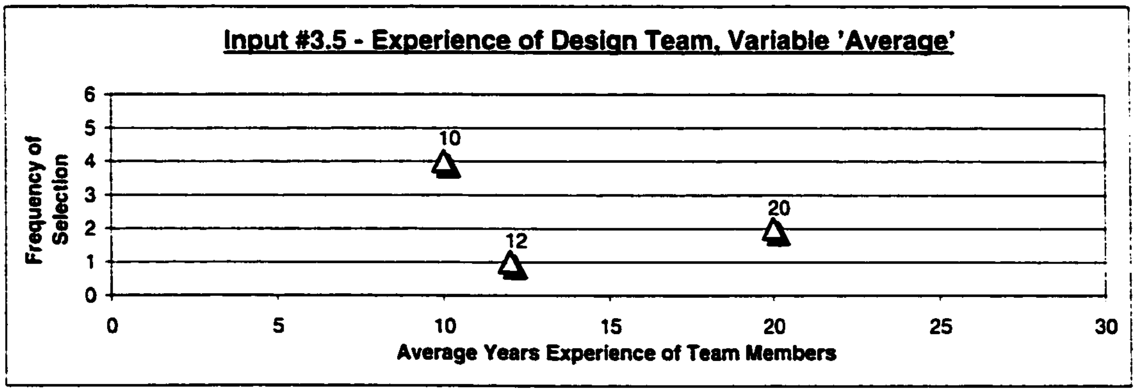

Input #3.3 - Ratio of Sr/Jr Designers, Variable 'Small'
Frequency of Selection
% of Total Designers that are Senior
50% 56%
0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%
0 1 2 3

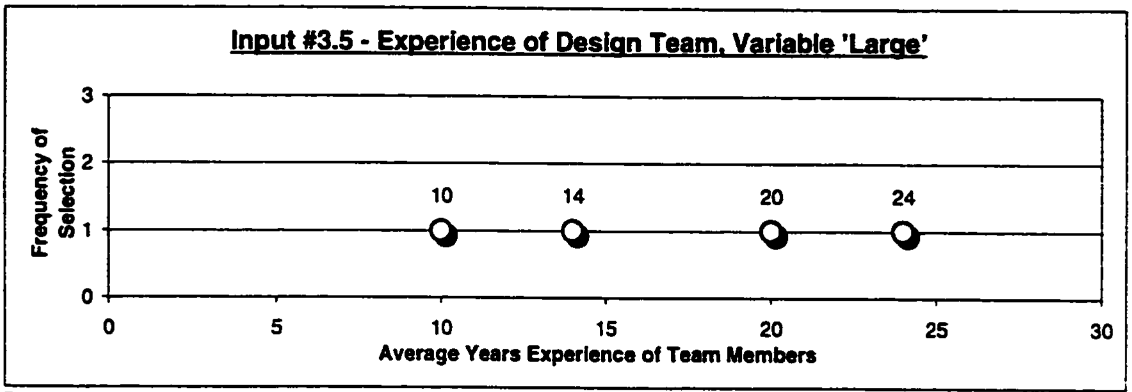

Input #3.3 - Ratio of Sr/Jr Designers, Variable 'Average'
Frequency of Selection
% of Total Designers that are Senior
50%
17% 29%
60% 67%
80%
0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%
0 1 2 3 4


Input #3.3 - Ratio of Sr/Jr Designers, Variable 'Large'
Frequency of Selection
40%
50%
0%
10%
20%
30%
40%
50%
60%
70%
80%
90%
100%
% of Total Designers that are Senior

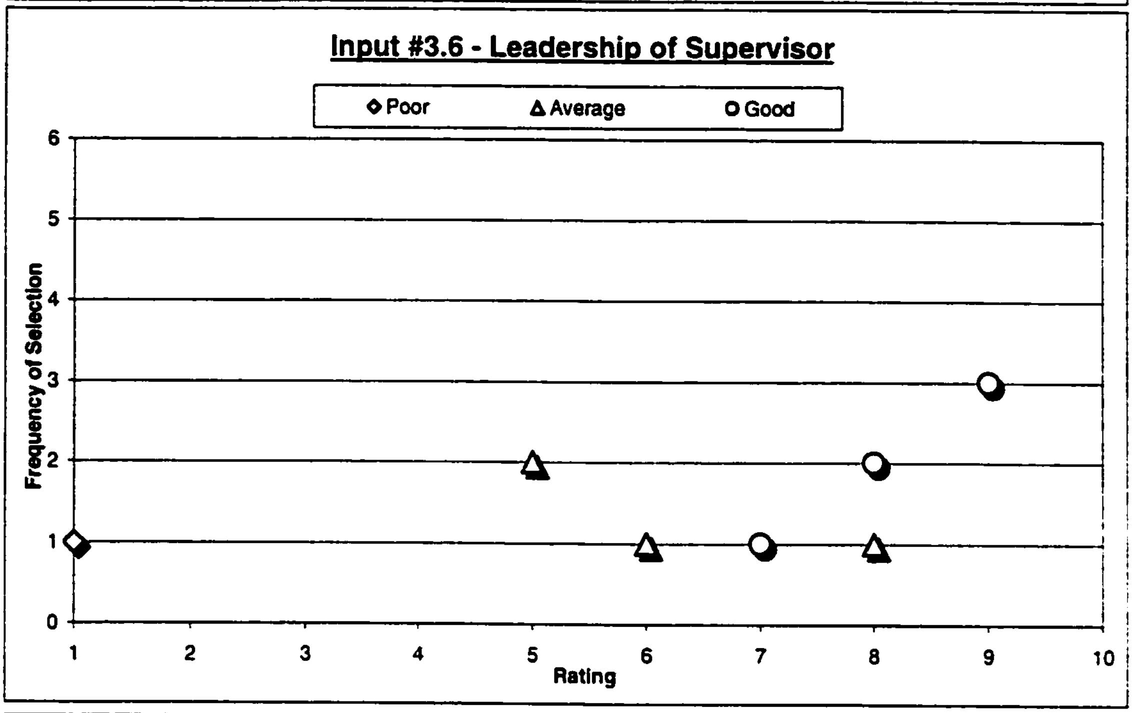

Input #3.4 - Skill of Design Team
Poor
Average
Good
Frequency of Selection
Rating


Input #3.5 - Experience of Design Team, Variable 'Small'
Frequency of Selection
Average Years Experience of Team Members

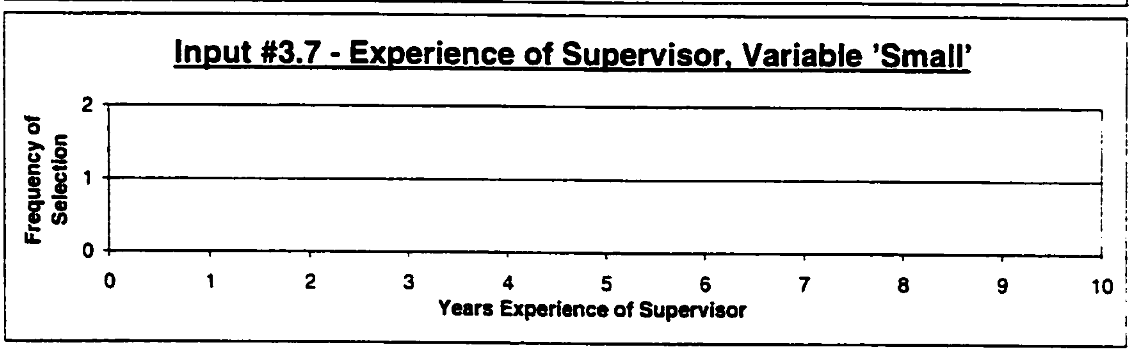

Input #3.5 - Experience of Design Team, Variable 'Average'
Frequency of Selection
10
12
20
Average Years Experience of Team Members



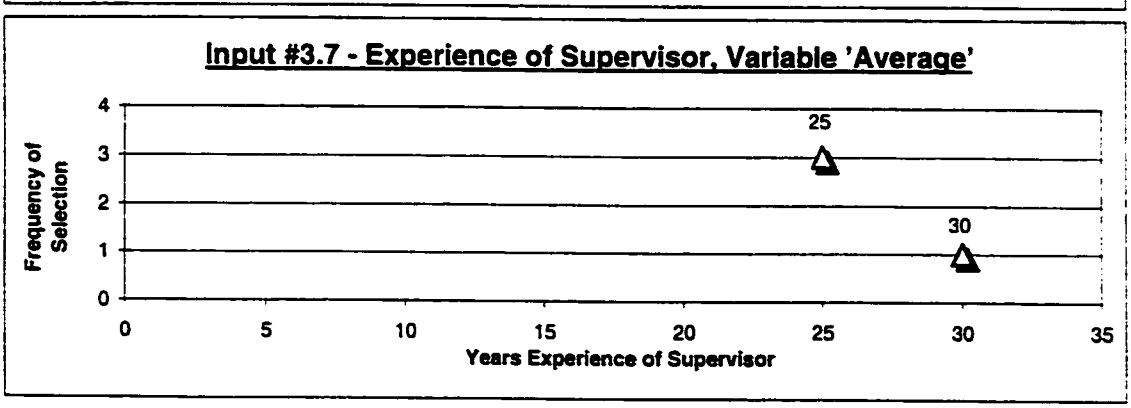
Input #3.5 - Experience of Design Team, Variable 'Large'
Frequency of Selection
Average Years Experience of Team Members
10
14
20
24
0
1
2
3
5
15
25
30



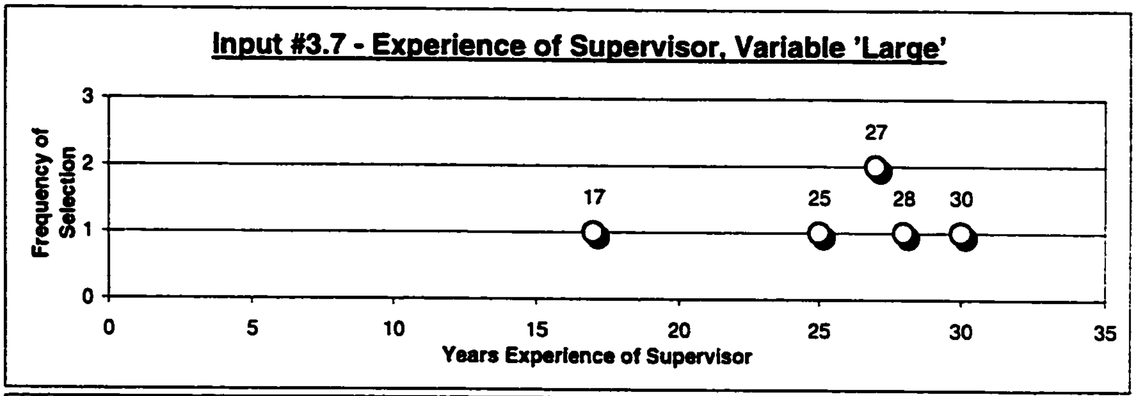
Input #3.6 - Leadership of Supervisor
Poor
Average
Good
Frequency of Selection
Rating
0
1
2
3
4
5
6
7
8
9
10



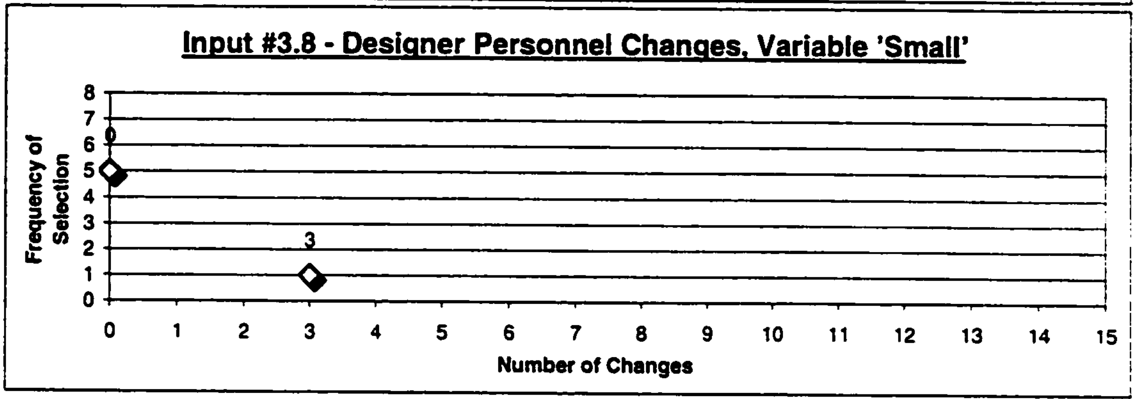
Input #3.7 - Experience of Supervisor, Variable 'Small'
Frequency of Selection
Years Experience of Supervisor
0
1
2
3
4
5
6
7
8
9
10



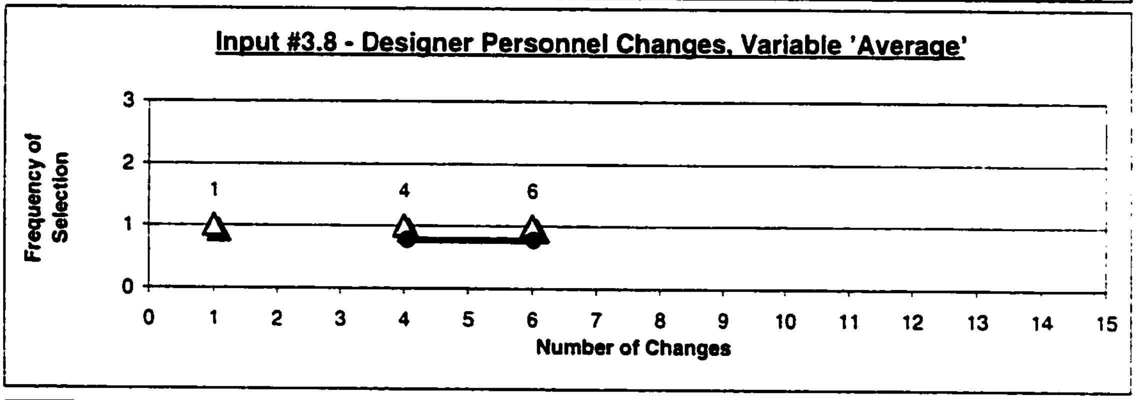
Input #3.7 - Experience of Supervisor, Variable 'Average'
Frequency of Selection
Years Experience of Supervisor
25
30
0
1
2
3
4
5
10
15
20
35



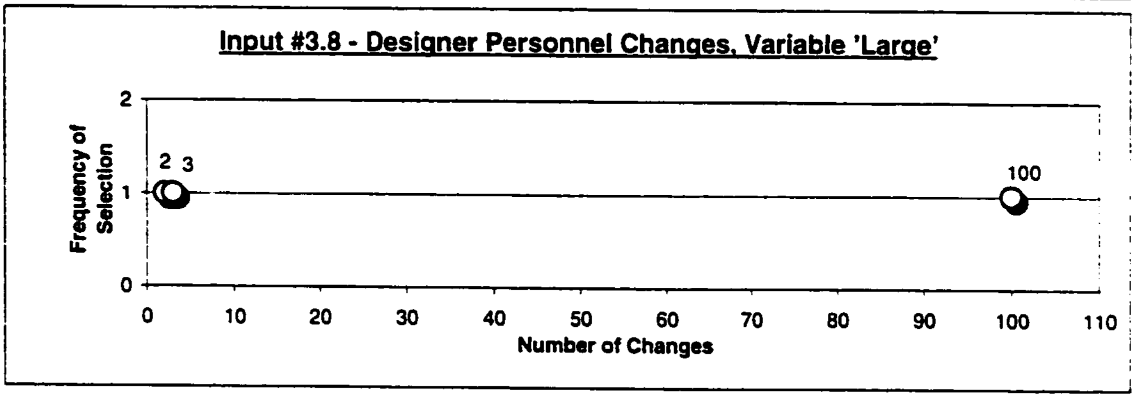
Input #3.7 - Experience of Supervisor, Variable 'Large'
Frequency of Selection
Years Experience of Supervisor
17
25
27
28
30
0
5
10
15
20
25
30
35



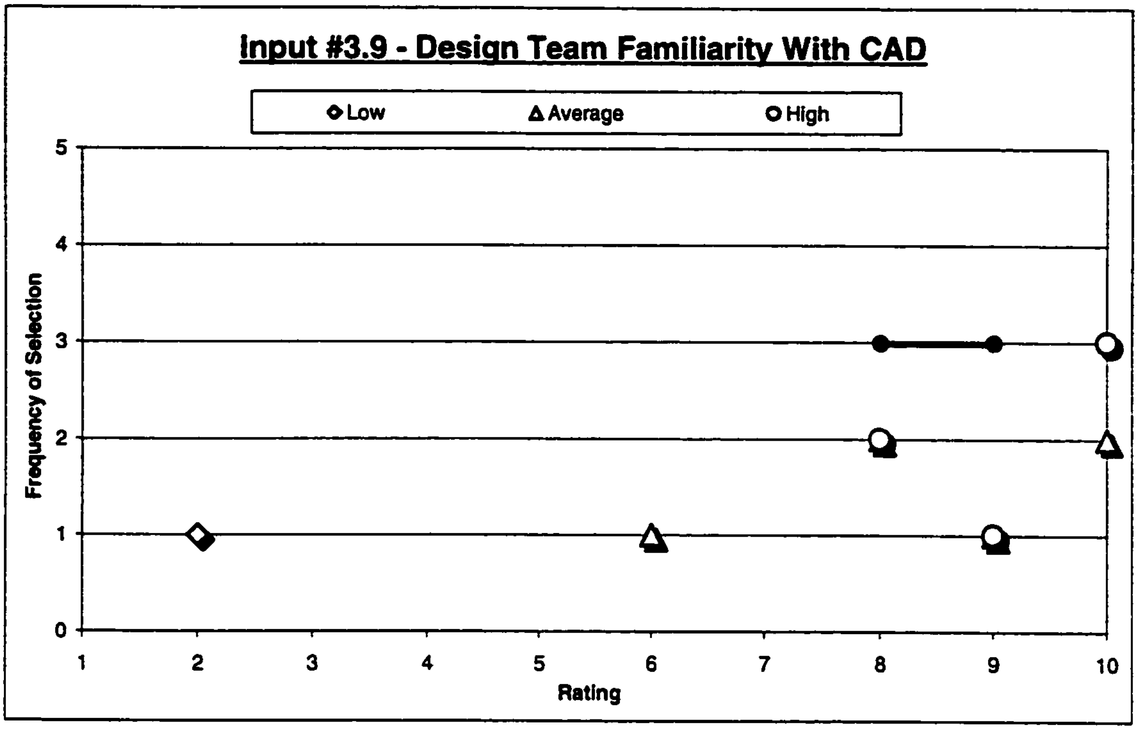
Input #3.8 - Designer Personnel Changes, Variable 'Small'
Frequency of Selection
Number of Changes
0
3



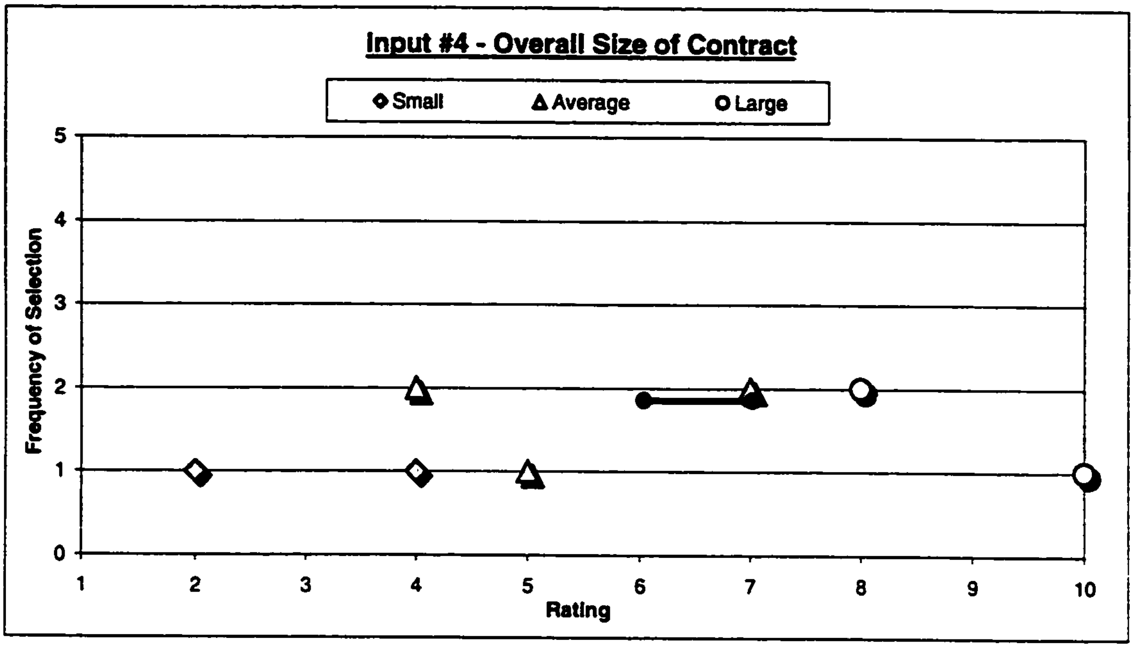
Input #3.8 - Designer Personnel Changes, Variable 'Average'
Frequency of Selection
Number of Changes
1
4
6

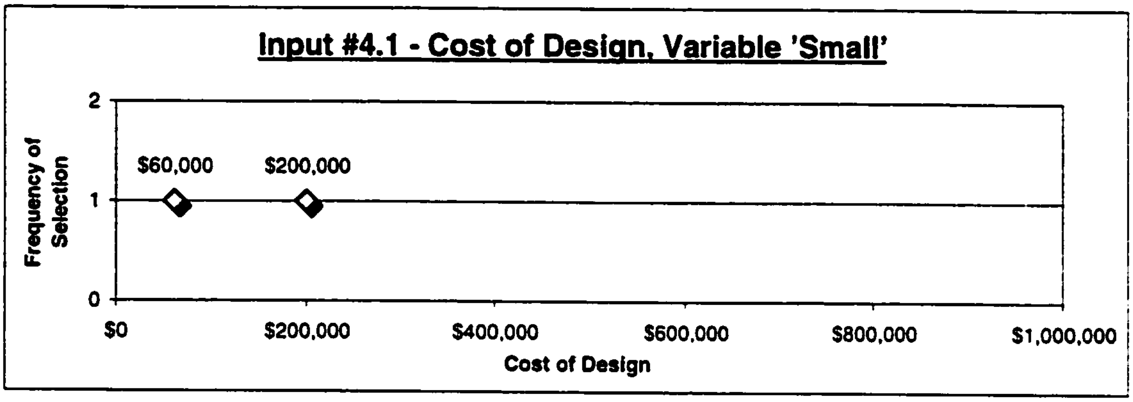


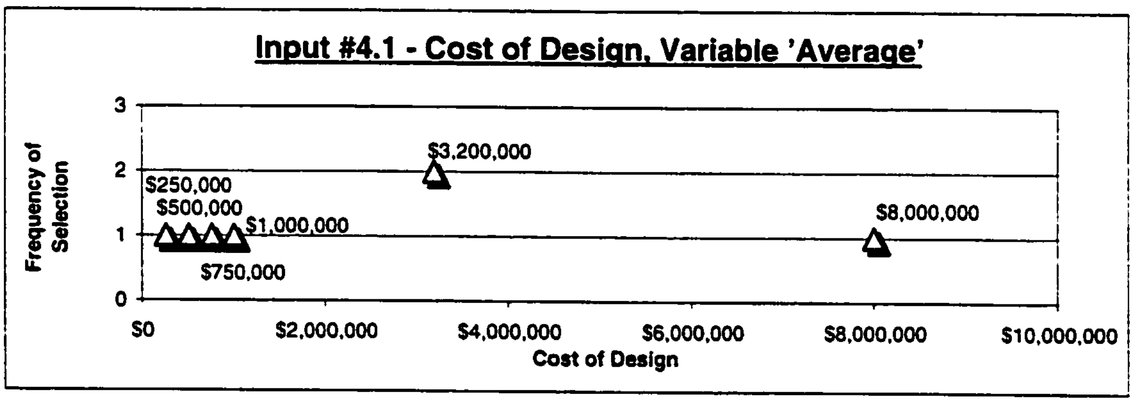
Input #3.8 - Designer Personnel Changes, Variable 'Large'
Frequency of Selection
Number of Changes
2 3
100

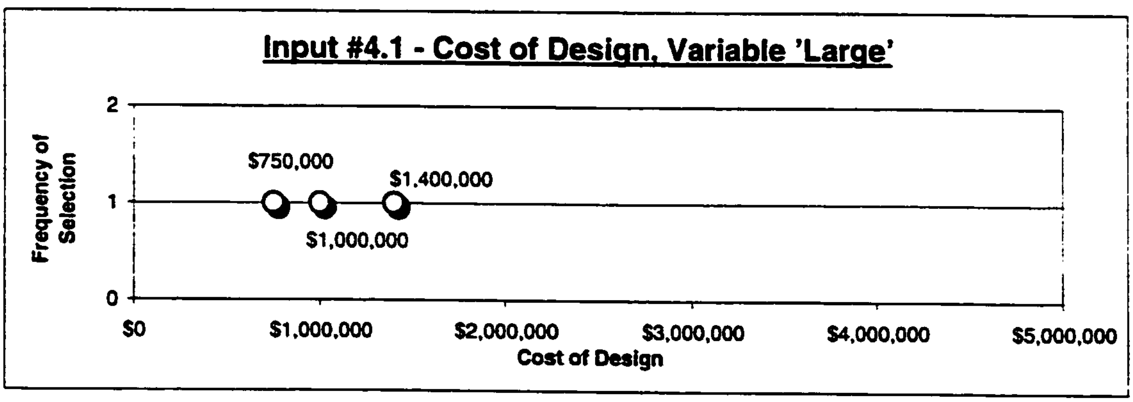


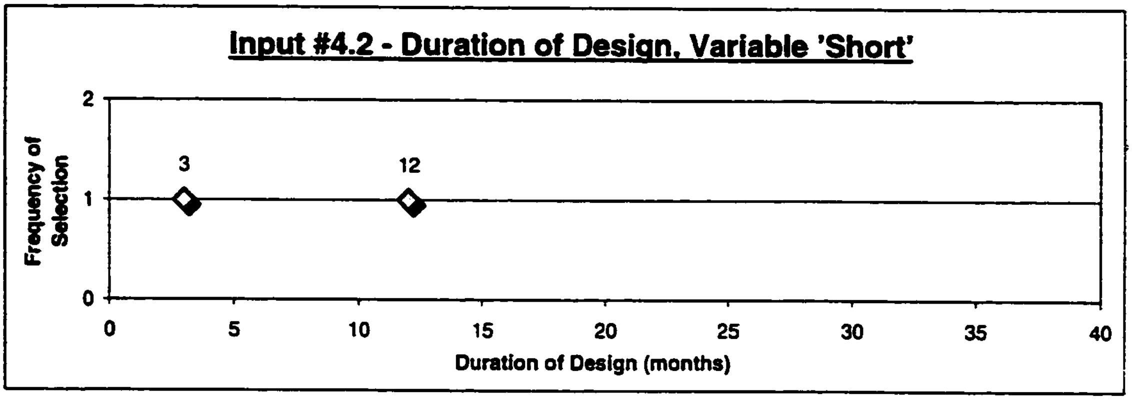
Input #3.9 - Design Team Familiarity With CAD
Low
Average
High
Frequency of Selection
Rating
0
1
2
3
4
5
1
2
3
4
5
6
7
8
9
10


Input #4 - Overall Size of Contract
Small
Average
Large
Frequency of Selection
Rating
0
1
2
3
4
5
1
2
3
4
5
6
7
8
9
10

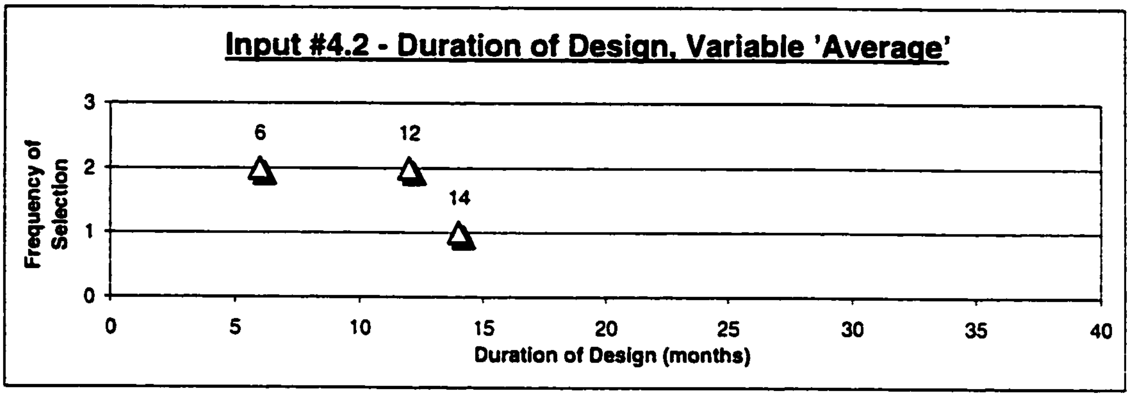

Input #4.1 - Cost of Design, Variable 'Small'
Frequency of Selection
$60,000
$200,000
0
1
2
$0
$200,000
$400,000
$600,000
$800,000
$1,000,000
Cost of Design


Input #4.1 - Cost of Design, Variable 'Average'
Frequency of Selection
$250,000
$500,000
$750,000
$1,000,000
$3,200,000
$8,000,000
0
1
2
3
$0
$2,000,000
$4,000,000
$6,000,000
$8,000,000
$10,000,000
Cost of Design

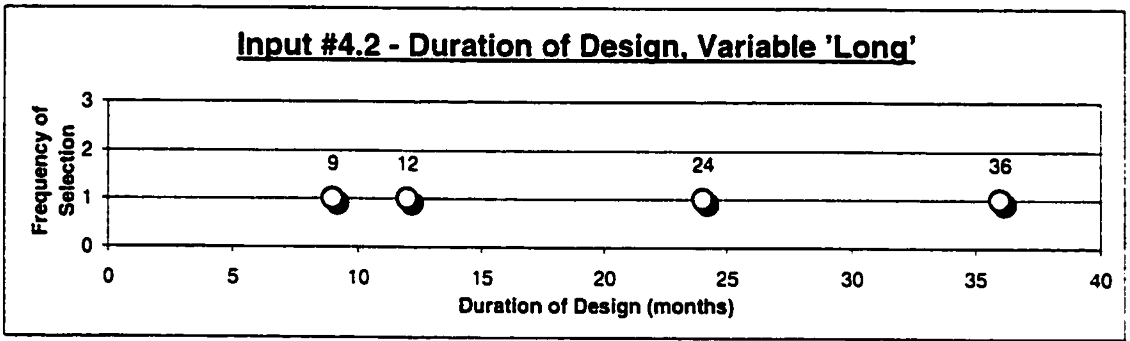

Input #4.1 - Cost of Design, Variable 'Large'
Frequency of Selection
$750,000
$1,000,000
$1,400,000
0
1
2
$0
$1,000,000
$2,000,000
$3,000,000
$4,000,000
$5,000,000
Cost of Design



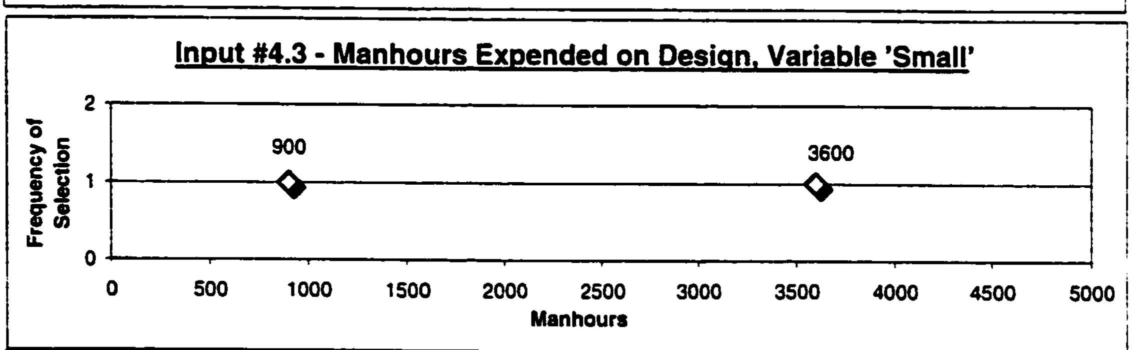
Input #4.2 - Duration of Design, Variable 'Short'
Frequency of Selection
0
1
2
3
12
0 5 10 15 20 25 30 35 40
Duration of Design (months)



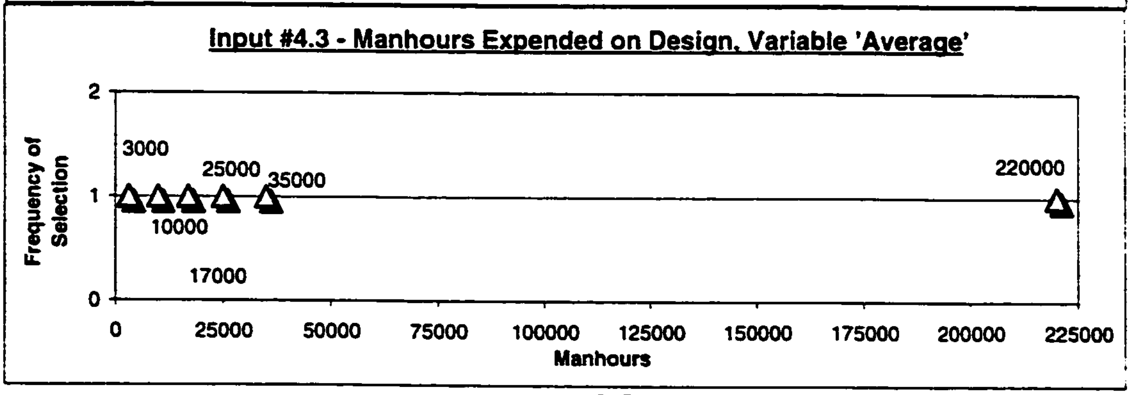
Input #4.2 - Duration of Design, Variable 'Average'
Frequency of Selection
0
1
2
3
6
12
14
0 5 10 15 20 25 30 35 40
Duration of Design (months)



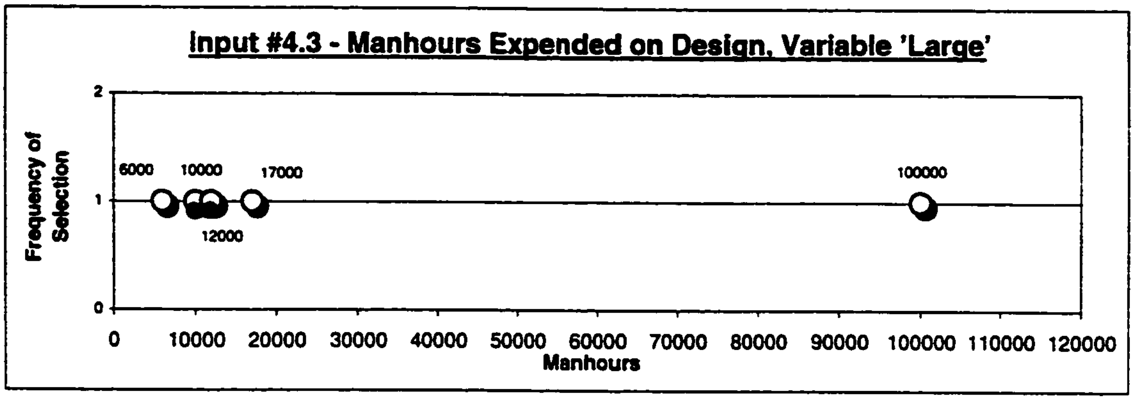
Input #4.2 - Duration of Design, Variable 'Long'
Frequency of Selection
0
1
2
3
9
12
24
36
0 5 10 15 20 25 30 35 40
Duration of Design (months)

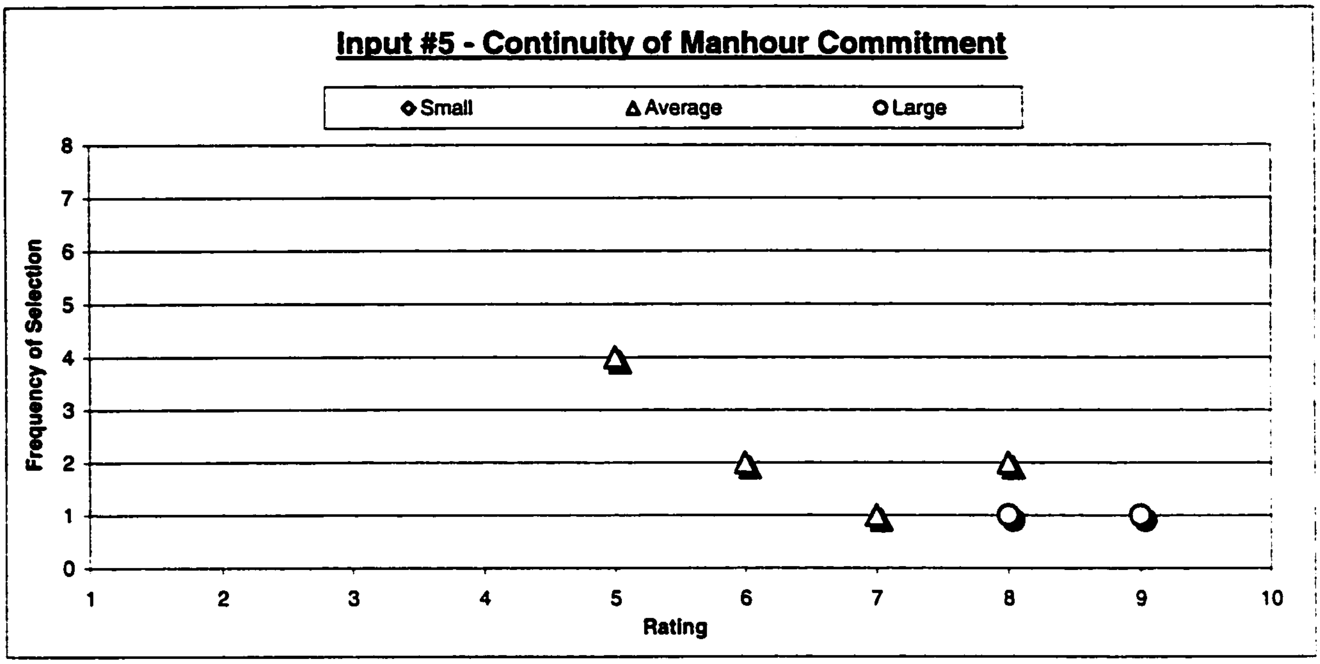


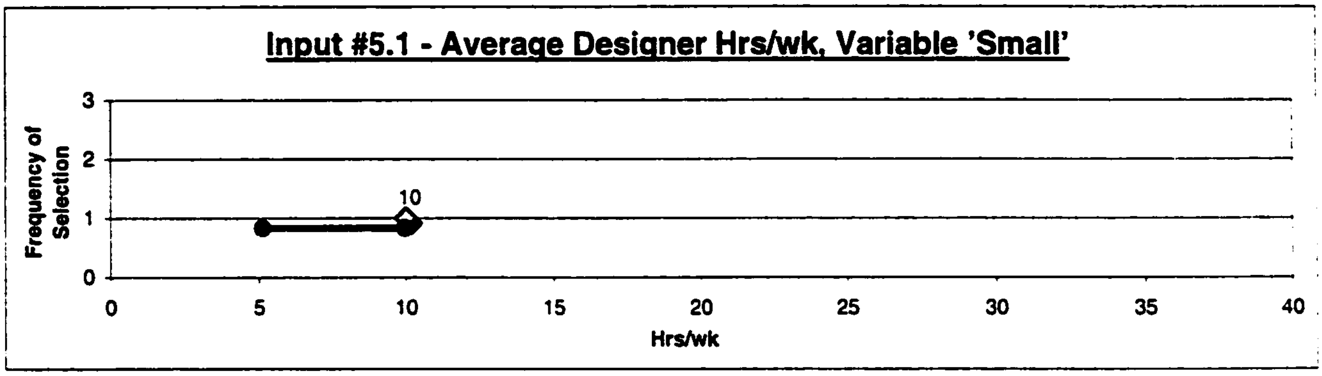
Input #4.3 - Manhours Expended on Design, Variable 'Small'
Frequency of Selection
0
1
2
900
3600
0 500 1000 1500 2000 2500 3000 3500 4000 4500 5000
Manhours

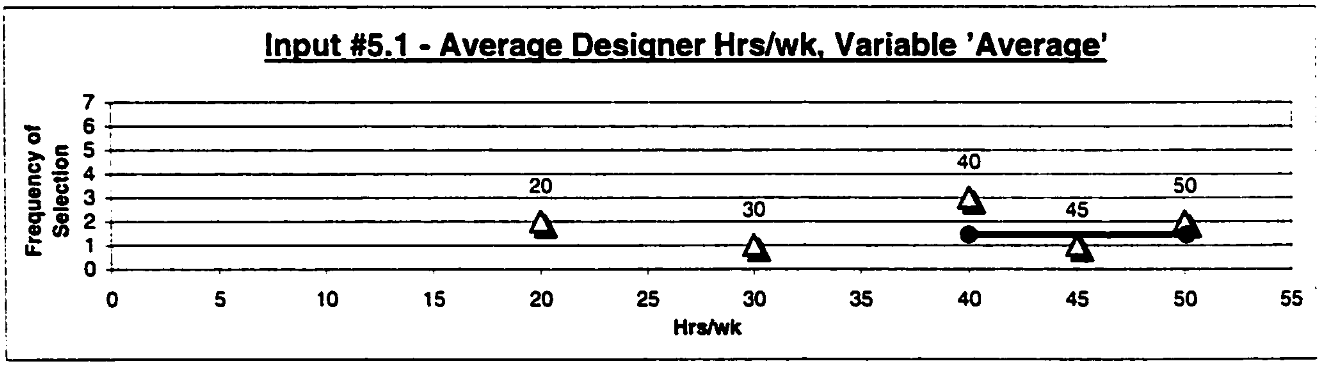


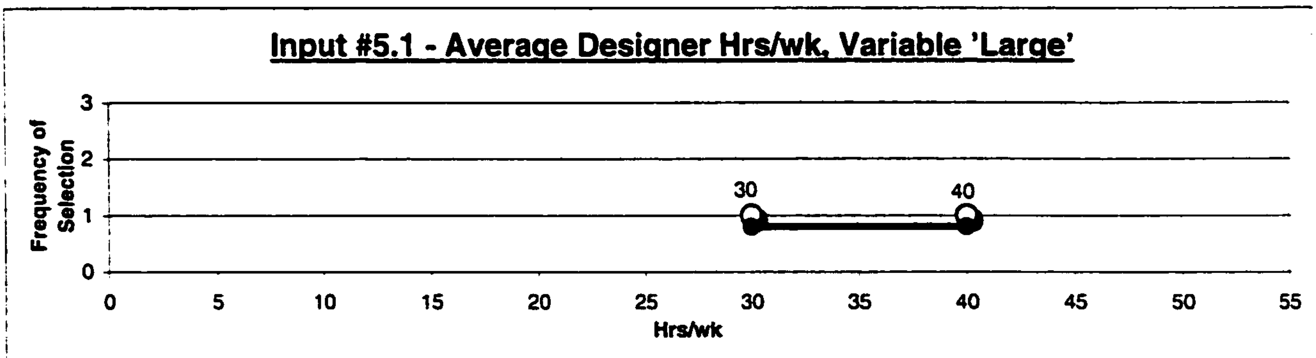
Input #4.3 - Manhours Expended on Design, Variable 'Average'
Frequency of Selection
0
1
2
3000
25000
35000
220000
10000
17000
0 25000 50000 75000 100000 125000 150000 175000 200000 225000
Manhours

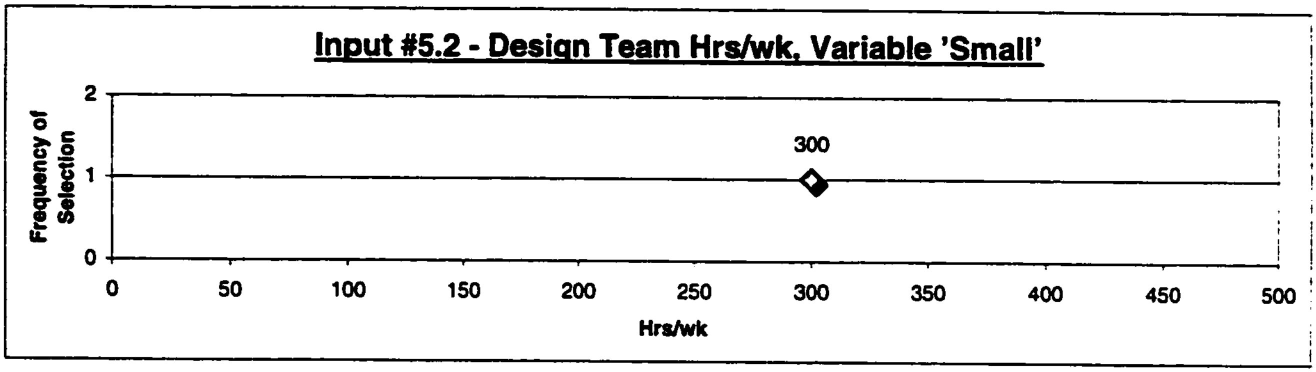

Input #4.3 - Manhours Expended on Design, Variable 'Large'
Frequency of Selection
Manhours
6000
10000
12000
17000
100000
0
1
2
0
10000
20000
30000
40000
50000
60000
70000
80000
90000
100000
110000
120000

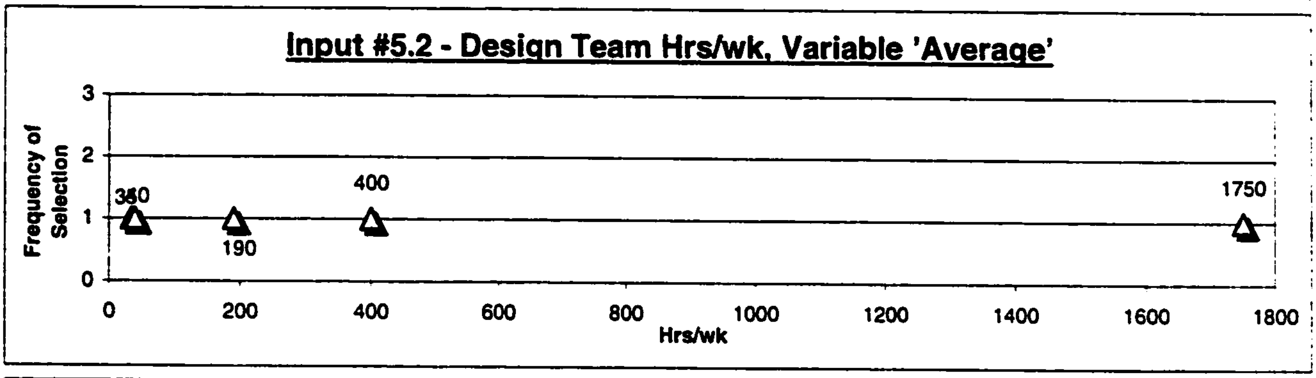

Input #5 - Continuity of Manhour Commitment
Small
Average
Large
Frequency of Selection
8
7
6
5
4
3
2
1
0
1
2
3
4
5
6
7
8
9
10
Rating

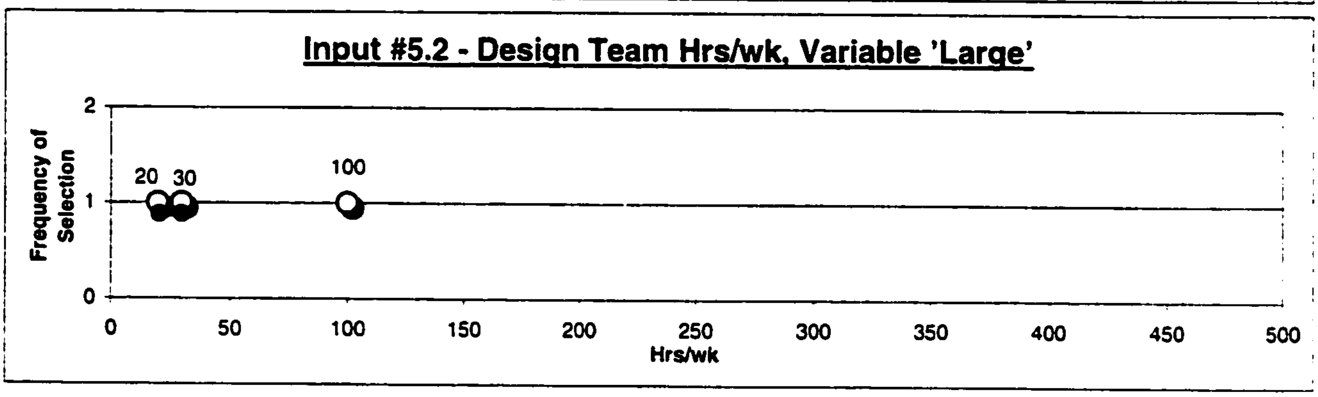

Input #5.1 - Average Designer Hrs/wk, Variable 'Small'
Frequency of Selection
3
2
1
0
10
0
5
10
15
20
25
30
35
40
Hrs/wk


Input #5.1 - Average Designer Hrs/wk, Variable 'Average'
Frequency of Selection
7
6
5
4
3
2
1
0
20
30
40
45
50
0
5
10
15
20
25
30
35
40
45
50
55
Hrs/wk

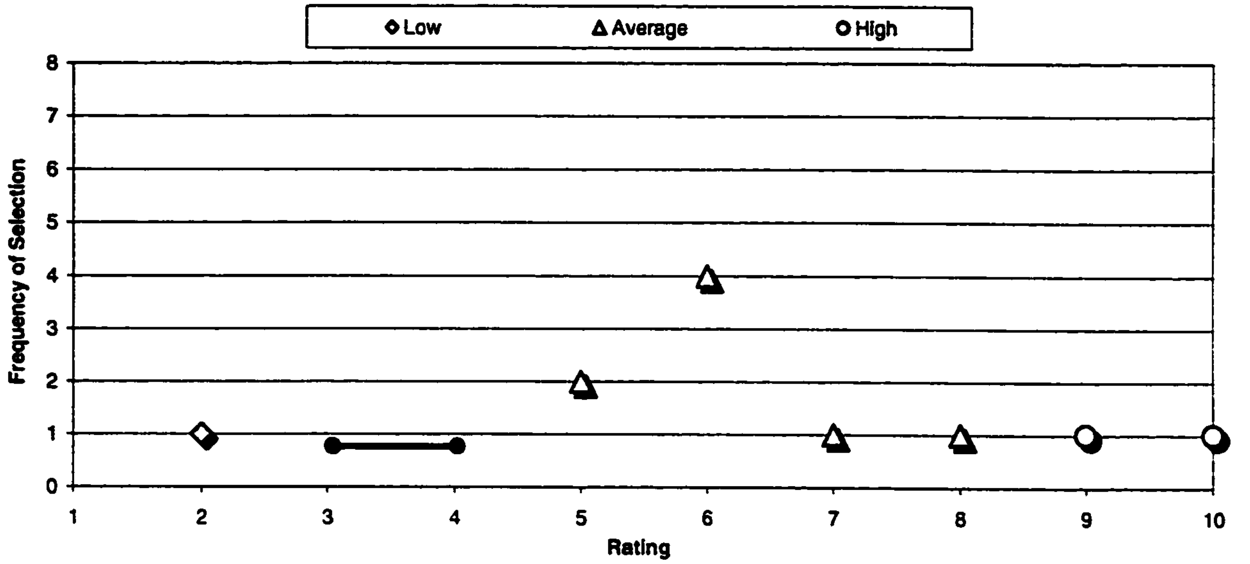

Input #5.1 - Average Designer Hrs/wk, Variable 'Large'
Frequency of Selection
3
2
1
0
30
40
0
5
10
15
20
25
30
35
40
45
50
55
Hrs/wk


Input #5.2 - Design Team Hrs/wk, Variable 'Small'
Frequency of Selection
Hrs/wk
300
0
1
2
0
50
100
150
200
250
300
350
400
450
500

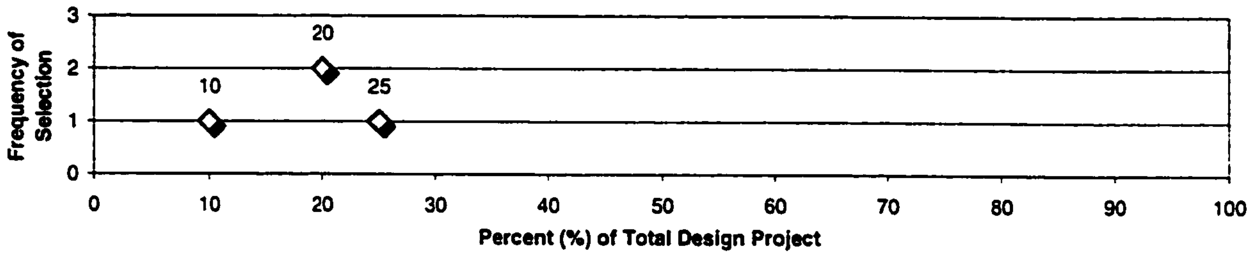

Input #5.2 - Design Team Hrs/wk, Variable 'Average'
Frequency of Selection
Hrs/wk
350
190
400
1750
0
1
2
3
0
200
400
600
800
1000
1200
1400
1600
1800


Input #5.2 - Design Team Hrs/wk, Variable 'Large'
Frequency of Selection
Hrs/wk
20
30
100
0
1
2
0
50
100
150
200
250
300
350
400
450
500

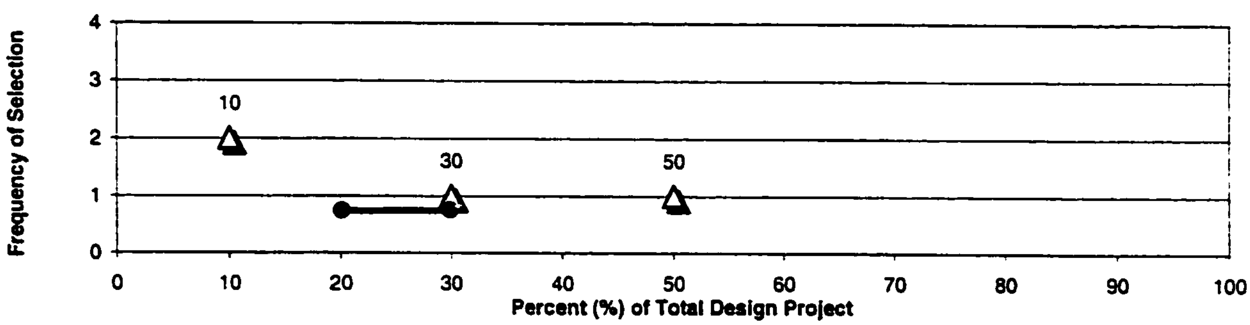
Input #7 - Complexity of Function of the Project



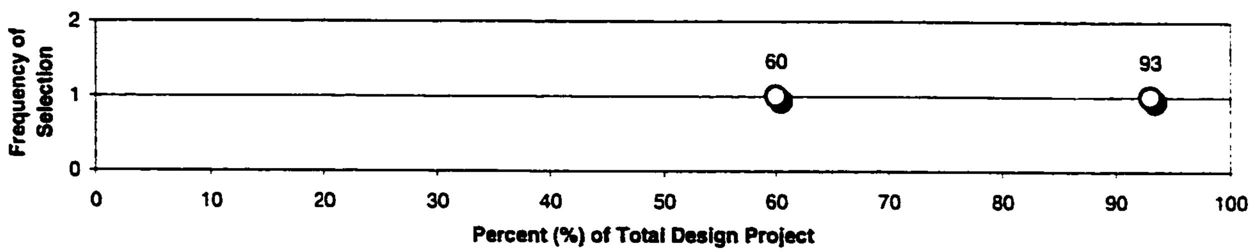
Input #7.1 - Repetition of Design, Variable 'Small'

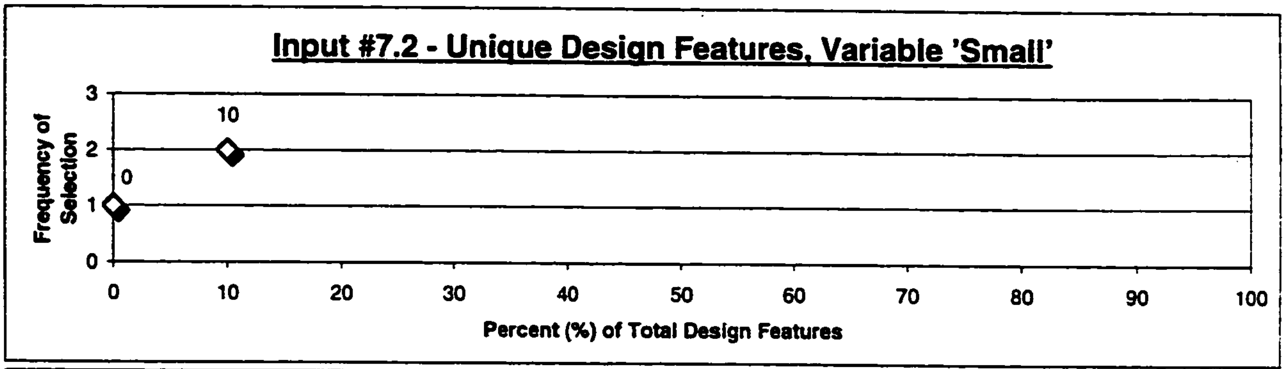


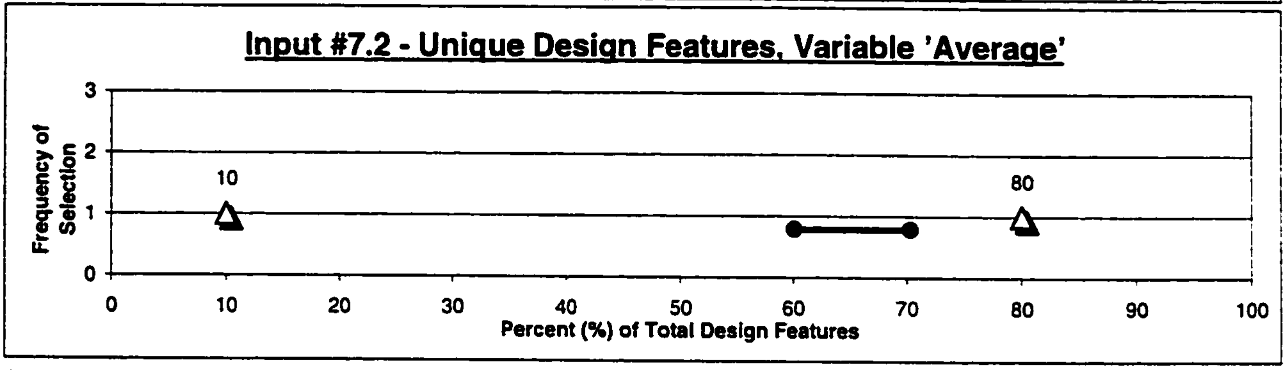
Input #7.1 - Repetition of Design, Variable 'Average'

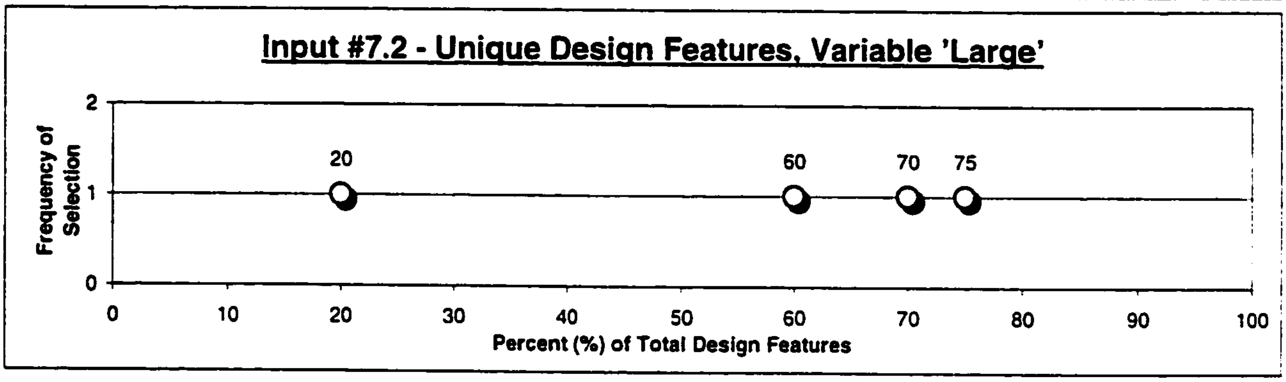


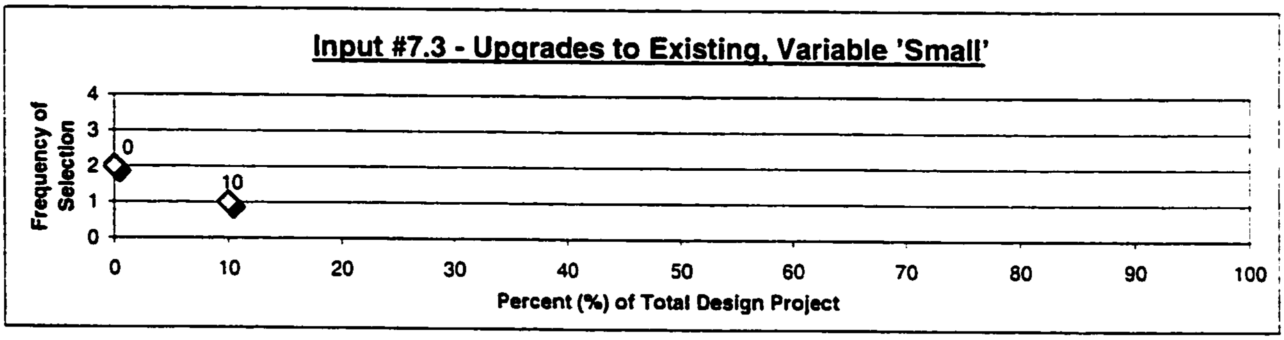
Input #7.1 - Repetition of Design, Variable 'Large'

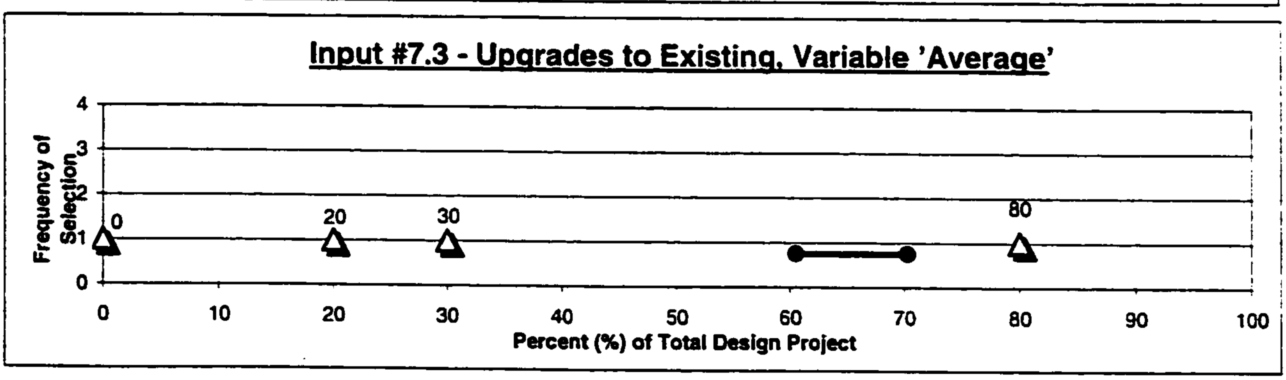


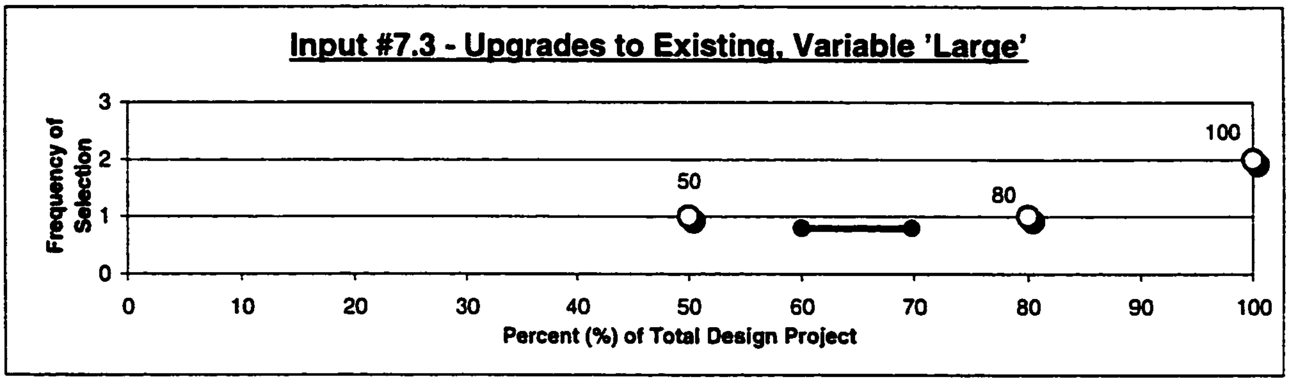

Input #7.2 - Unique Design Features, Variable 'Small'
Frequency of Selection
Percent (%) of Total Design Features
0
10

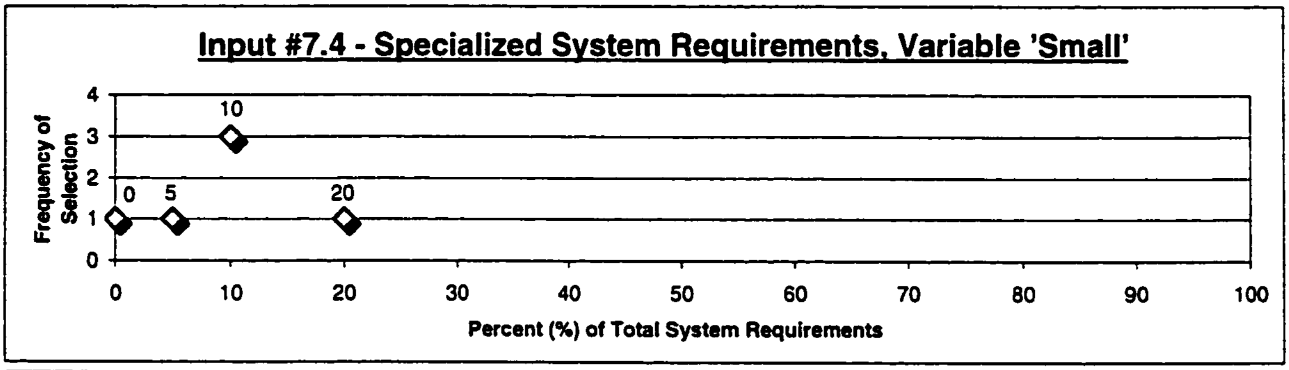

Input #7.2 - Unique Design Features, Variable 'Average'
Frequency of Selection
Percent (%) of Total Design Features
10
80

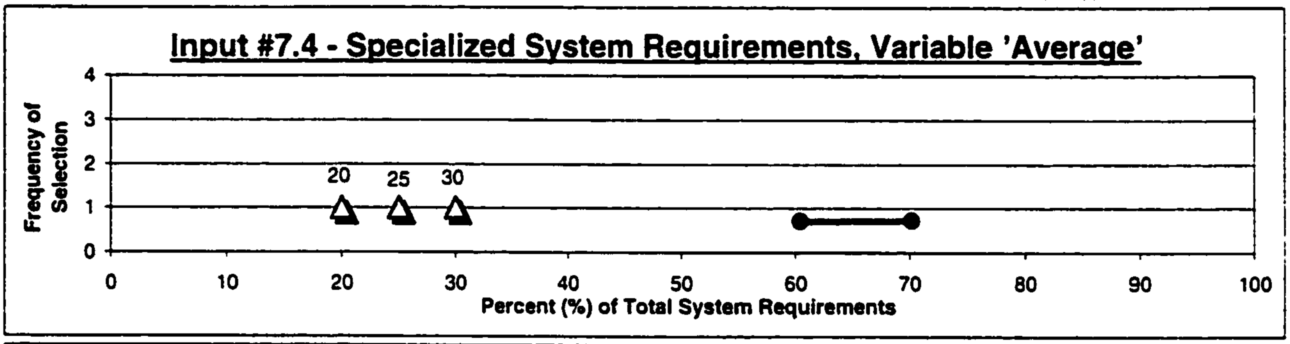

Input #7.2 - Unique Design Features, Variable 'Large'
Frequency of Selection
Percent (%) of Total Design Features
20
60
70
75

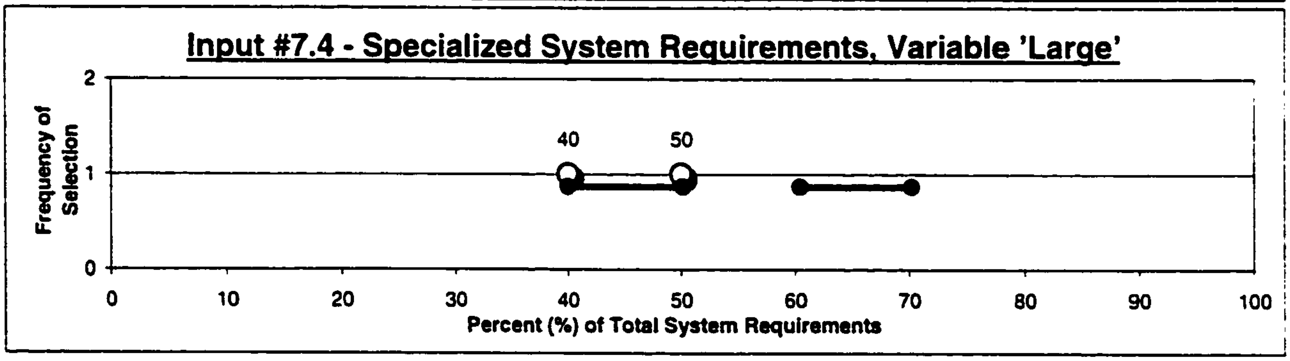

Input #7.3 - Upgrades to Existing, Variable 'Small'
Frequency of Selection
Percent (%) of Total Design Project
0
10

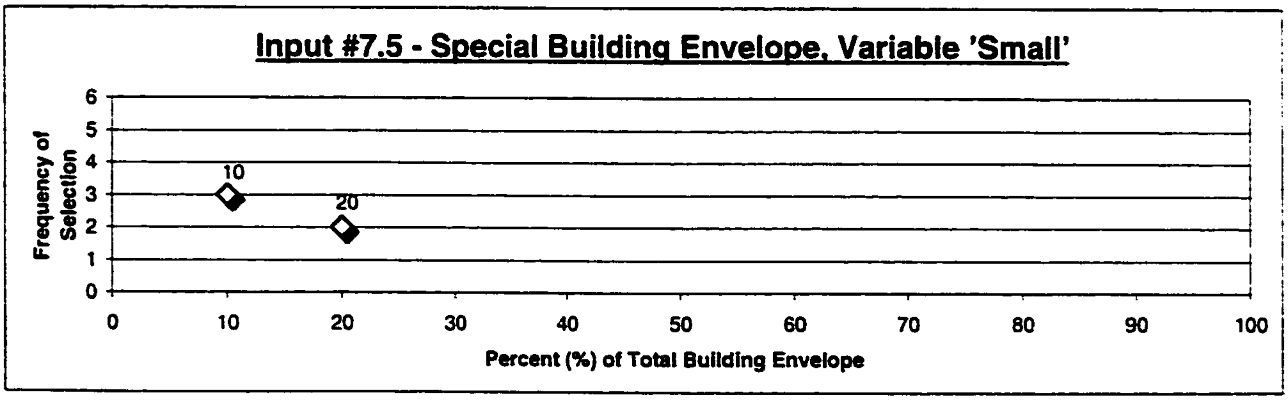

Input #7.3 - Upgrades to Existing, Variable 'Average'
Frequency of Selection
Percent (%) of Total Design Project
0
20
30
80

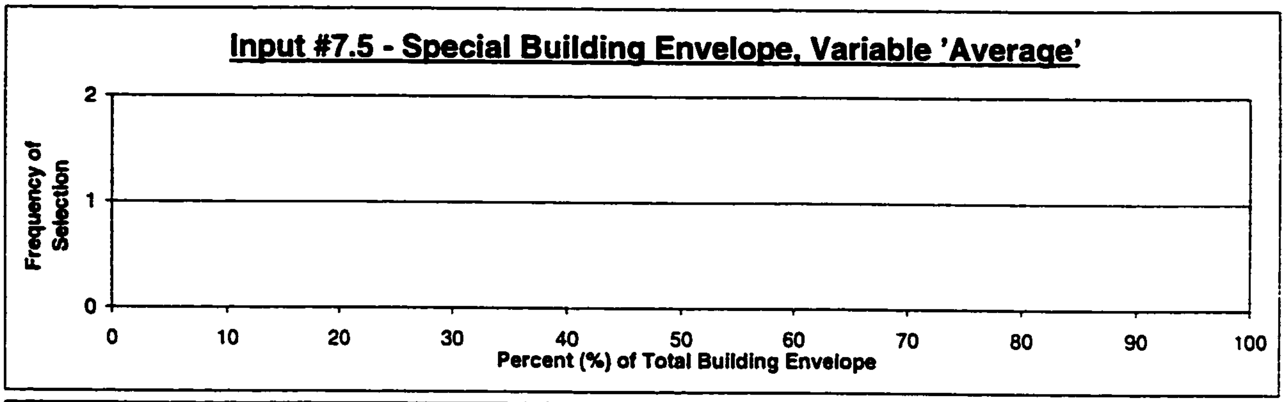

Input #7.3 - Upgrades to Existing, Variable 'Large'
Frequency of Selection
0
1
2
3
50
80
100
0
10
20
30
40
50
60
70
80
90
100
Percent (%) of Total Design Project


Input #7.4 - Specialized System Requirements, Variable 'Small'
Frequency of Selection
0
1
2
3
4
0
5
10
20
0
10
20
30
40
50
60
70
80
90
100
Percent (%) of Total System Requirements

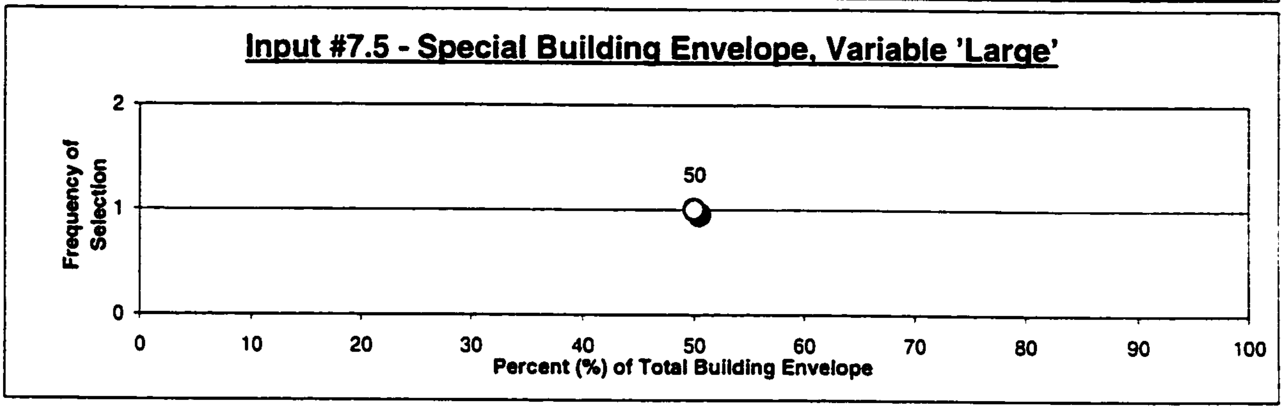

Input #7.4 - Specialized System Requirements, Variable 'Average'
Frequency of Selection
0
1
2
3
4
20
25
30
0
10
20
30
40
50
60
70
80
90
100
Percent (%) of Total System Requirements


Input #7.4 - Specialized System Requirements, Variable 'Large'
Frequency of Selection
0
1
2
40
50
0
10
20
30
40
50
60
70
80
90
100
Percent (%) of Total System Requirements

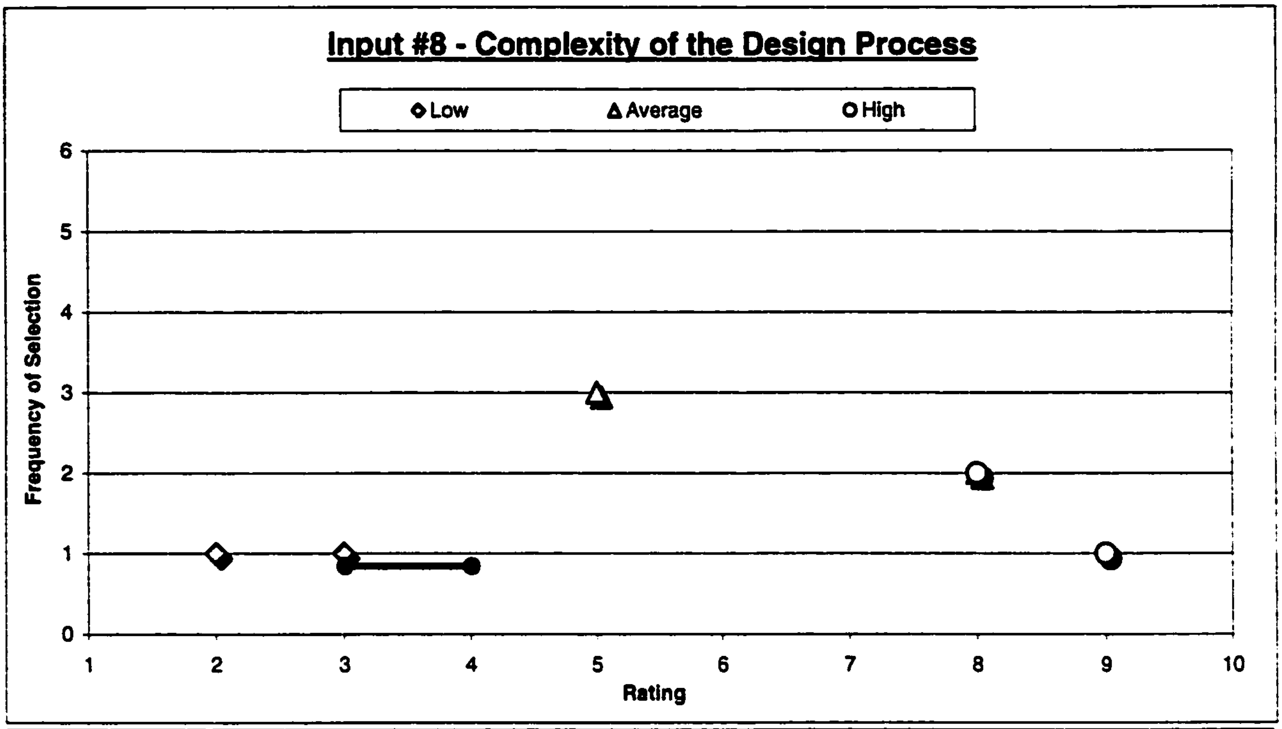

Input #7.5 - Special Building Envelope, Variable 'Small'
Frequency of Selection
0
1
2
3
4
5
6
10
20
0
10
20
30
40
50
60
70
80
90
100
Percent (%) of Total Building Envelope



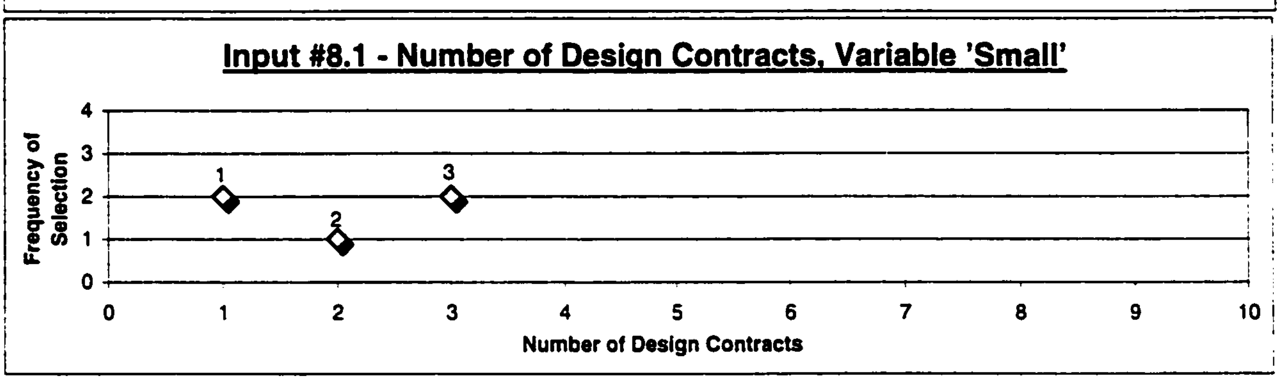
Input #7.5 - Special Building Envelope, Variable 'Average'
Frequency of Selection
2
1
0
0 10 20 30 40 50 60 70 80 90 100
Percent (%) of Total Building Envelope



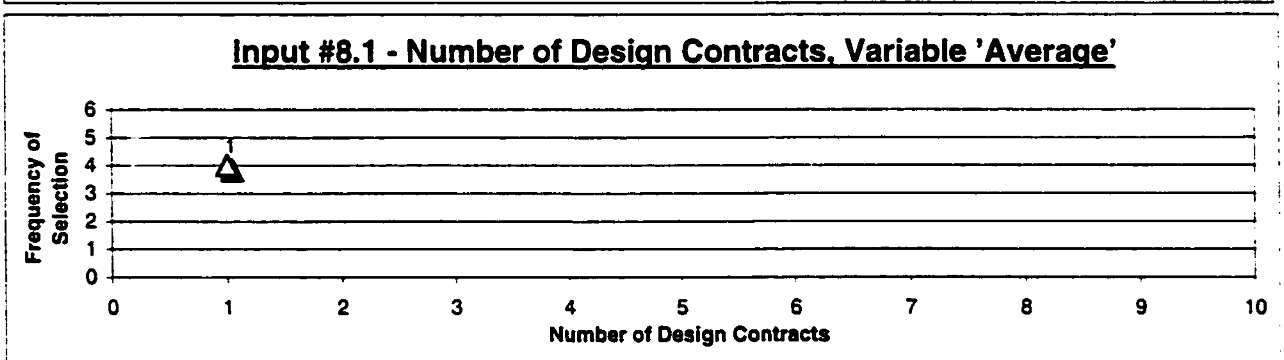
Input #7.5 - Special Building Envelope, Variable 'Large'
Frequency of Selection
2
1
0
50
0 10 20 30 40 50 60 70 80 90 100
Percent (%) of Total Building Envelope



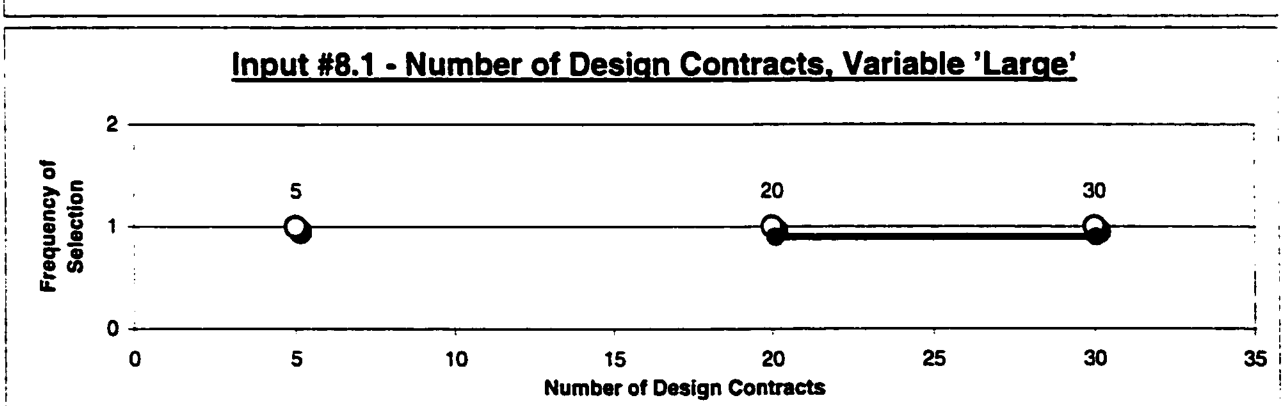
Input #8 - Complexity of the Design Process
Low
Average
High
Frequency of Selection
Rating



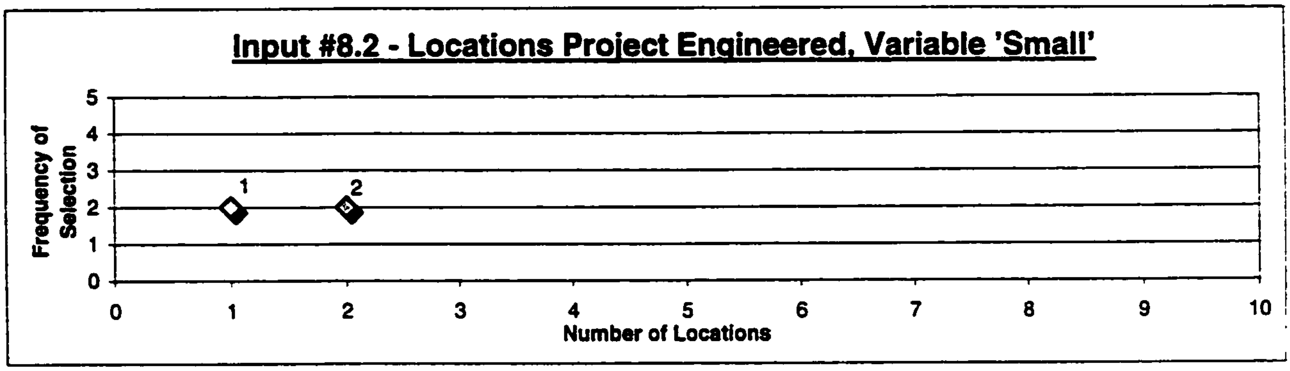
Input #8.1 - Number of Design Contracts, Variable 'Small'
Frequency of Selection
Number of Design Contracts

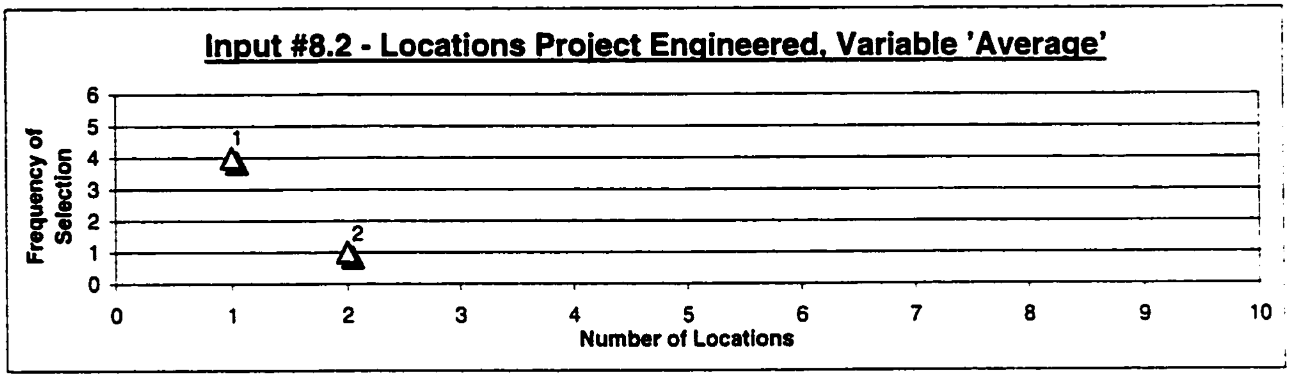


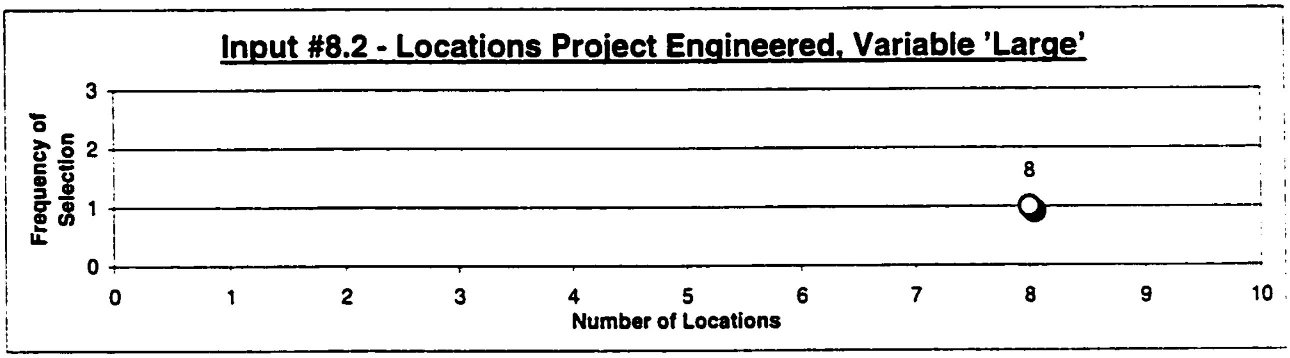
Input #8.1 - Number of Design Contracts, Variable 'Average'
Frequency of Selection
Number of Design Contracts

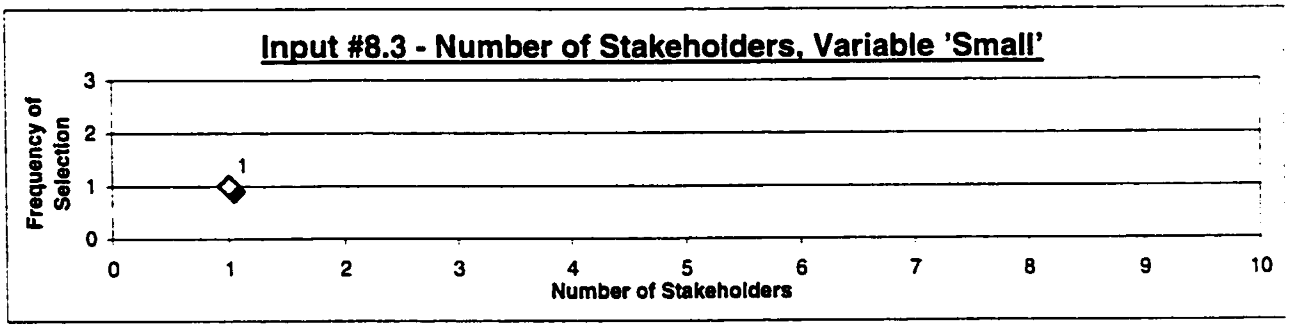


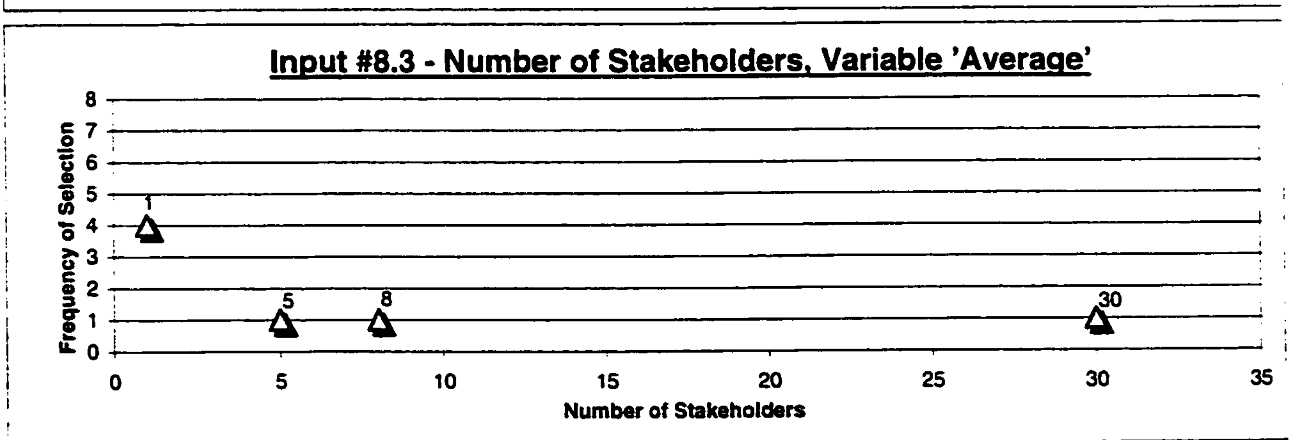
Input #8.1 - Number of Design Contracts, Variable 'Large'
Frequency of Selection
Number of Design Contracts

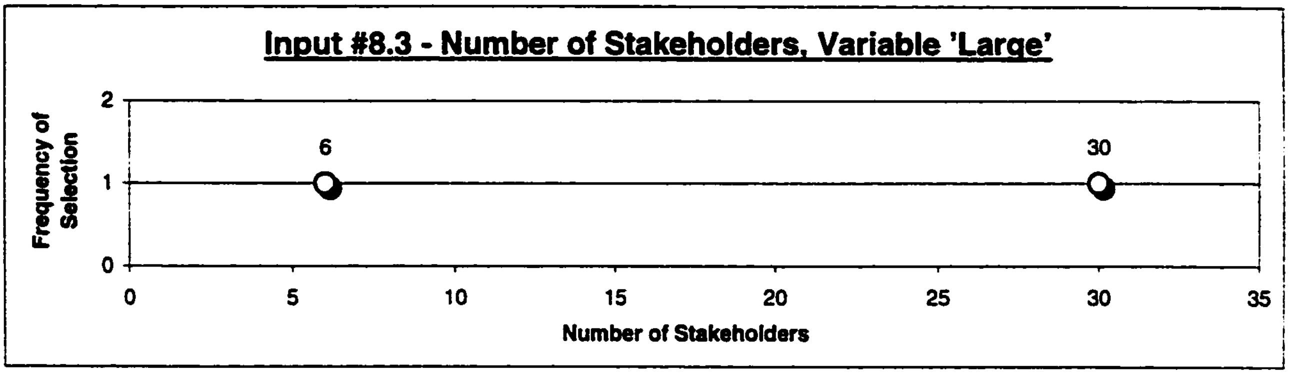

Input #8.2 - Locations Project Engineered, Variable 'Small'
Frequency of Selection
Number of Locations
0 1 2 3 4 5 6 7 8 9 10
0 1 2 3 4 5
1
2

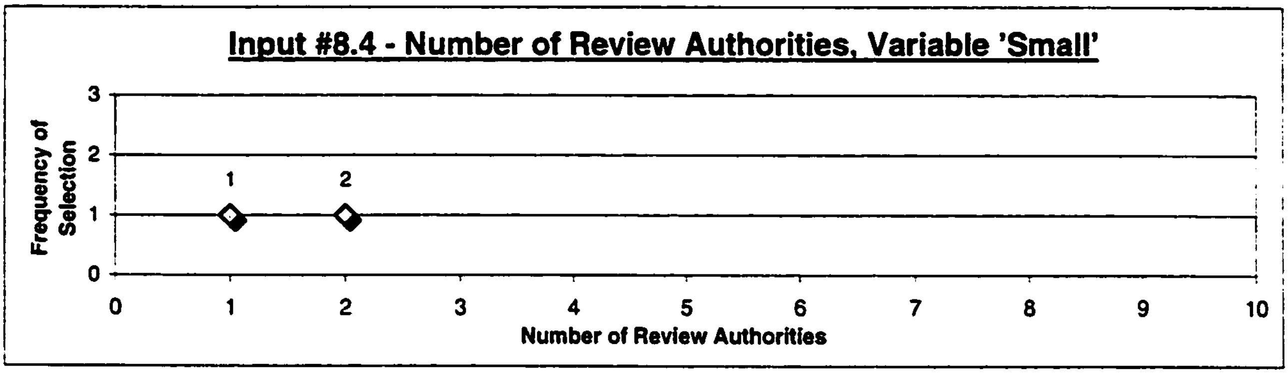

Input #8.2 - Locations Project Engineered, Variable 'Average'
Frequency of Selection
Number of Locations
0 1 2 3 4 5 6 7 8 9 10
0 1 2 3 4 5 6
1
2

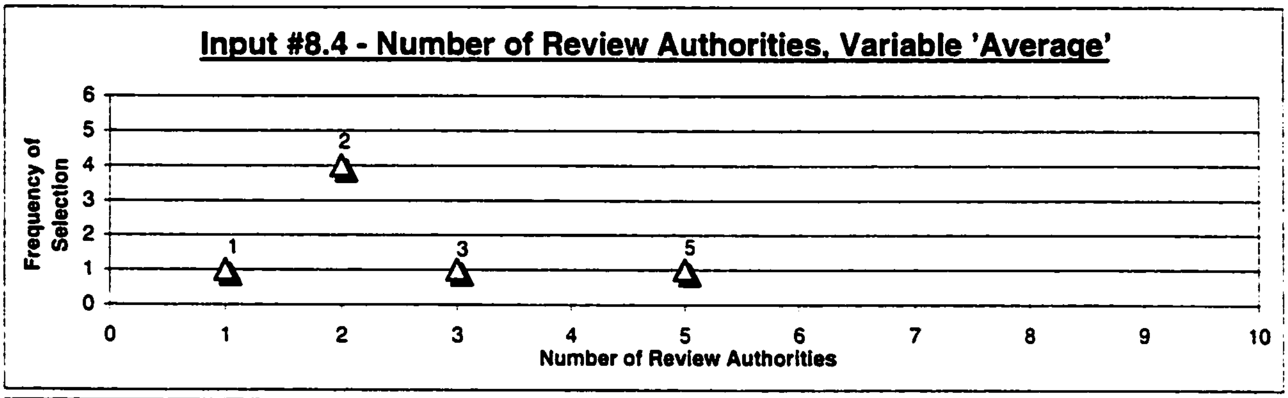

Input #8.2 - Locations Project Engineered, Variable 'Large'
Frequency of Selection
Number of Locations
0 1 2 3 4 5 6 7 8 9 10
0 1 2 3
8

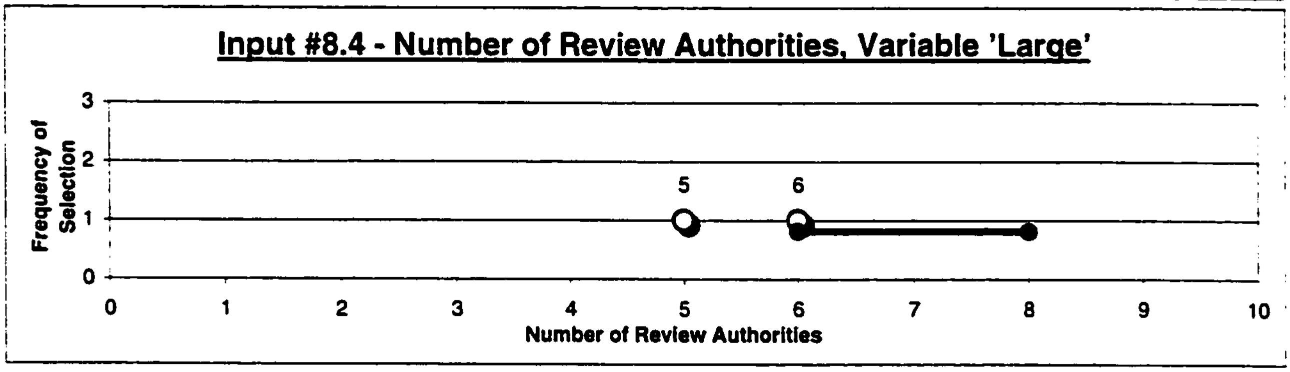

Input #8.3 - Number of Stakeholders, Variable 'Small'
Frequency of Selection
Number of Stakeholders
0 1 2 3 4 5 6 7 8 9 10
0 1 2 3
1

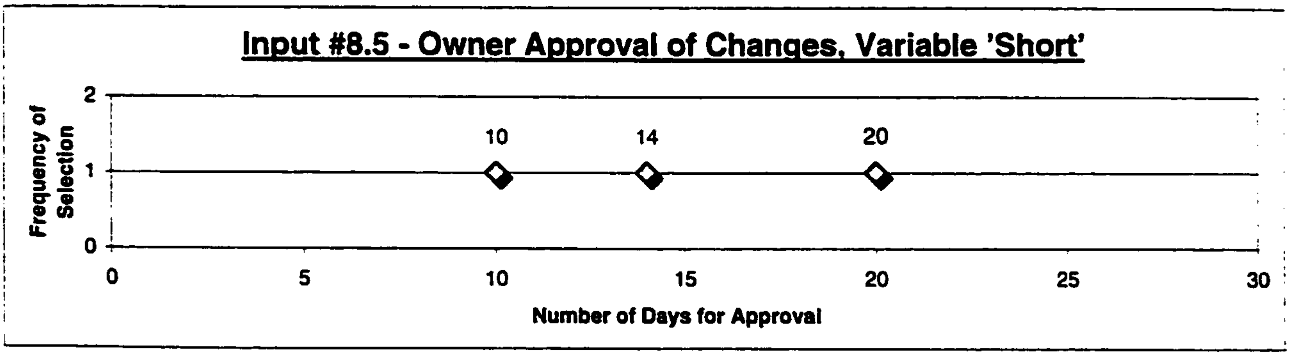

Input #8.3 - Number of Stakeholders, Variable 'Average'
Frequency of Selection
Number of Stakeholders
0 5 10 15 20 25 30 35
0 1 2 3 4 5 6 7 8
1
5
8
30


Input #8.3 - Number of Stakeholders, Variable 'Large'
Frequency of Selection
Number of Stakeholders
6
30

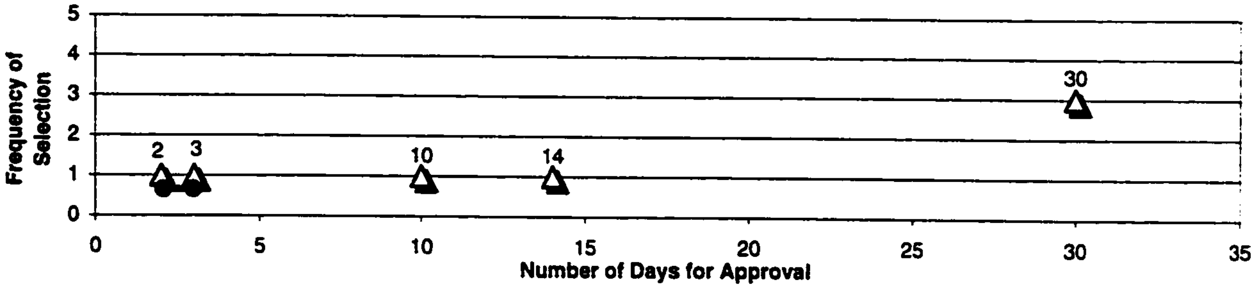

Input #8.4 - Number of Review Authorities, Variable 'Small'
Frequency of Selection
Number of Review Authorities
1
2


Input #8.4 - Number of Review Authorities, Variable 'Average'
Frequency of Selection
Number of Review Authorities
1
2
3
5

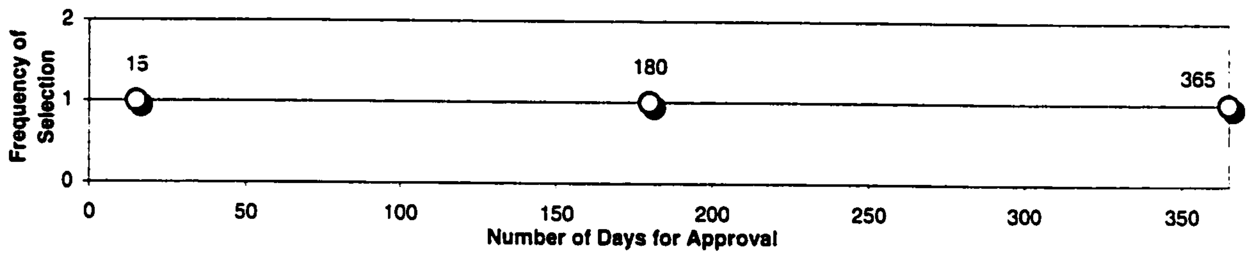

Input #8.4 - Number of Review Authorities, Variable 'Large'
Frequency of Selection
Number of Review Authorities
5
6


Input #8.5 - Owner Approval of Changes, Variable 'Short'
Frequency of Selection
Number of Days for Approval
10
14
20

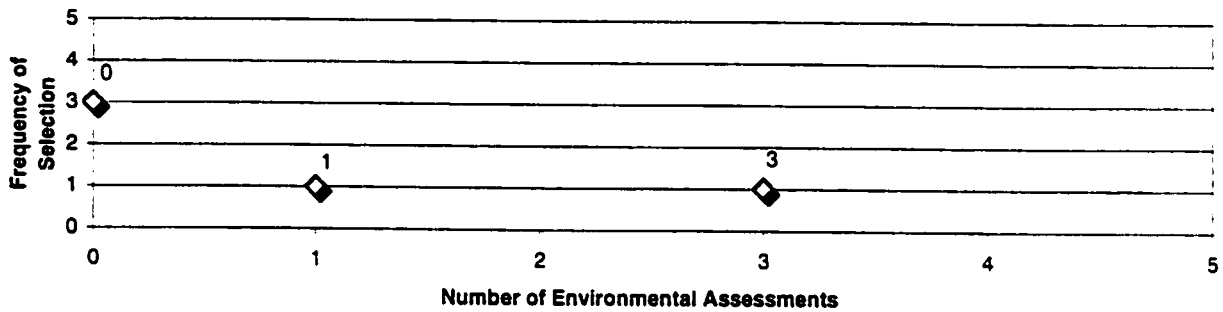
Input #8.5 - Owner Approval of Changes, Variable 'Average'



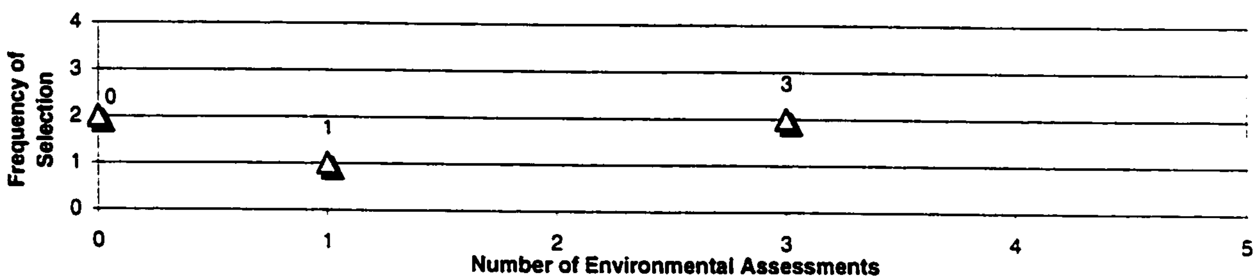
Input #8.5 - Owner Approval of Changes, Variable 'Long'



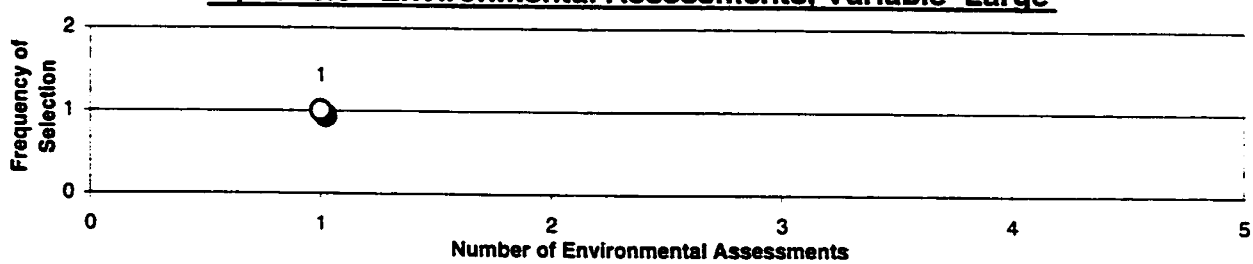
Input #8.6 - Environmental Assessments, Variable 'Small'

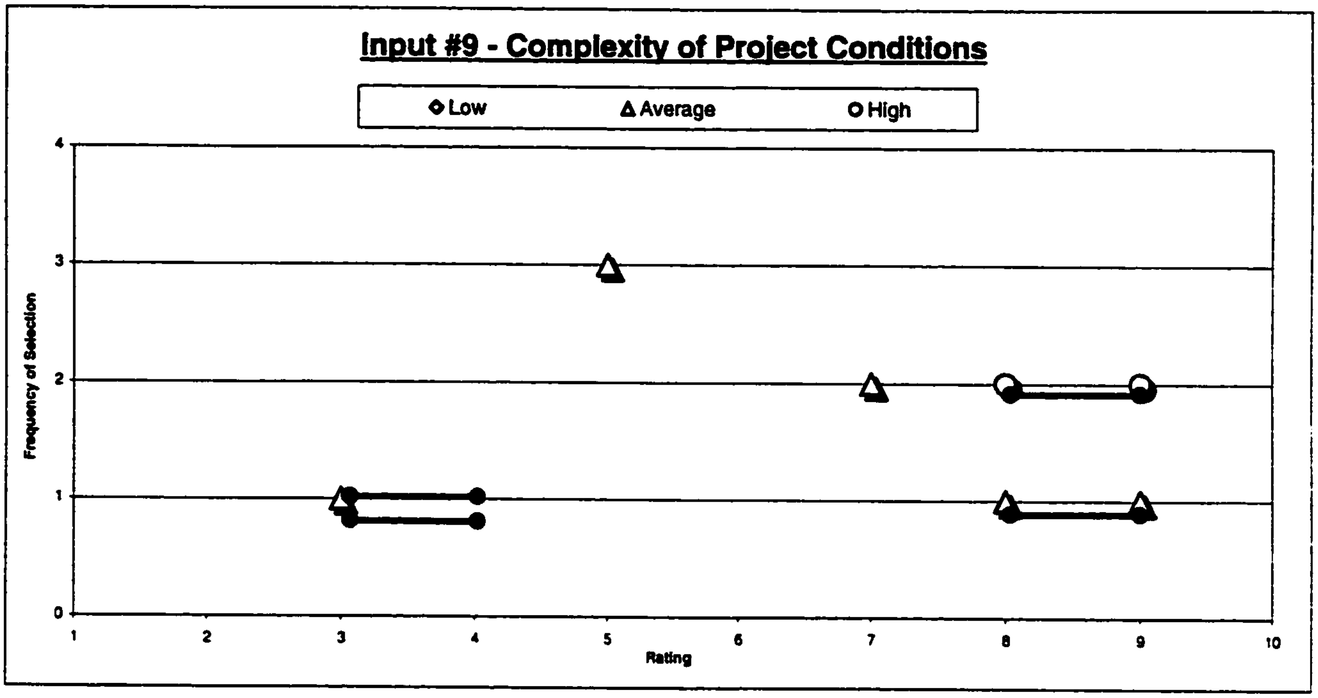


Input #8.6 - Environmental Assessments, Variable 'Average'

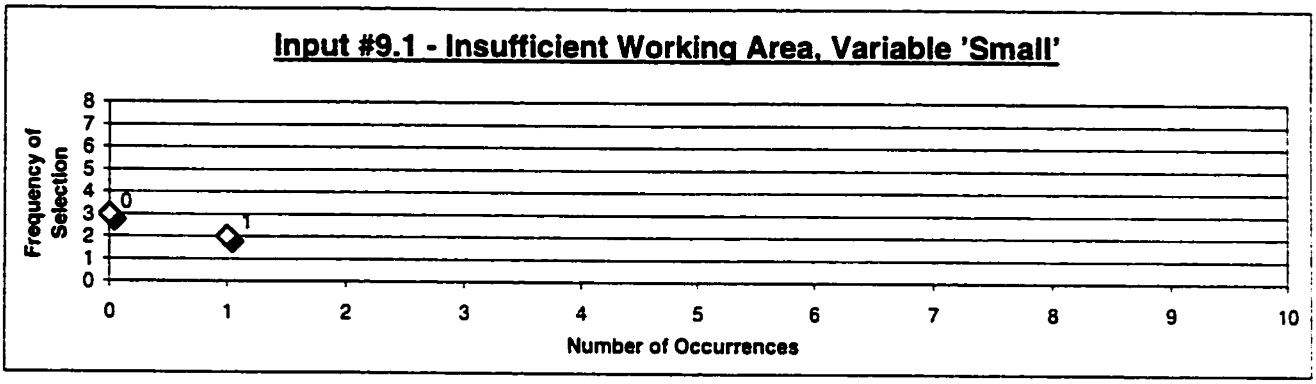


Input #8.6 - Environmental Assessments, Variable 'Large'

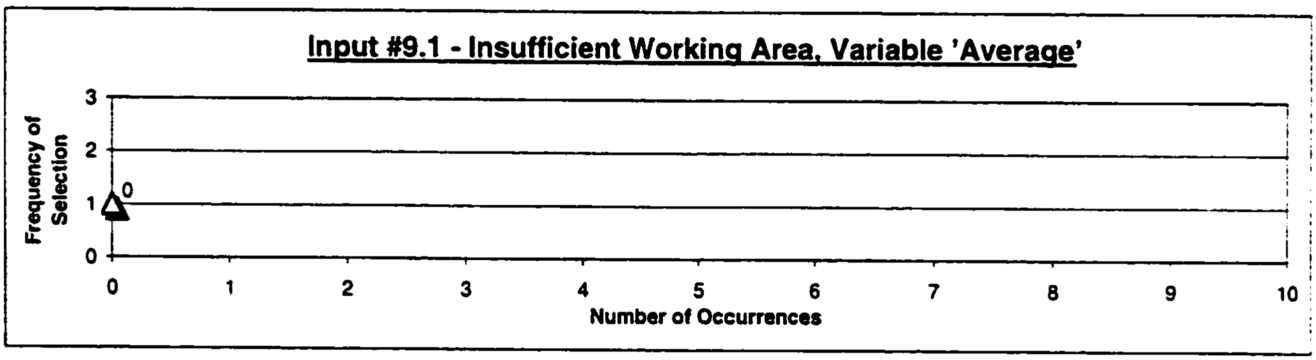




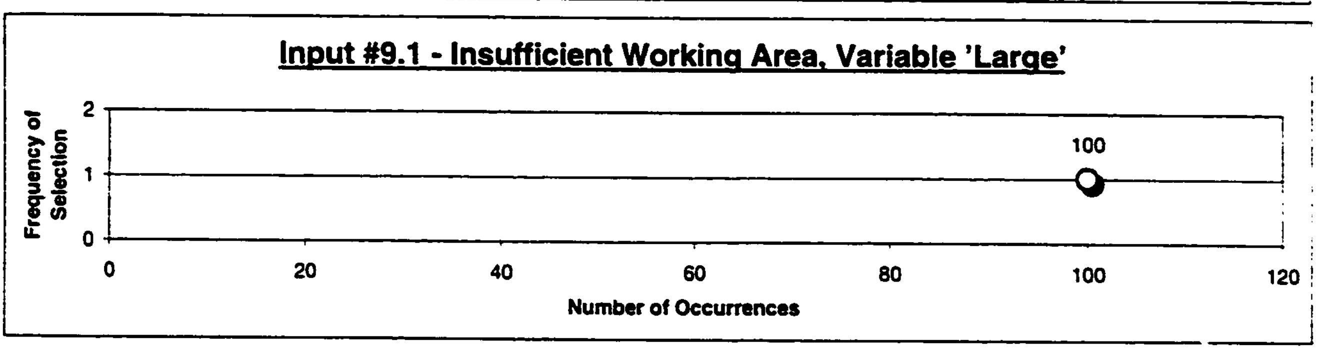
Input #9 - Complexity of Project Conditions
Low
Average
High
Frequency of Selection
Rating



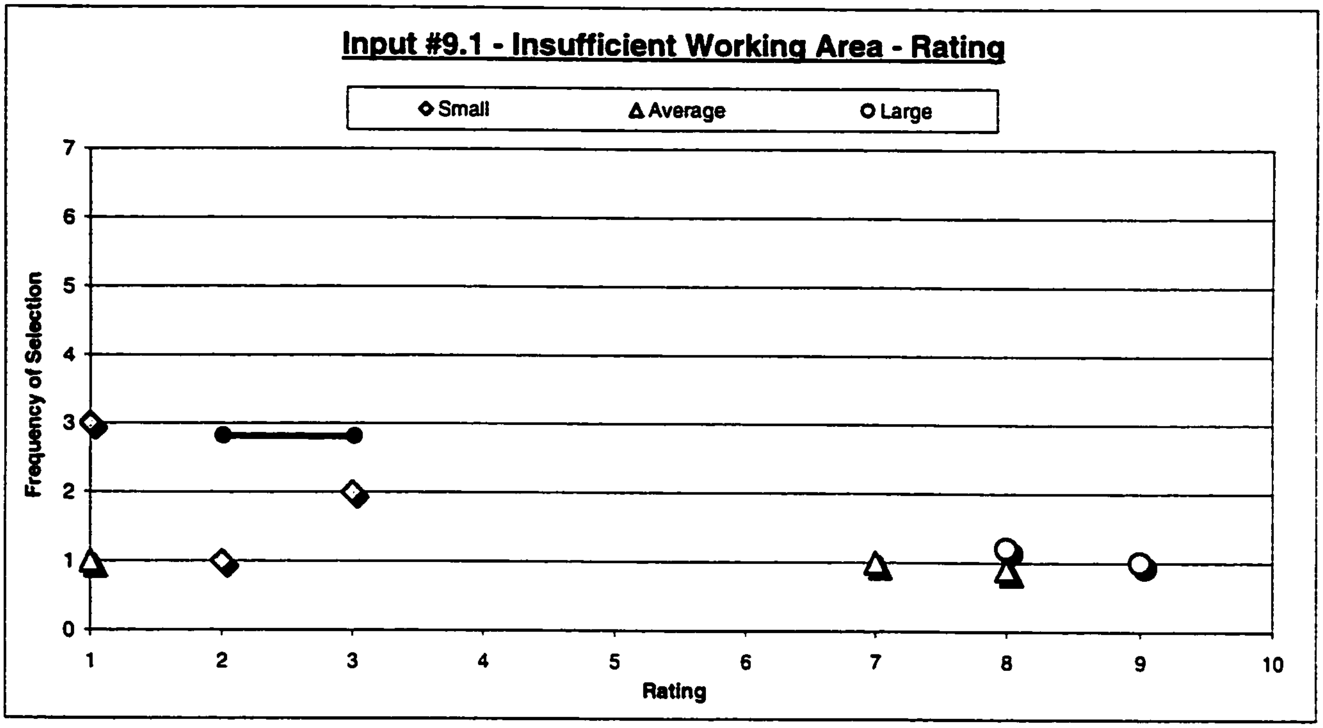
Input #9.1 - Insufficient Working Area, Variable 'Small'
Frequency of Selection
Number of Occurrences

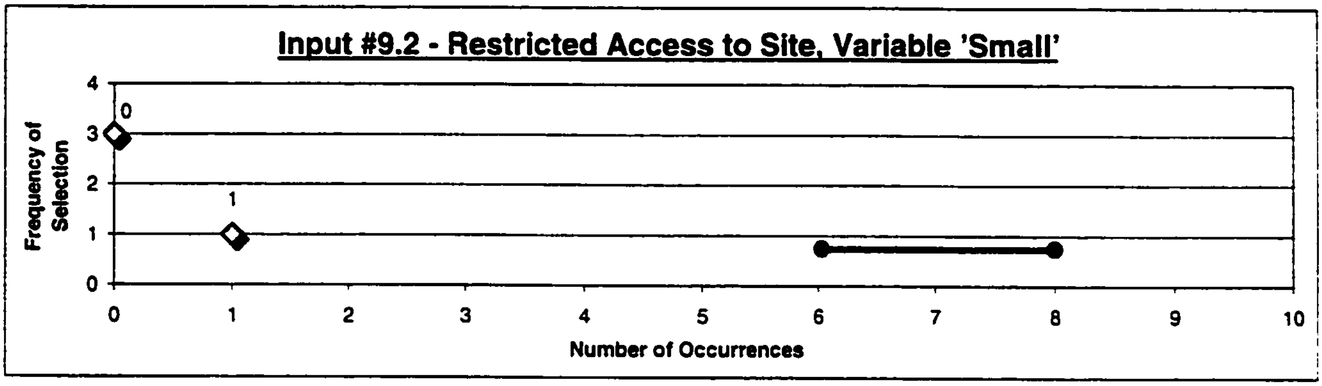


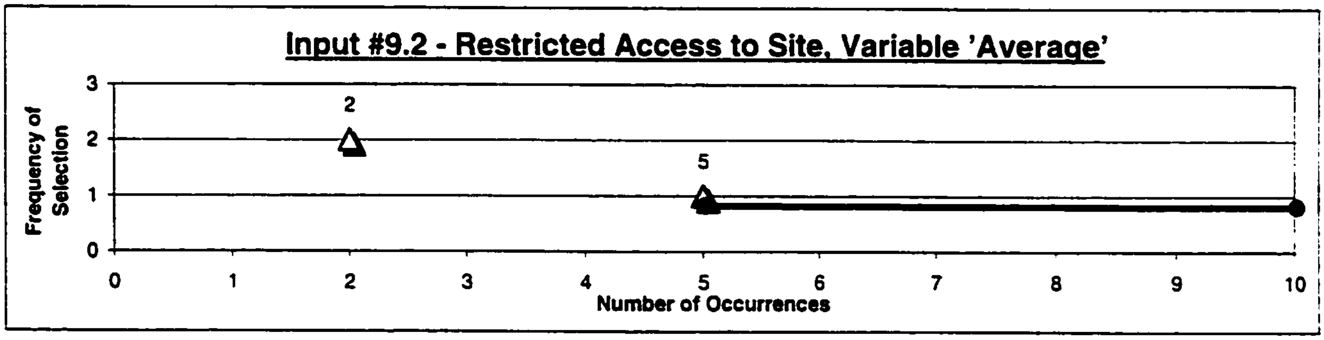
Input #9.1 - Insufficient Working Area, Variable 'Average'
Frequency of Selection
Number of Occurrences

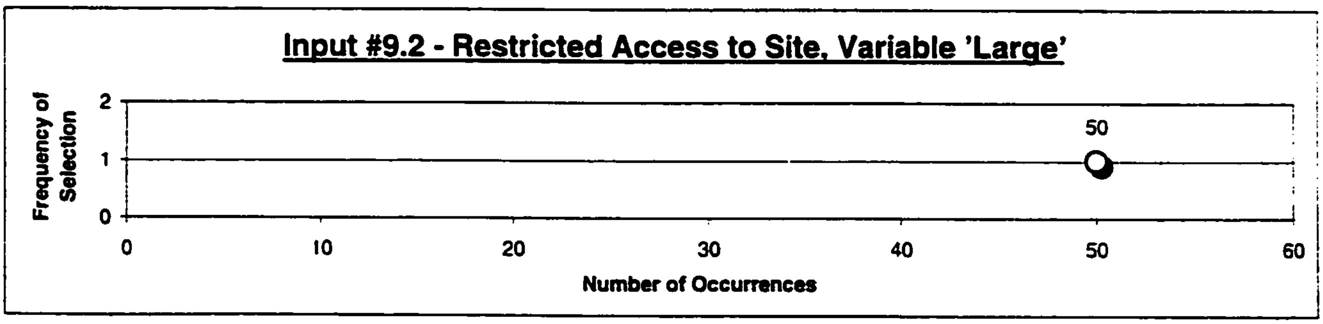


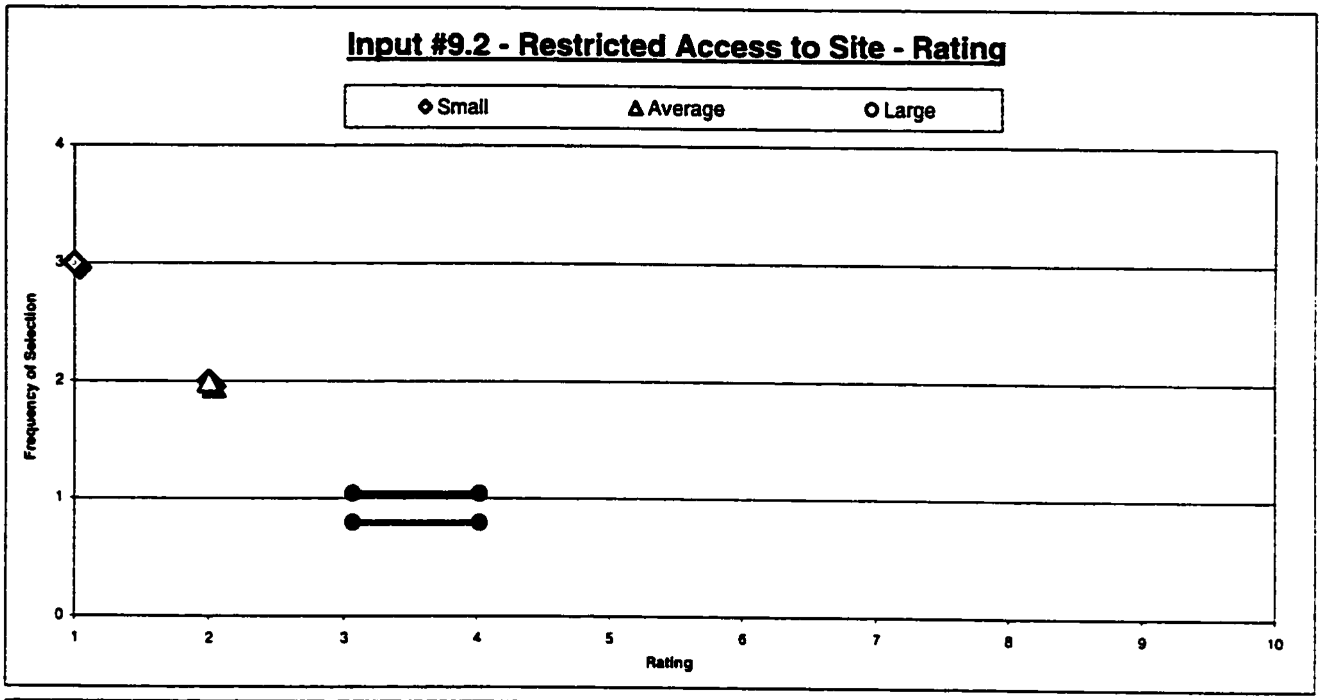
Input #9.1 - Insufficient Working Area, Variable 'Large'
Frequency of Selection
Number of Occurrences

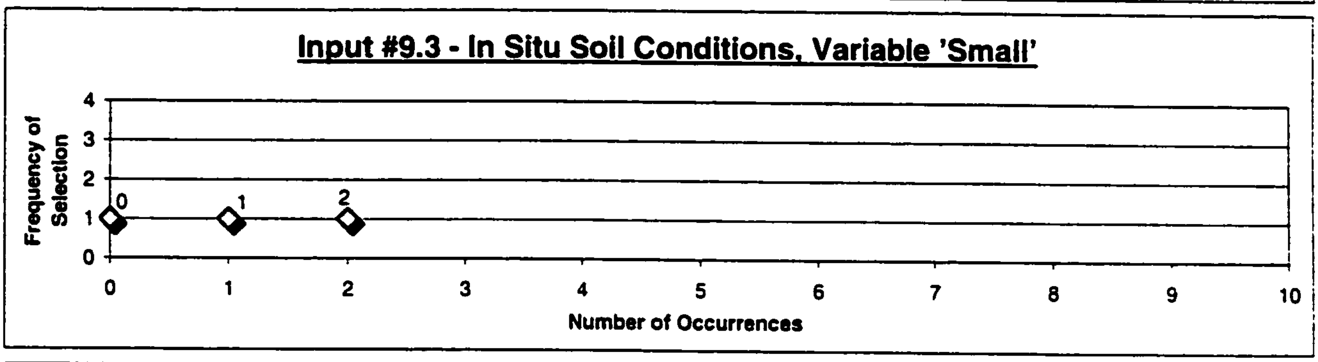

Input #9.1 - Insufficient Working Area - Rating
Small
Average
Large
Frequency of Selection
Rating
0
1
2
3
4
5
6
7
1
2
3
4
5
6
7
8
9
10

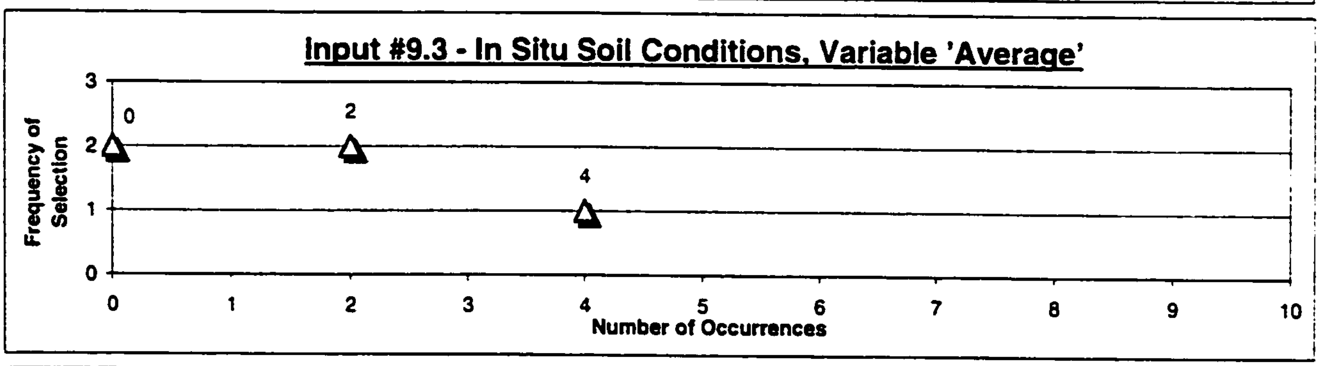

Input #9.2 - Restricted Access to Site, Variable 'Small'
Frequency of Selection
Number of Occurrences
0
1
0
1
2
3
4
0
1
2
3
4
5
6
7
8
9
10

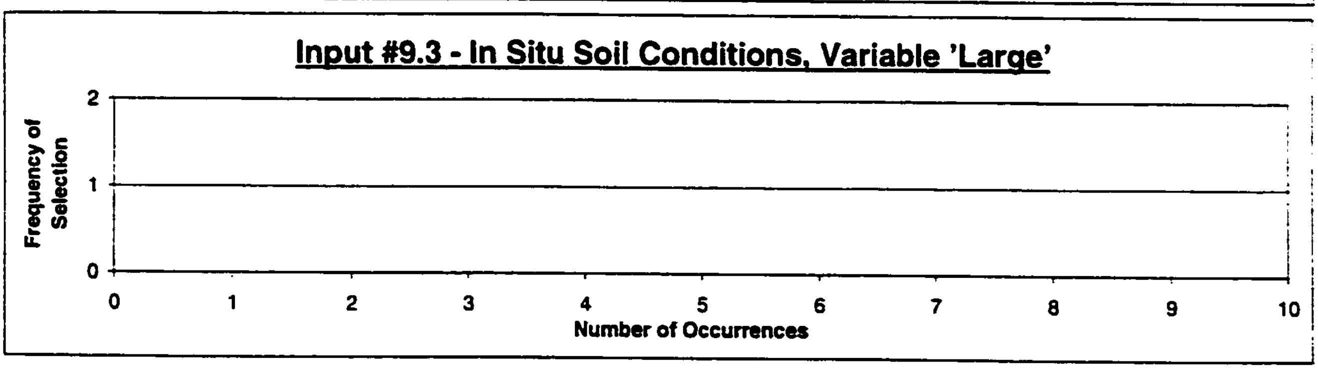

Input #9.2 - Restricted Access to Site, Variable 'Average'
Frequency of Selection
Number of Occurrences
2
5
0
1
2
3
0
1
2
3
4
5
6
7
8
9
10

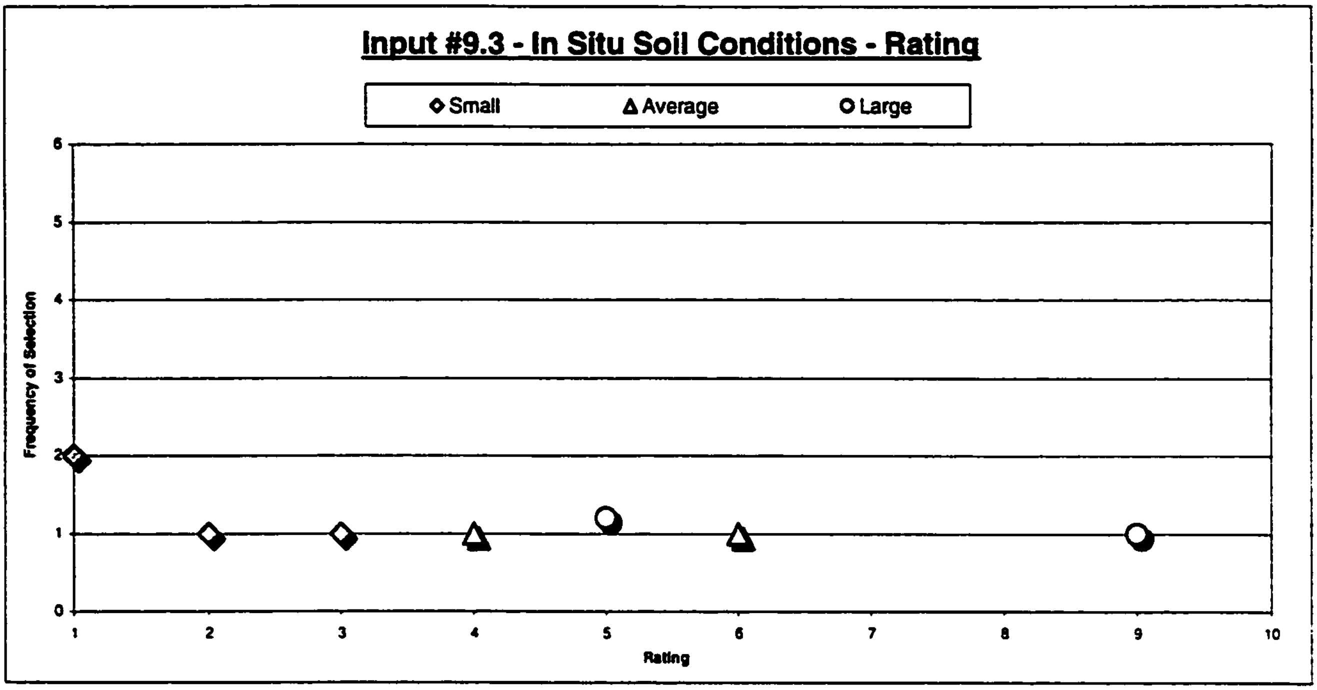

Input #9.2 - Restricted Access to Site, Variable 'Large'
Frequency of Selection
Number of Occurrences
50
0
1
2
0
10
20
30
40
50
60

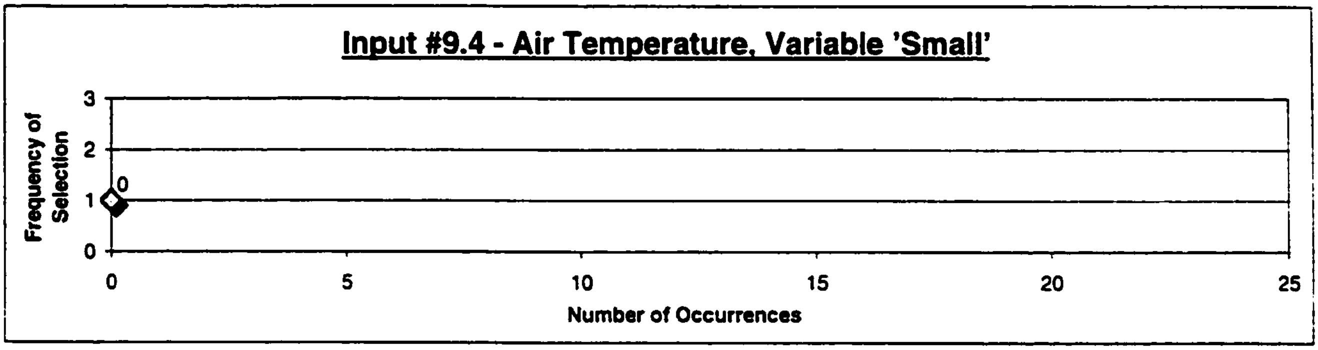

Input #9.2 - Restricted Access to Site - Rating
Small
Average
Large
Frequency of Selection
Rating

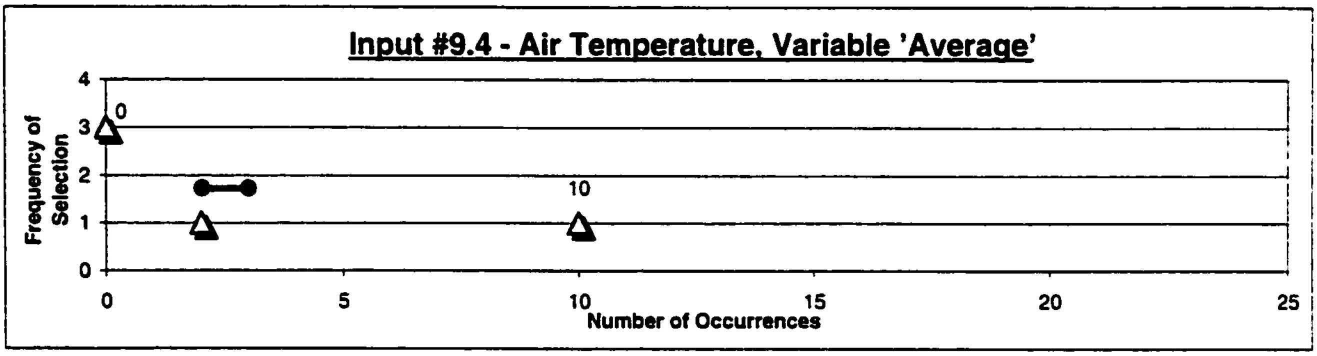

Input #9.3 - In Situ Soil Conditions, Variable 'Small'
Frequency of Selection
Number of Occurrences

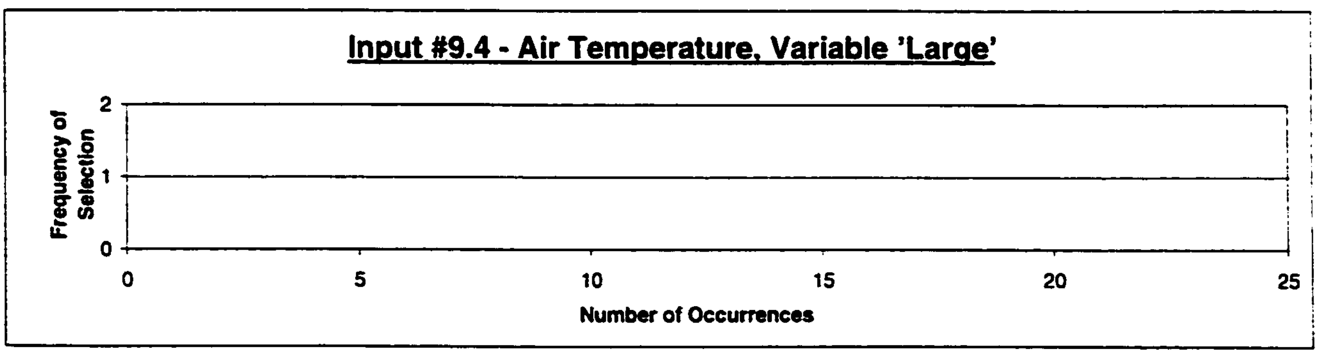

Input #9.3 - In Situ Soil Conditions, Variable 'Average'
Frequency of Selection
Number of Occurrences

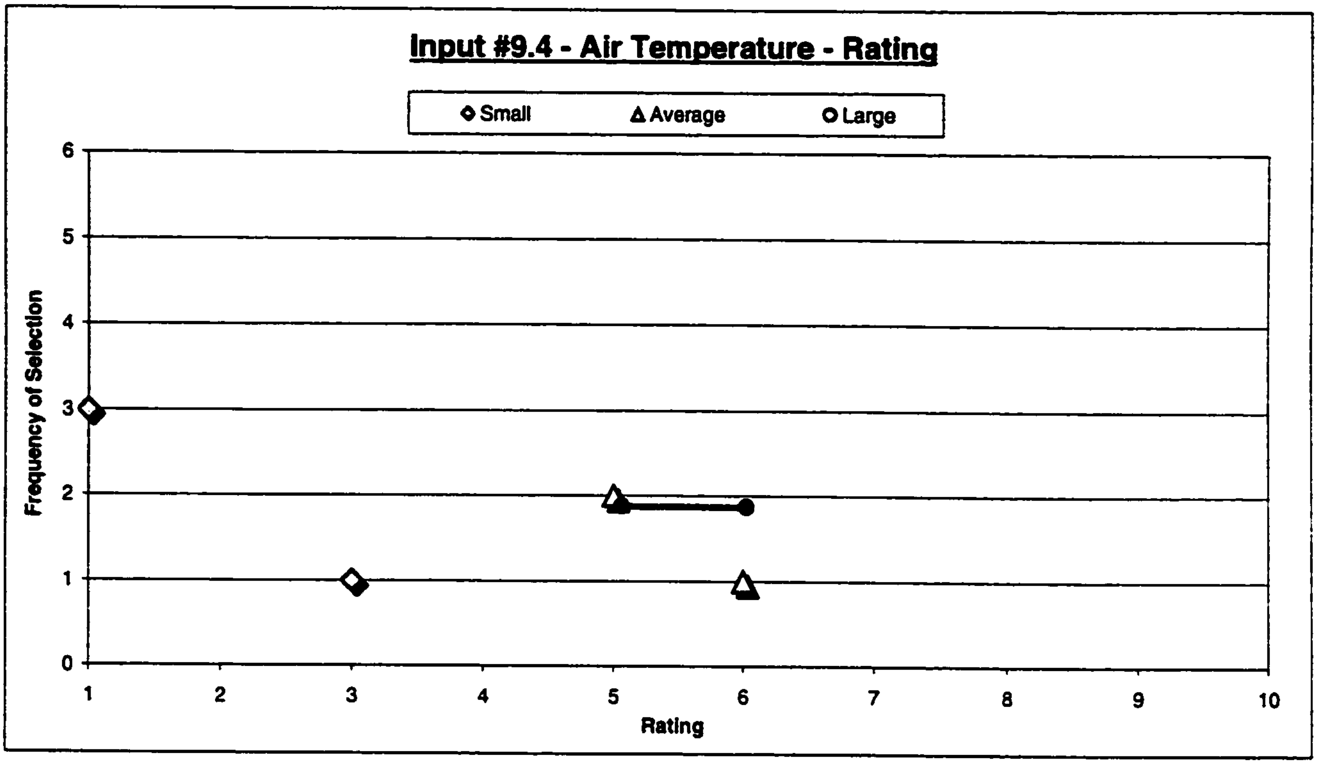

Input #9.3 - In Situ Soil Conditions, Variable 'Large'
Frequency of Selection
Number of Occurrences


Input #9.3 - In Situ Soil Conditions - Rating
Small
Average
Large
Frequency of Selection
Rating
0
1
2
3
4
5
6
7
8
9
10

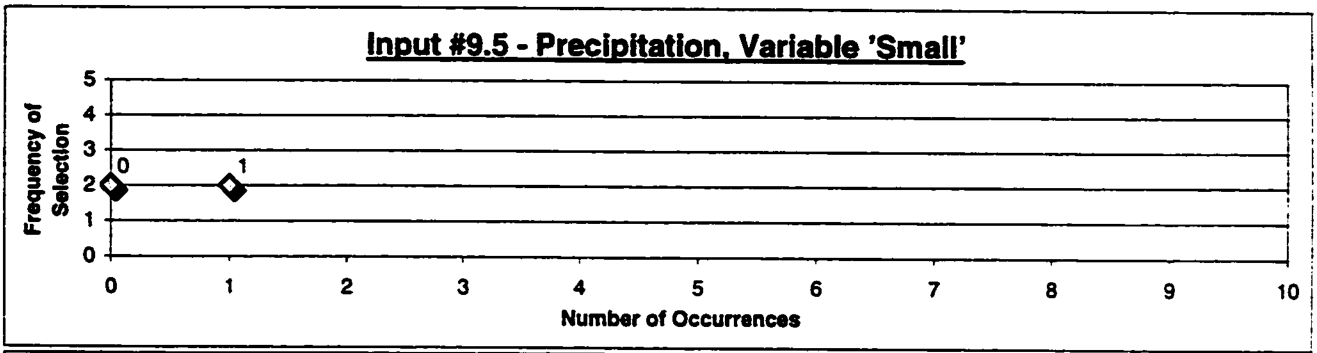

Input #9.4 - Air Temperature, Variable 'Small'
Frequency of Selection
Number of Occurrences
0
5
10
15
20
25


Input #9.4 - Air Temperature, Variable 'Average'
Frequency of Selection
Number of Occurrences
0
10
5
15
20
25

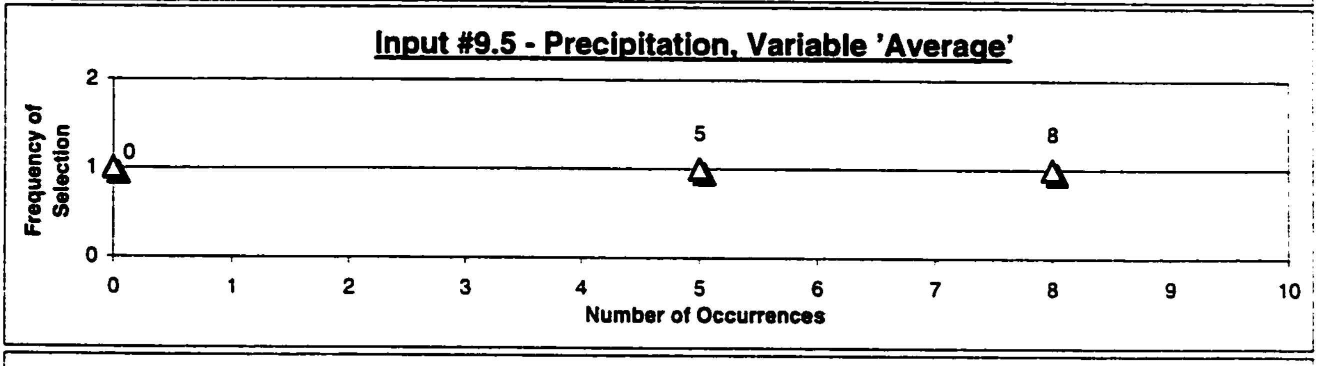

Input #9.4 - Air Temperature, Variable 'Large'
Frequency of Selection
Number of Occurrences
0
5
10
15
20
25



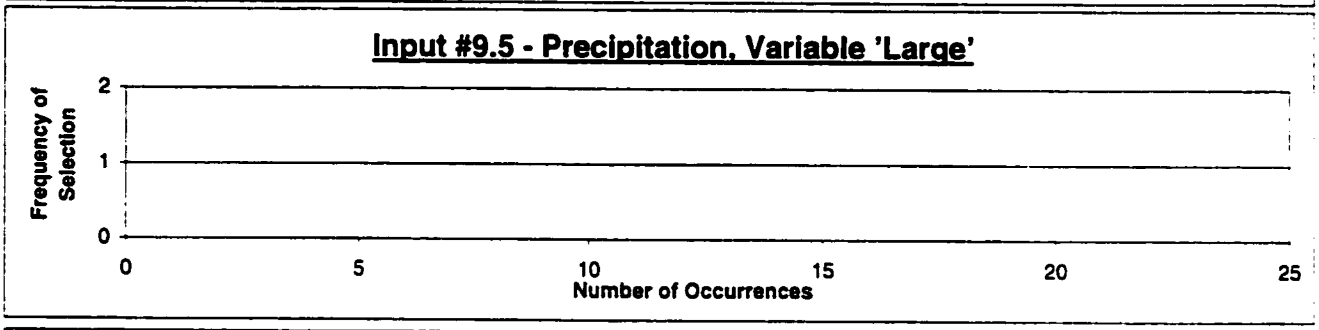
Input #9.4 - Air Temperature - Rating
Small
Average
Large
Frequency of Selection
Rating
0
1
2
3
4
5
6
1
2
3
4
5
6
7
8
9
10



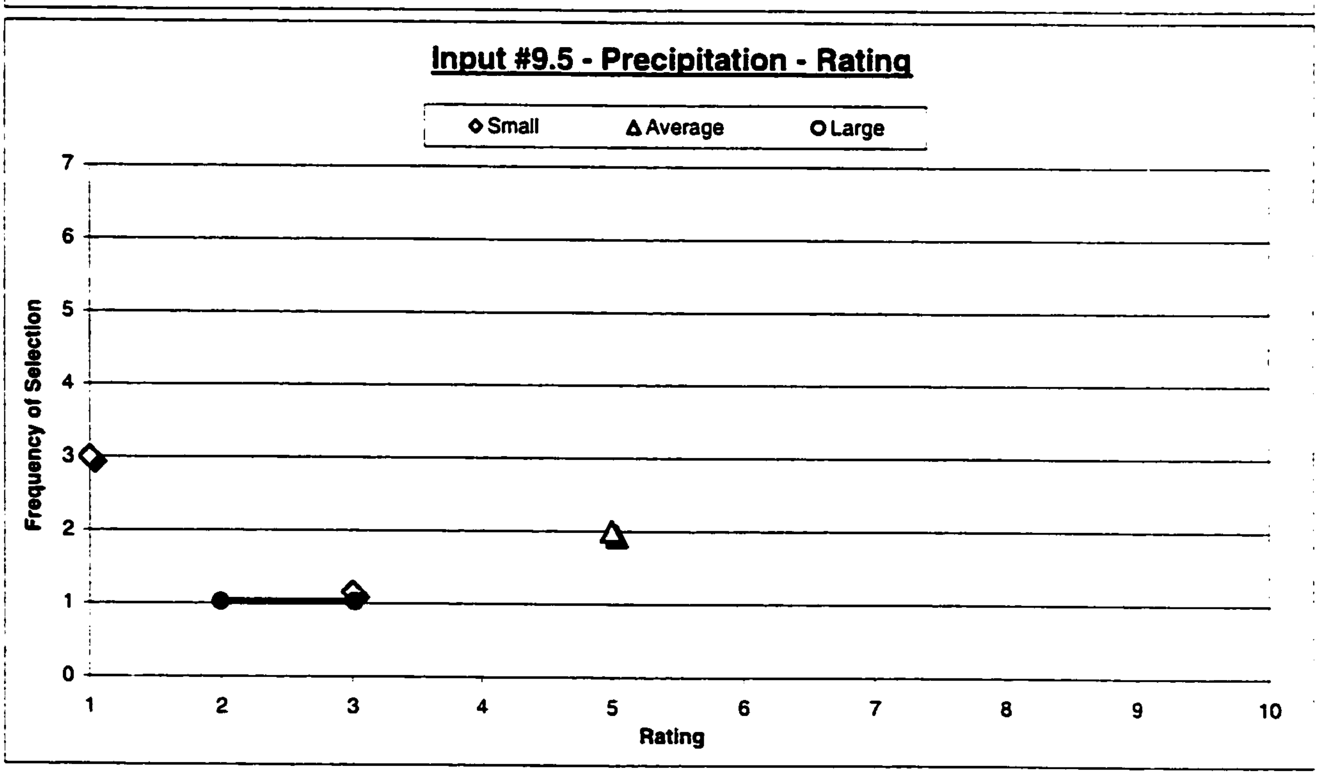
Input #9.5 - Precipitation, Variable 'Small'
Frequency of Selection
Number of Occurrences



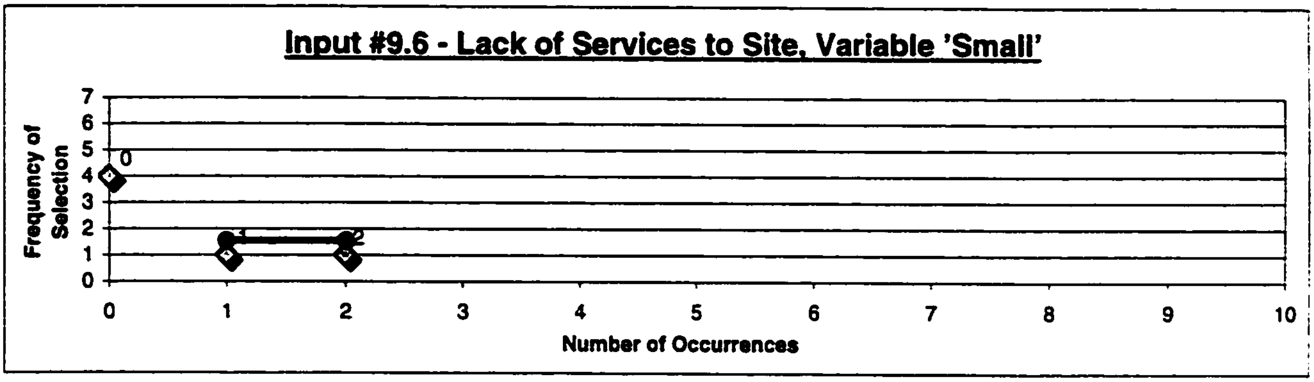
Input #9.5 - Precipitation, Variable 'Average'
Frequency of Selection
Number of Occurrences



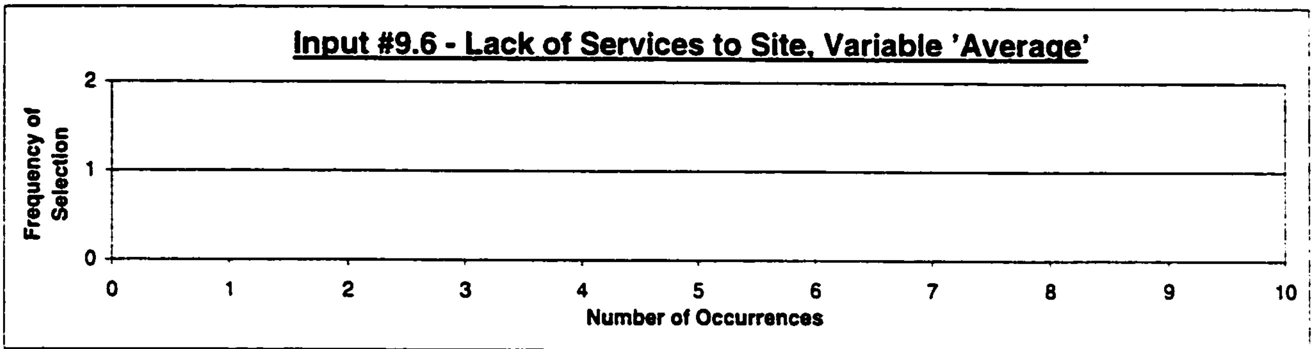
Input #9.5 - Precipitation, Variable 'Large'
Frequency of Selection
Number of Occurrences



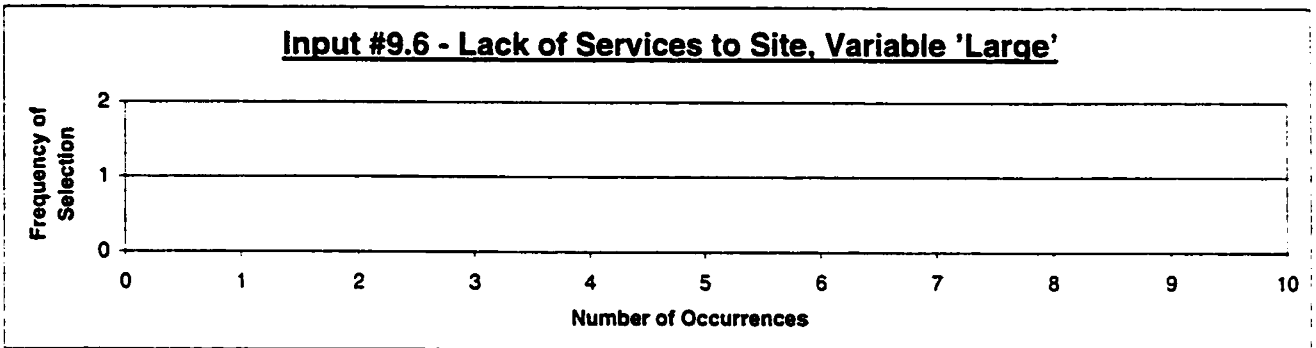
Input #9.5 - Precipitation - Rating
Small
Average
Large
Frequency of Selection
Rating



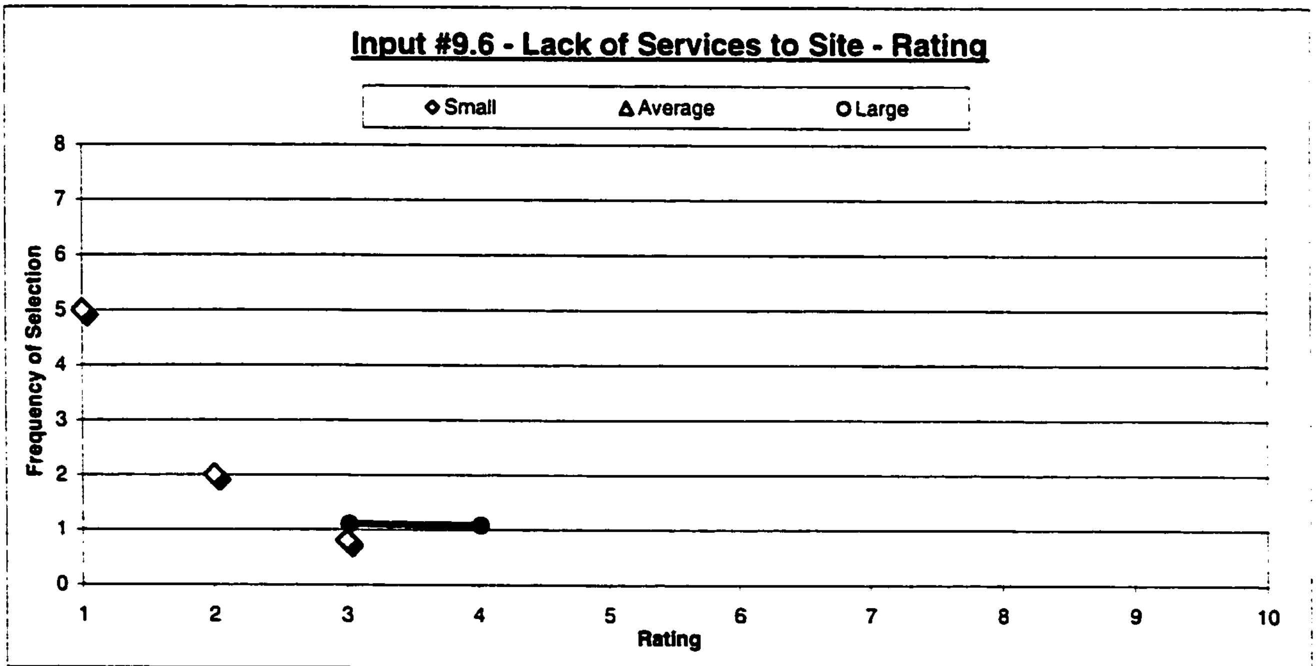
Input #9.6 - Lack of Services to Site, Variable 'Small'
Frequency of Selection
Number of Occurrences



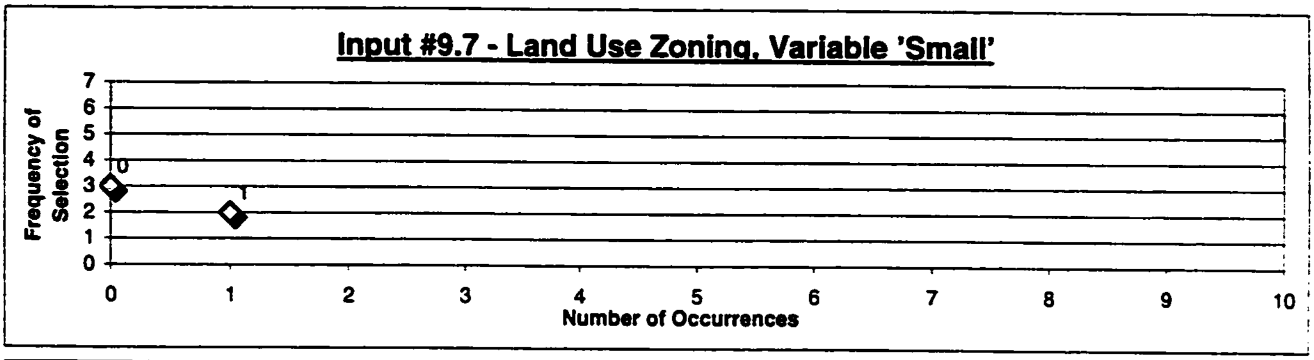
Input #9.6 - Lack of Services to Site, Variable 'Average'
Frequency of Selection
Number of Occurrences

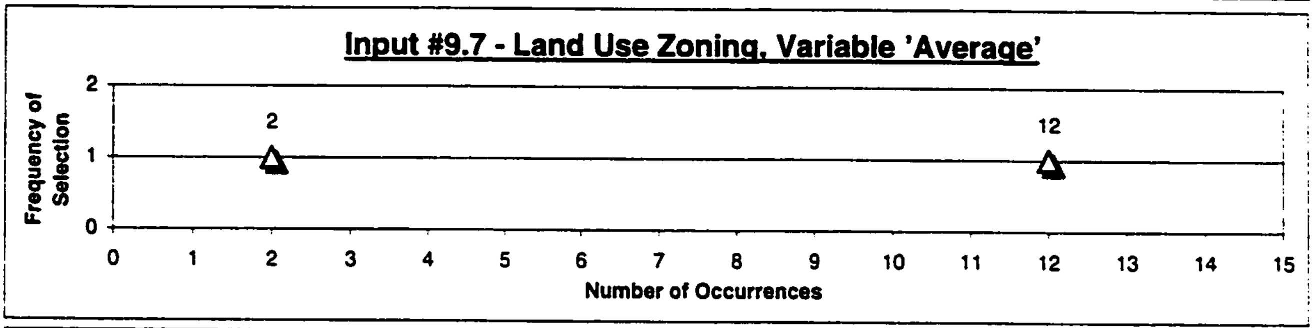


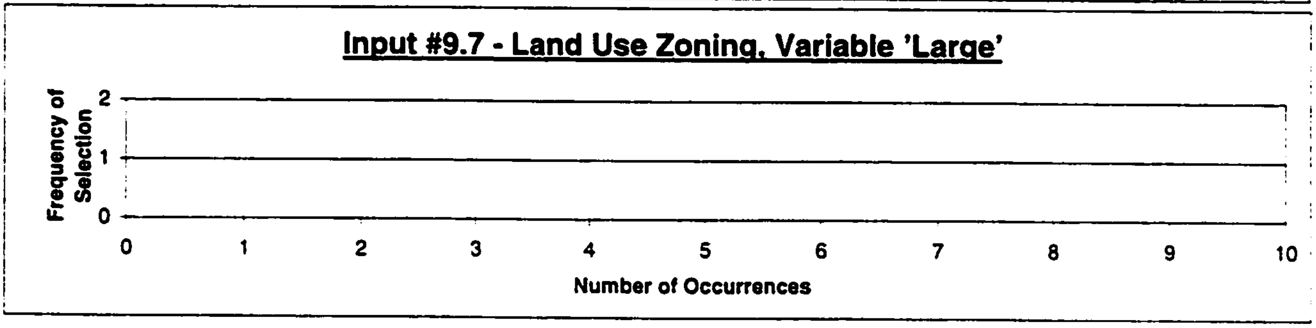
Input #9.6 - Lack of Services to Site, Variable 'Large'
Frequency of Selection
Number of Occurrences

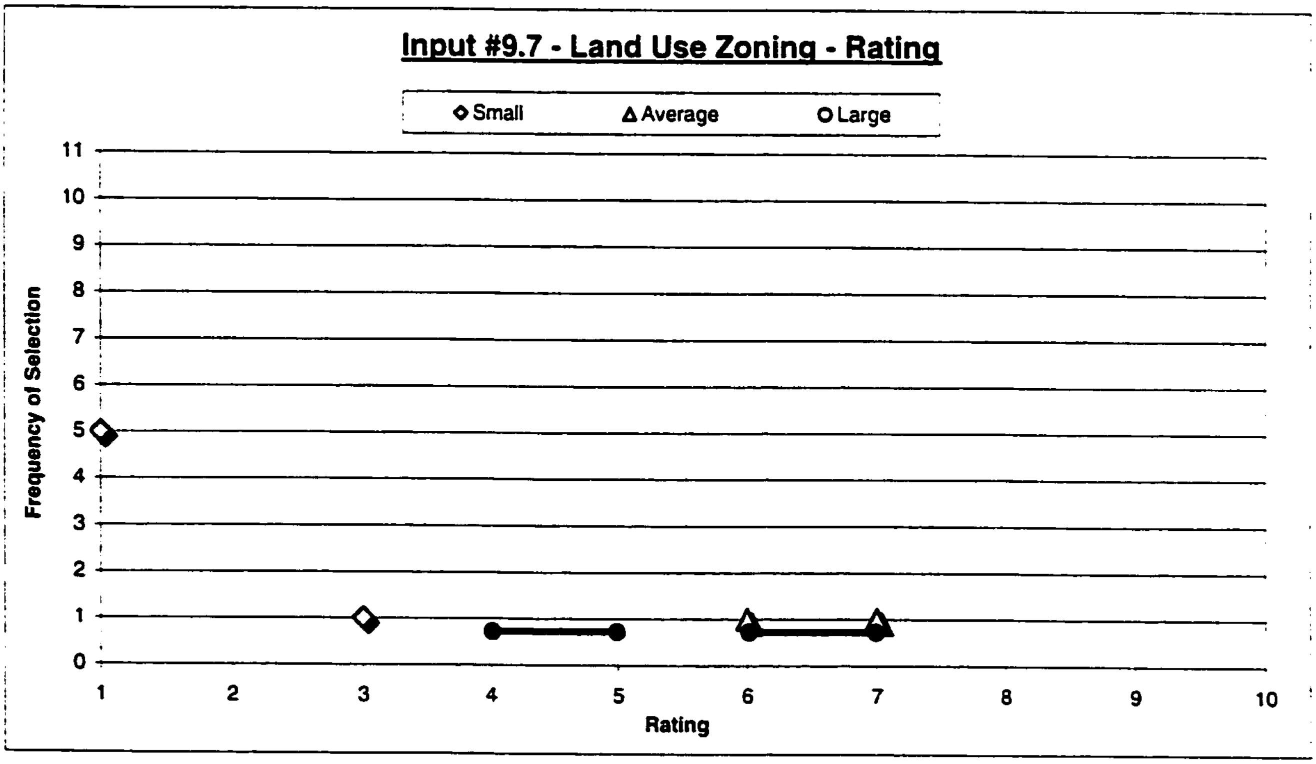


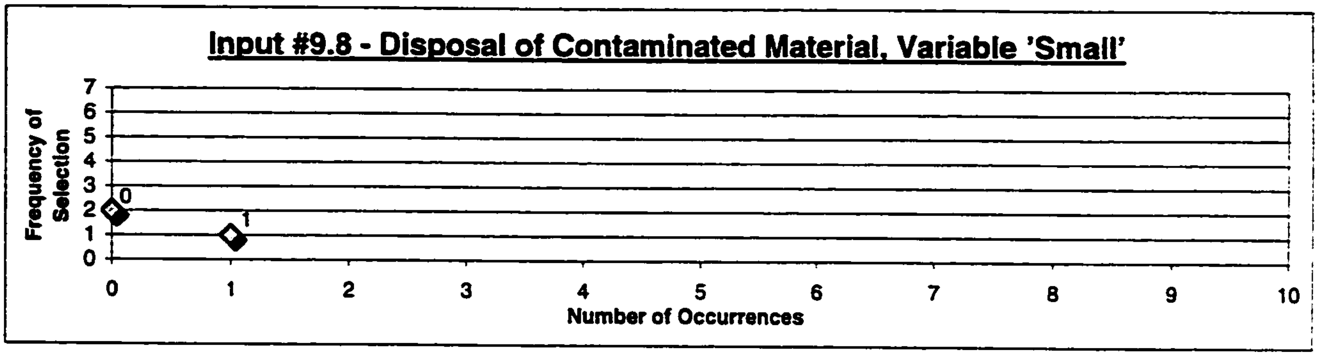
Input #9.6 - Lack of Services to Site - Rating
Small
Average
Large
Frequency of Selection
Rating


Input #9.7 - Land Use Zoning, Variable 'Small'
Frequency of Selection
Number of Occurrences
0
1

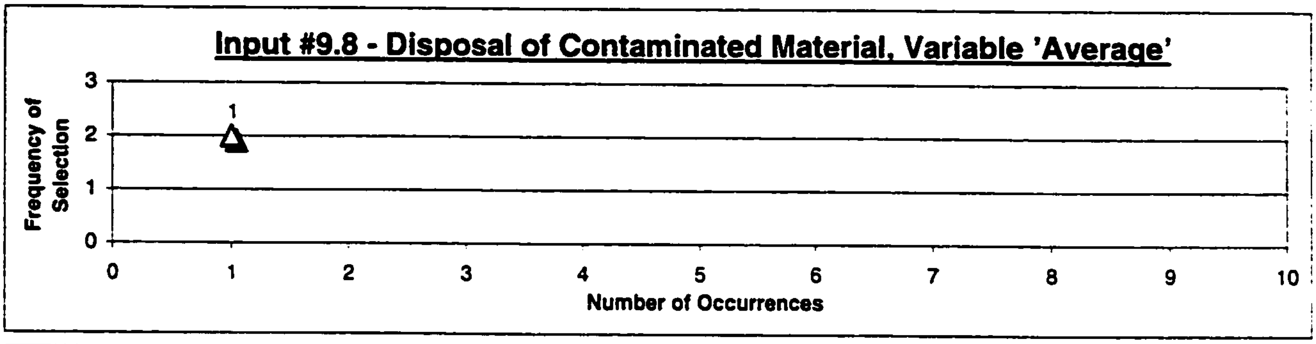

Input #9.7 - Land Use Zoning, Variable 'Average'
Frequency of Selection
Number of Occurrences
2
12


Input #9.7 - Land Use Zoning, Variable 'Large'
Frequency of Selection
Number of Occurrences

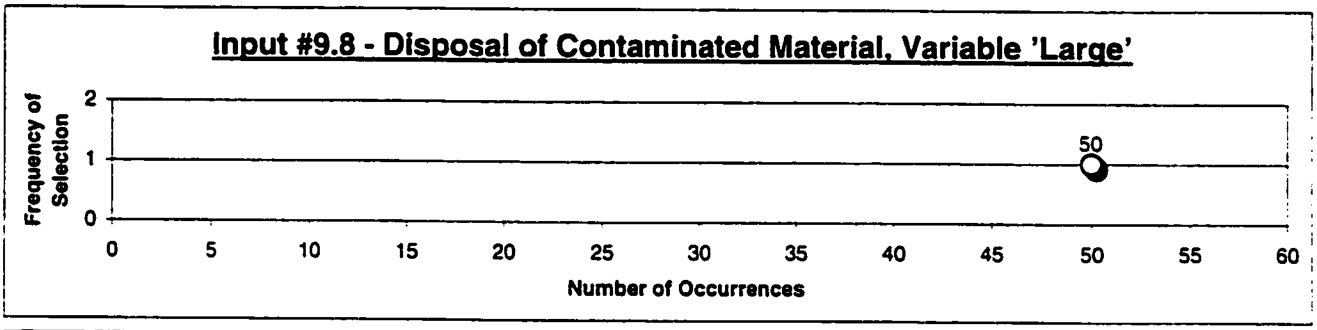

Input #9.7 - Land Use Zoning - Rating
Small
Average
Large
Frequency of Selection
Rating



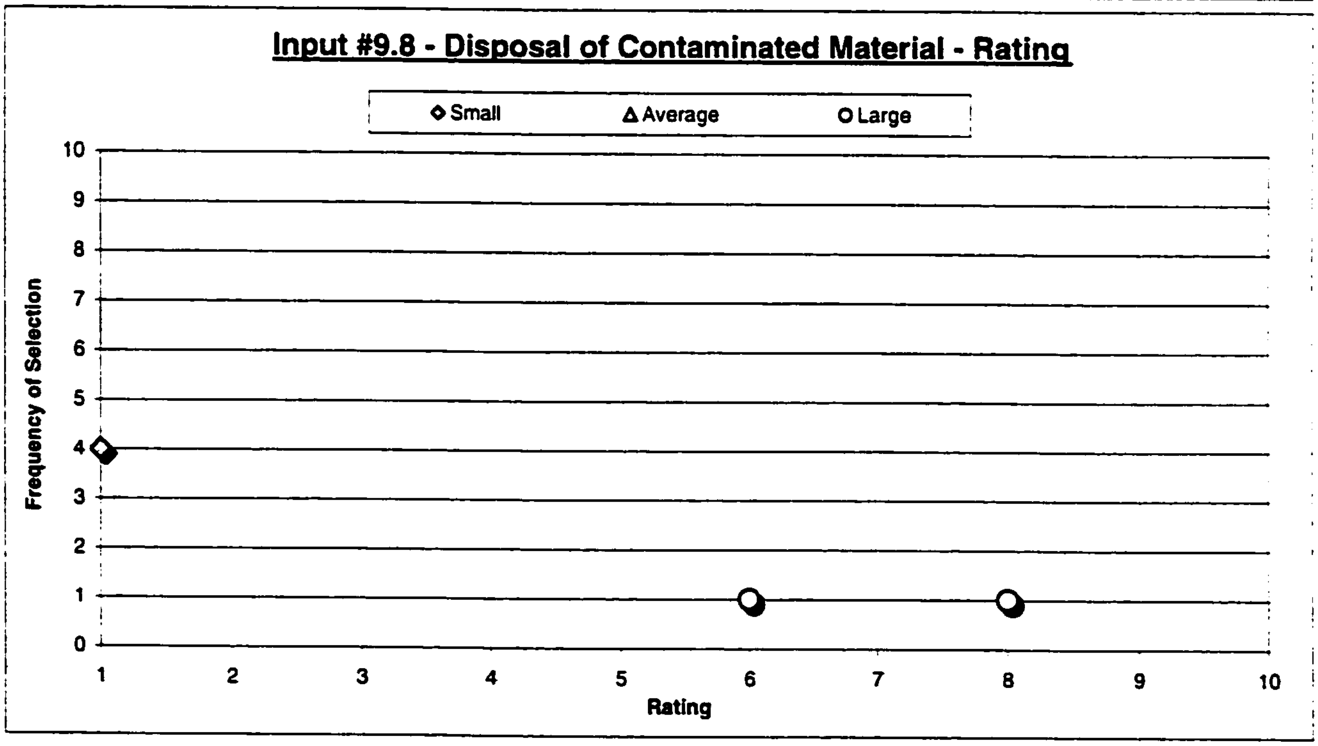
Input #9.8 - Disposal of Contaminated Material, Variable 'Small'
Frequency of Selection
Number of Occurrences
0
1



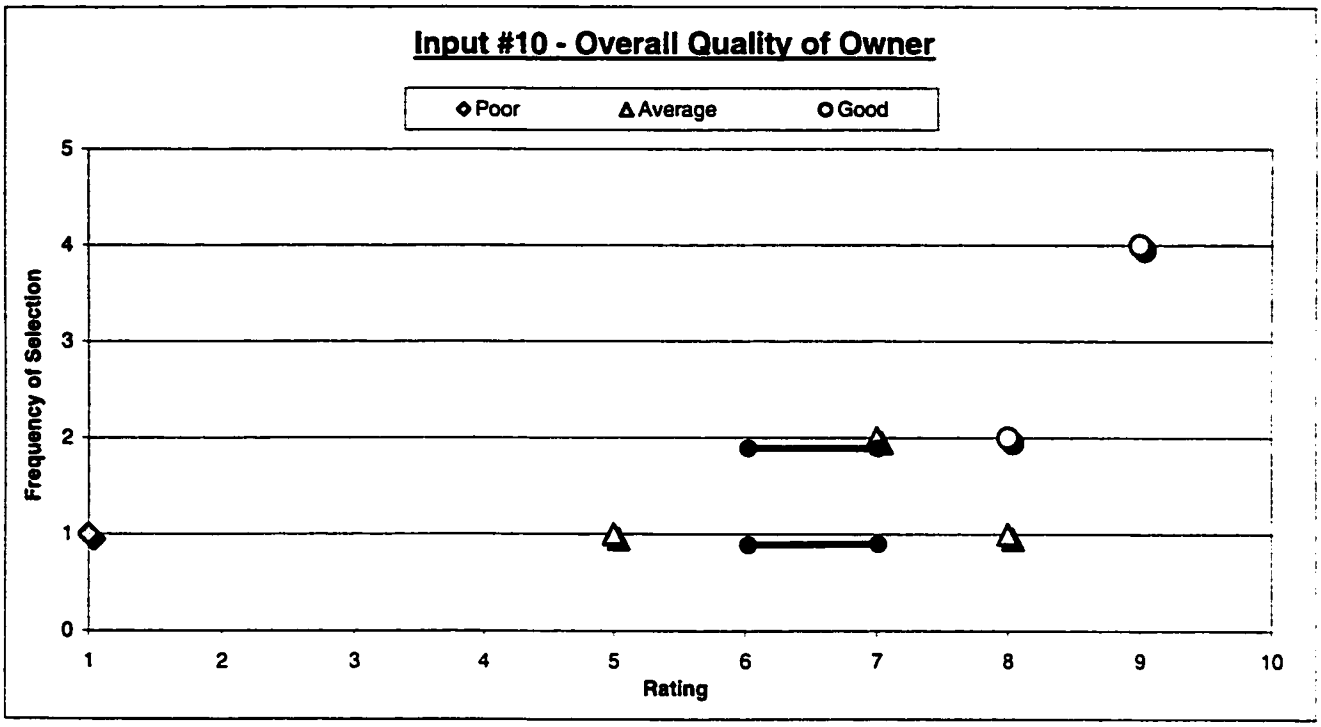
Input #9.8 - Disposal of Contaminated Material, Variable 'Average'
Frequency of Selection
Number of Occurrences
1



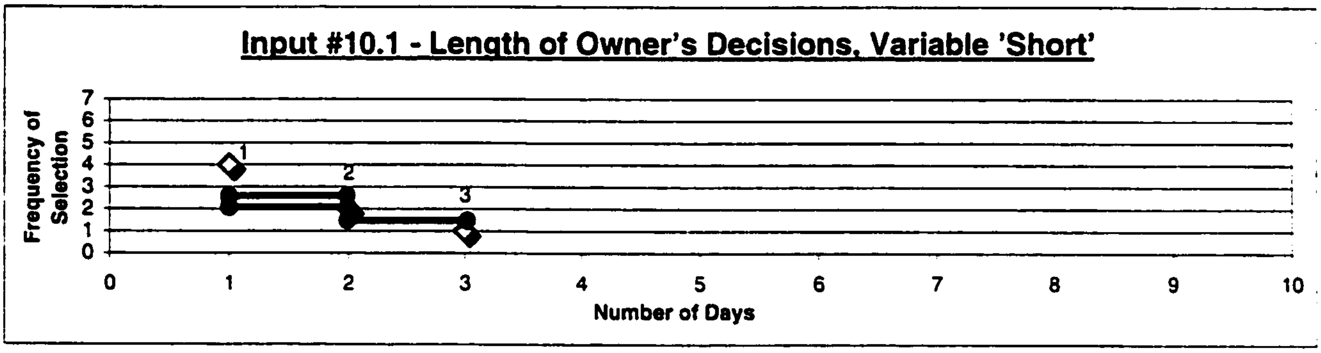
Input #9.8 - Disposal of Contaminated Material, Variable 'Large'
Frequency of Selection
Number of Occurrences
50



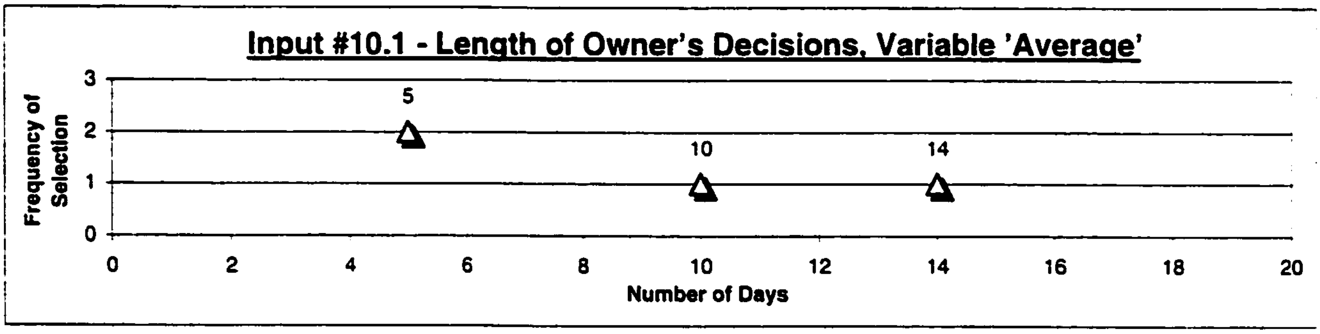
Input #9.8 - Disposal of Contaminated Material - Rating
Small
Average
Large
Frequency of Selection
Rating



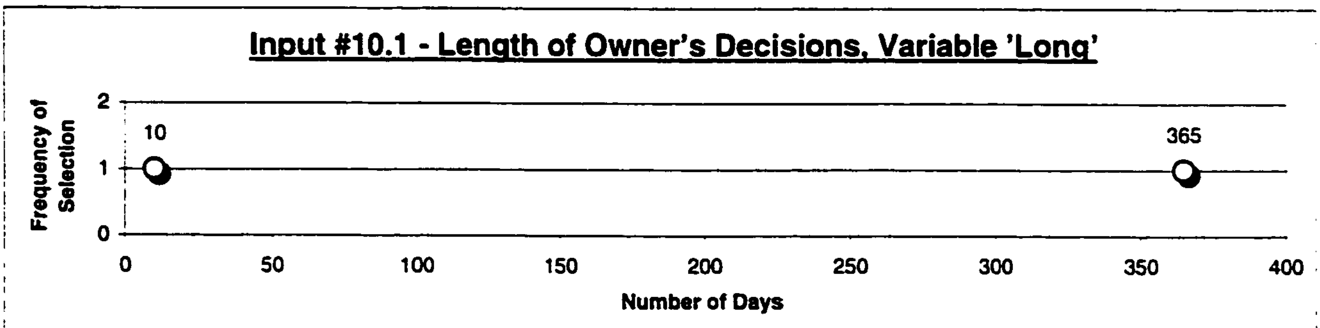
Input #10 - Overall Quality of Owner
Poor
Average
Good
Frequency of Selection
Rating



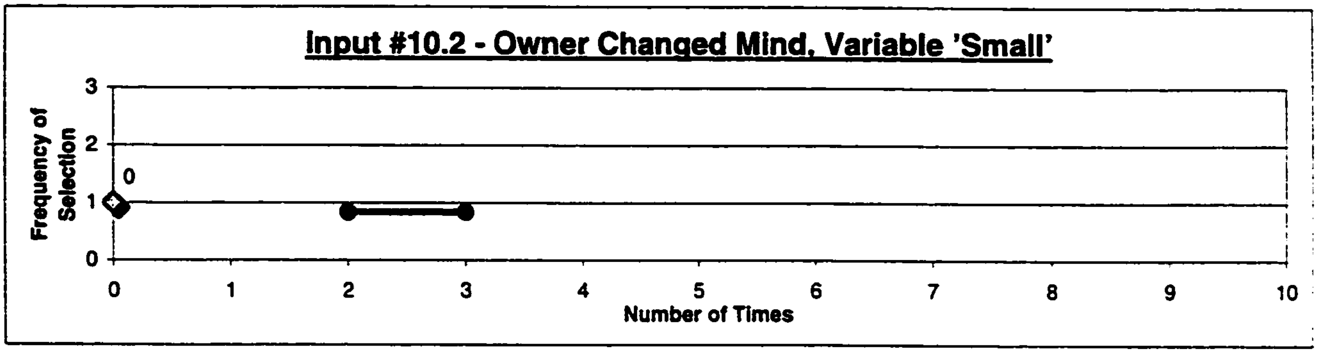
Input #10.1 - Length of Owner's Decisions, Variable 'Short'
Frequency of Selection
Number of Days



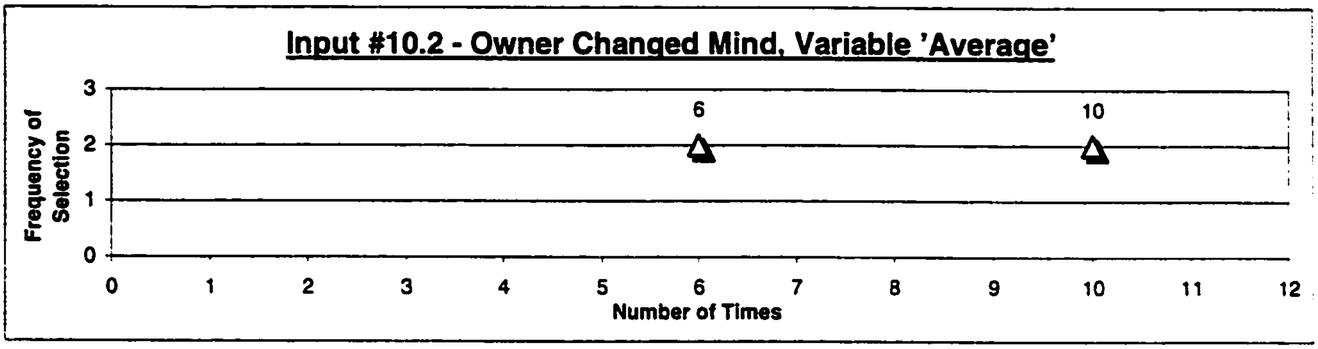
Input #10.1 - Length of Owner's Decisions, Variable 'Average'
Frequency of Selection
Number of Days



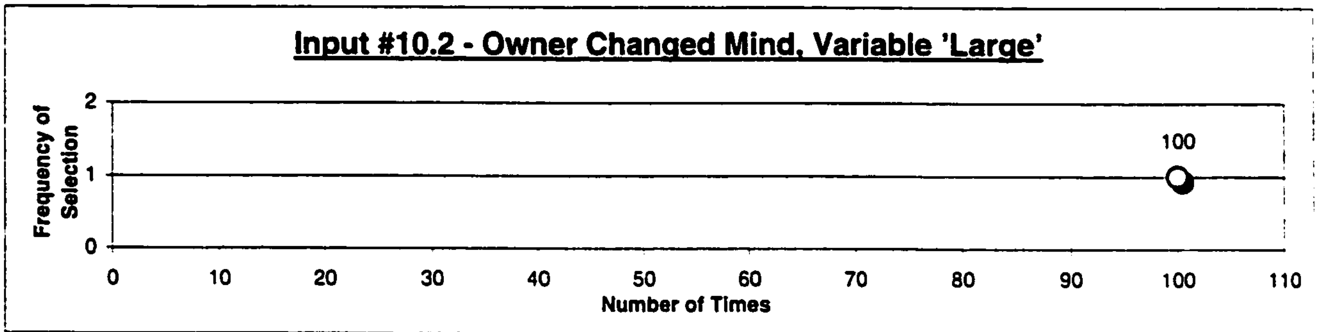
Input #10.1 - Length of Owner's Decisions, Variable 'Long'
Frequency of Selection
Number of Days



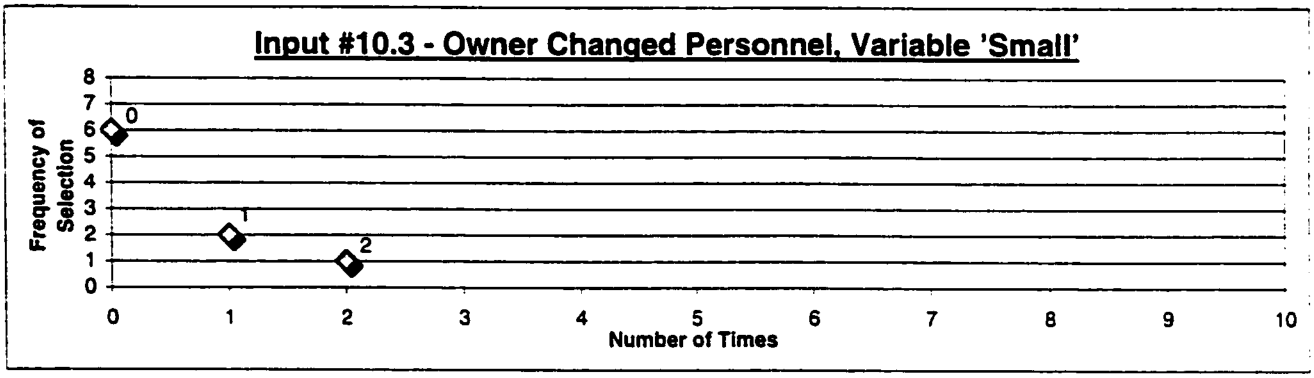
Input #10.2 - Owner Changed Mind, Variable 'Small'
Frequency of Selection
Number of Times
0



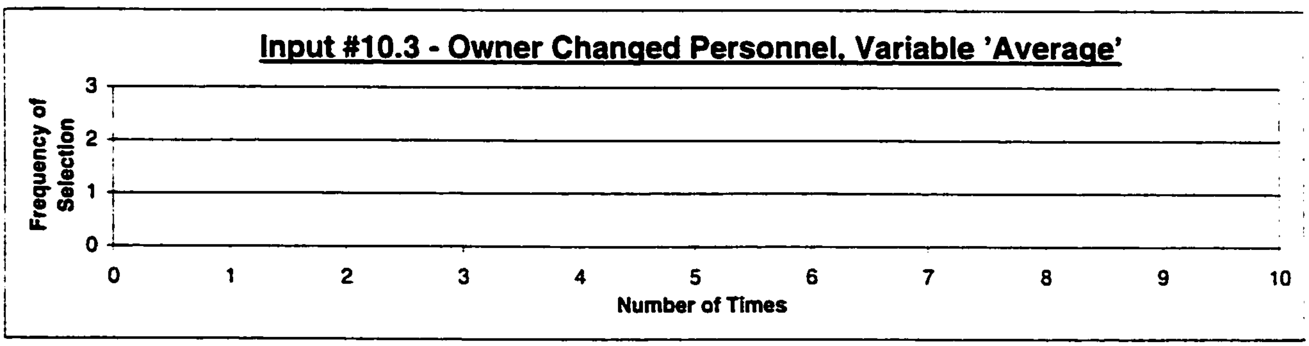
Input #10.2 - Owner Changed Mind, Variable 'Average'
Frequency of Selection
Number of Times
6
10



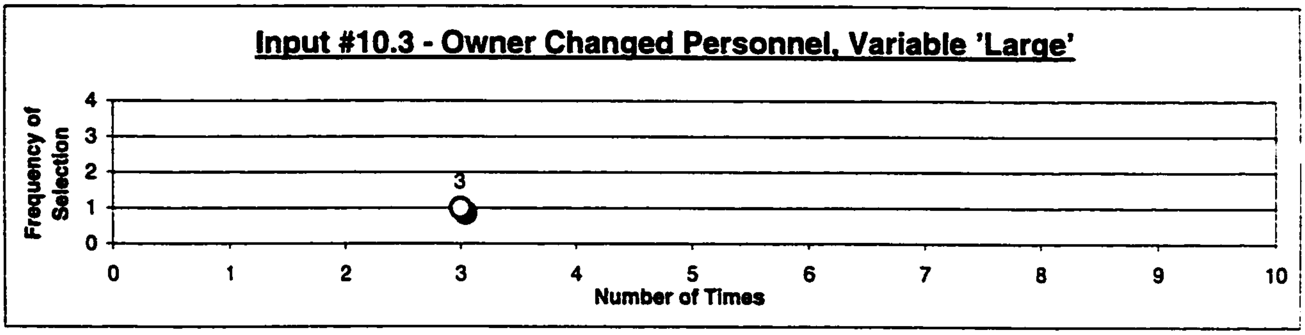
Input #10.2 - Owner Changed Mind, Variable 'Large'
Frequency of Selection
Number of Times
100



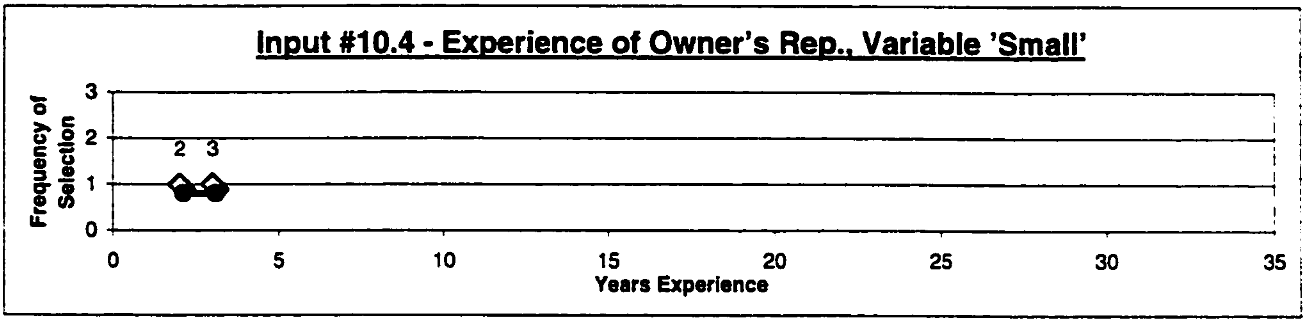
Input #10.3 - Owner Changed Personnel, Variable 'Small'
Frequency of Selection
Number of Times
0
1
2



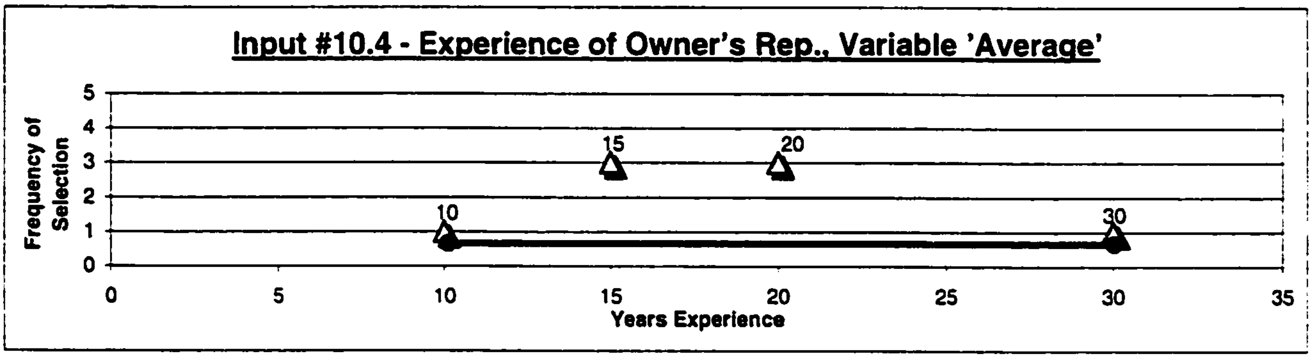
Input #10.3 - Owner Changed Personnel, Variable 'Average'
Frequency of Selection
Number of Times



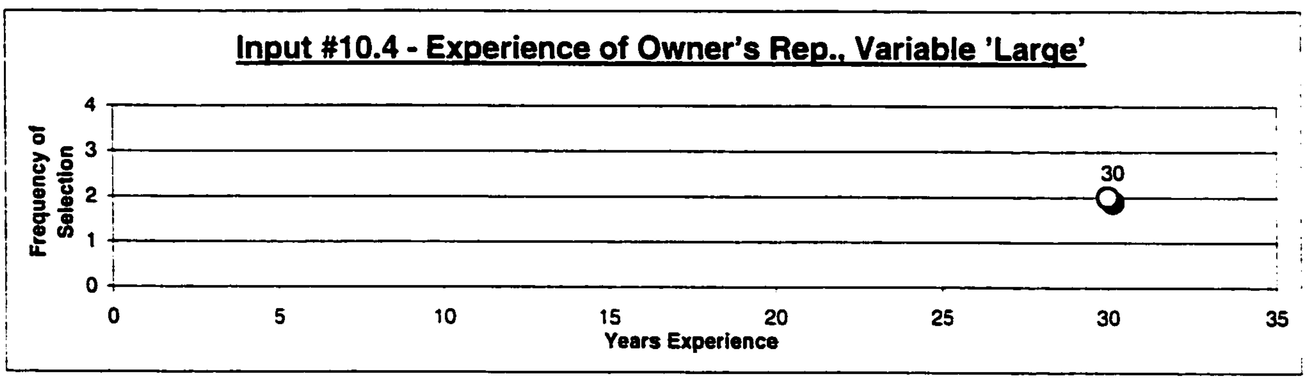
Input #10.3 - Owner Changed Personnel, Variable 'Large'
Frequency of Selection
Number of Times
3



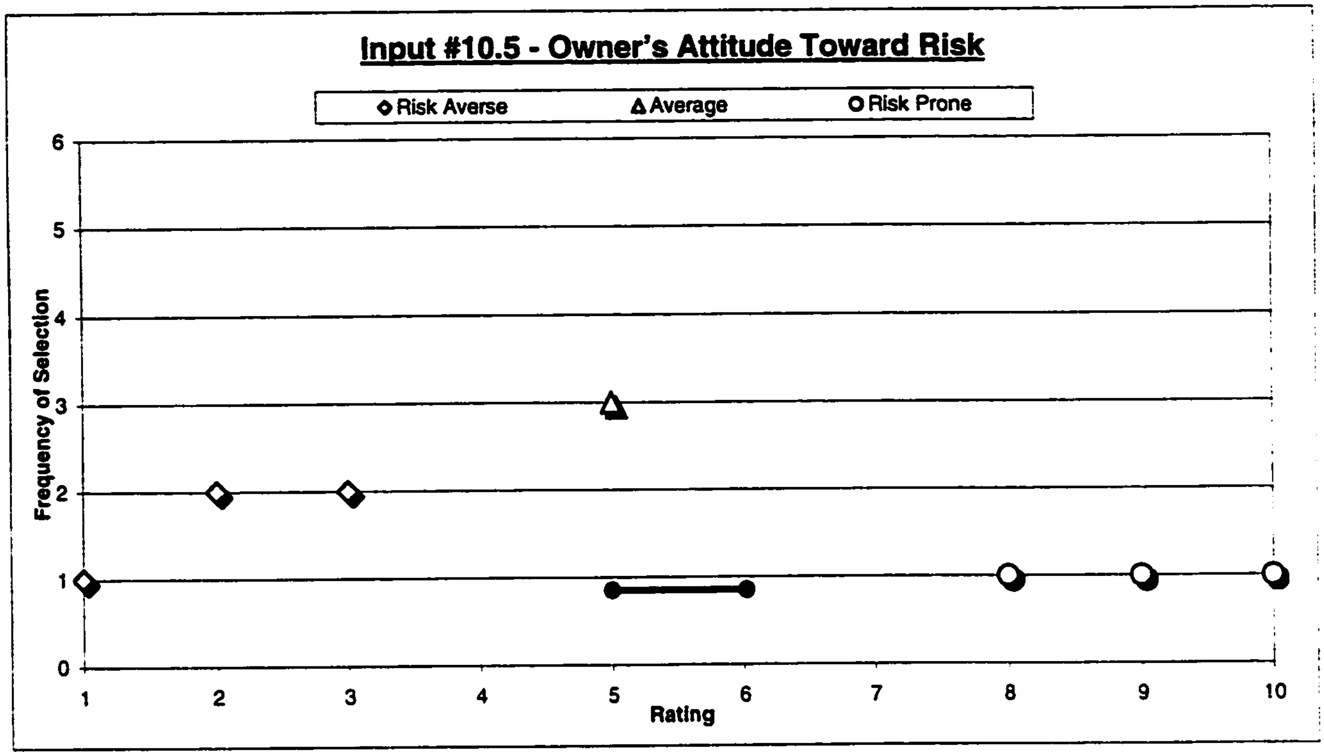
Input #10.4 - Experience of Owner's Rep., Variable 'Small'
Frequency of Selection
Years Experience
2
3



Input #10.4 - Experience of Owner's Rep., Variable 'Average'
Frequency of Selection
Years Experience
10
15
20
30

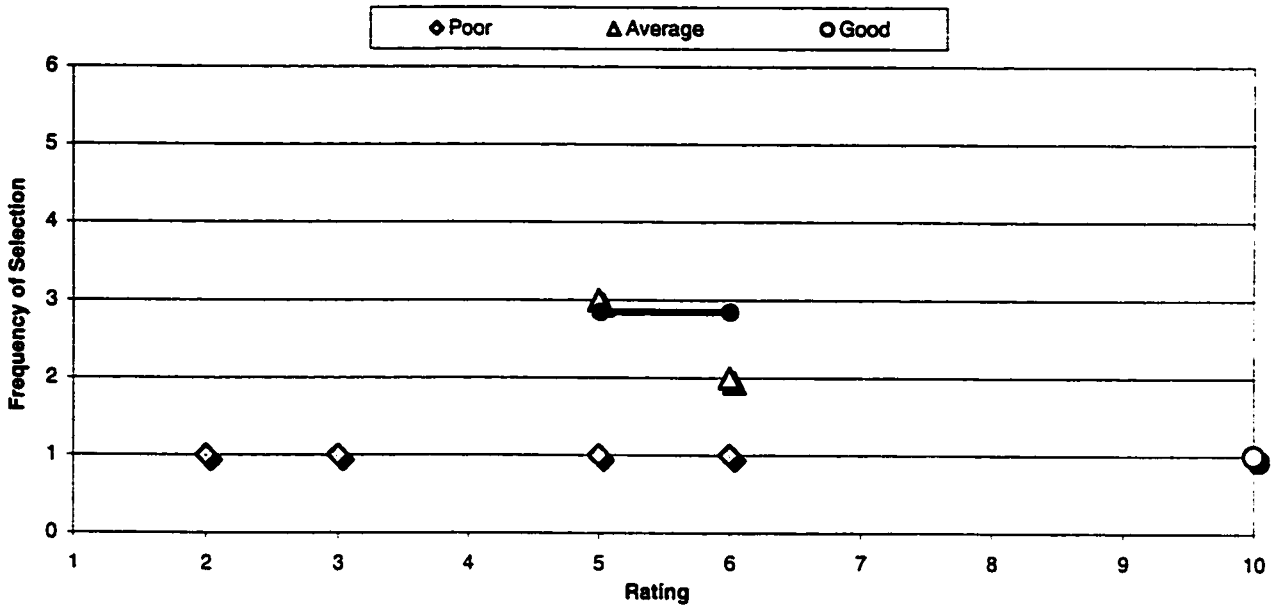


Input #10.4 - Experience of Owner's Rep., Variable 'Large'
Frequency of Selection
Years Experience
30

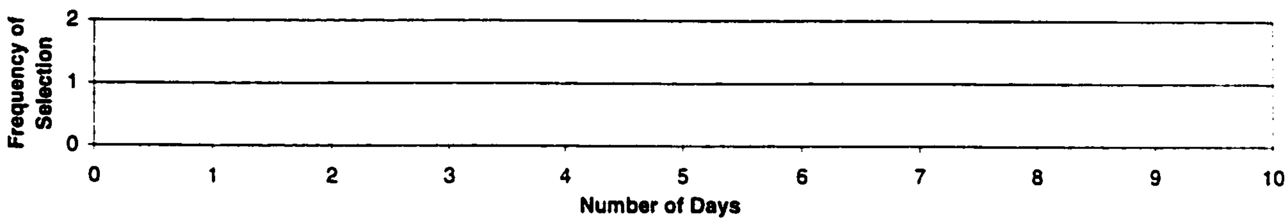


Input #10.5 - Owner's Attitude Toward Risk
Risk Averse
Average
Risk Prone
Frequency of Selection
0
1
2
3
4
5
6
1
2
3
4
5
6
7
8
9
10
Rating

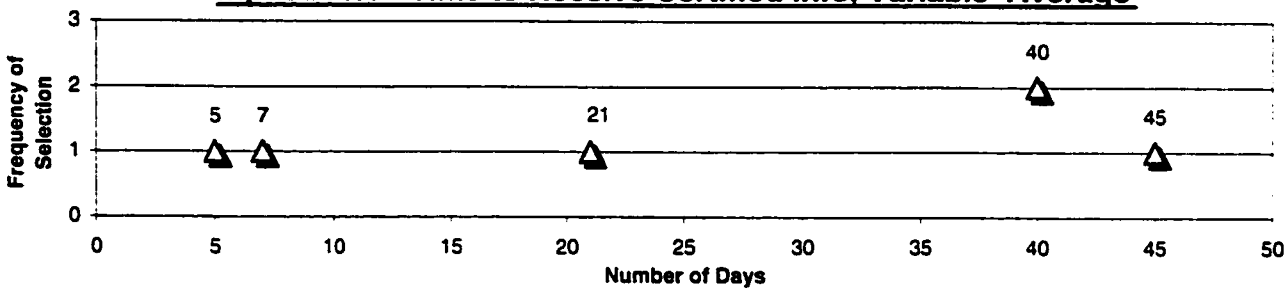
Input #11 - Quality of Vendors Profiles



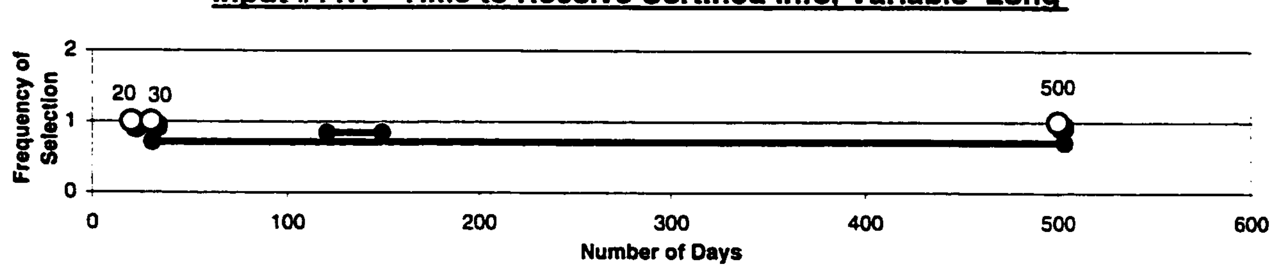
Input #11.1 - Time to Receive Certified Info, Variable 'Short'

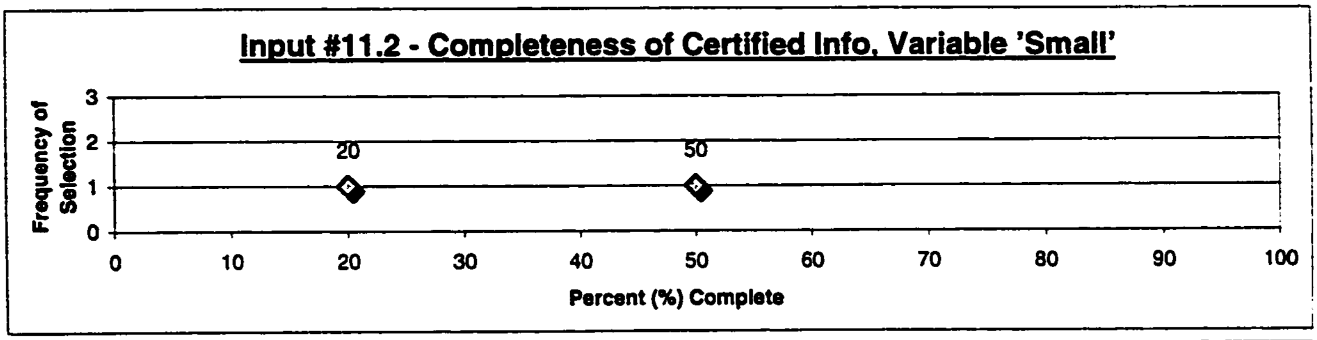


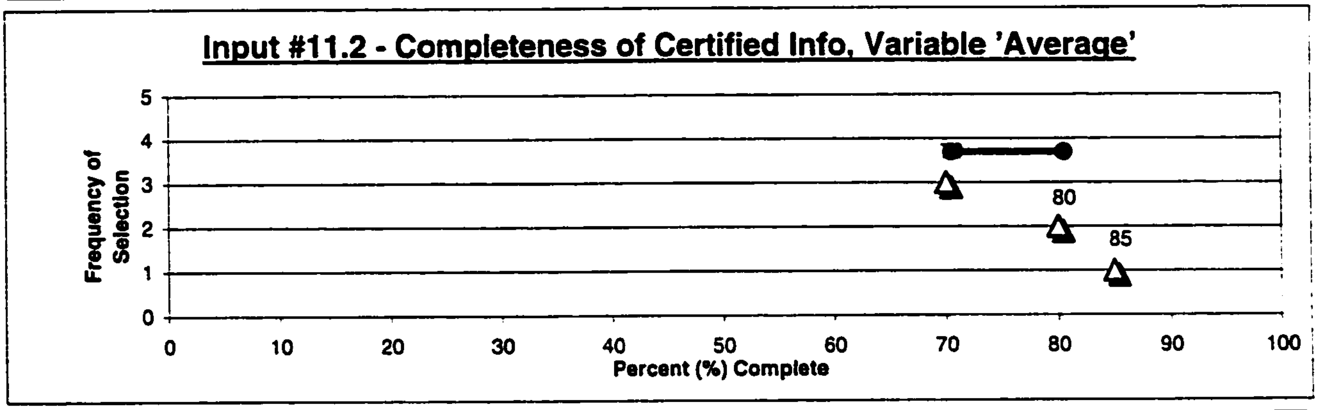
Input #11.1 - Time to Receive Certified Info, Variable 'Average'

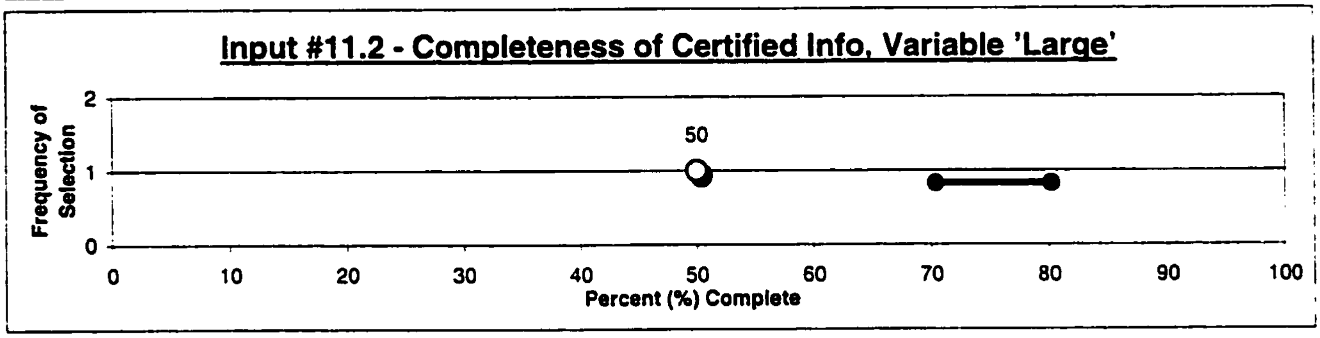


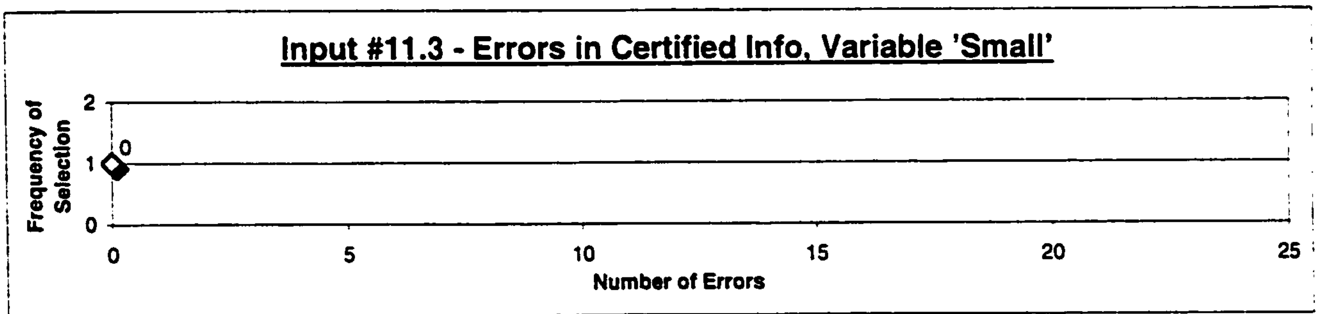
Input #11.1 - Time to Receive Certified Info, Variable 'Long'

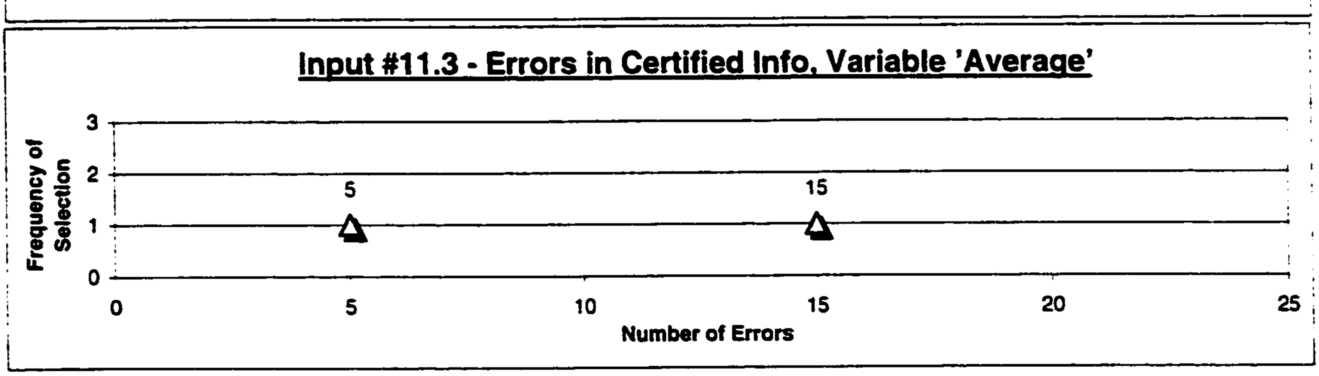


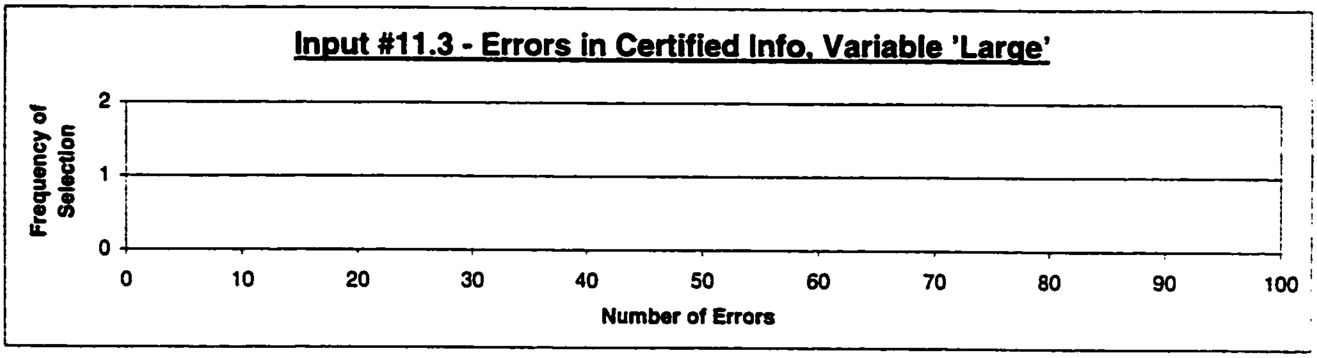

Input #11.2 - Completeness of Certified Info, Variable 'Small'
Frequency of Selection
Percent (%) Complete
20
50

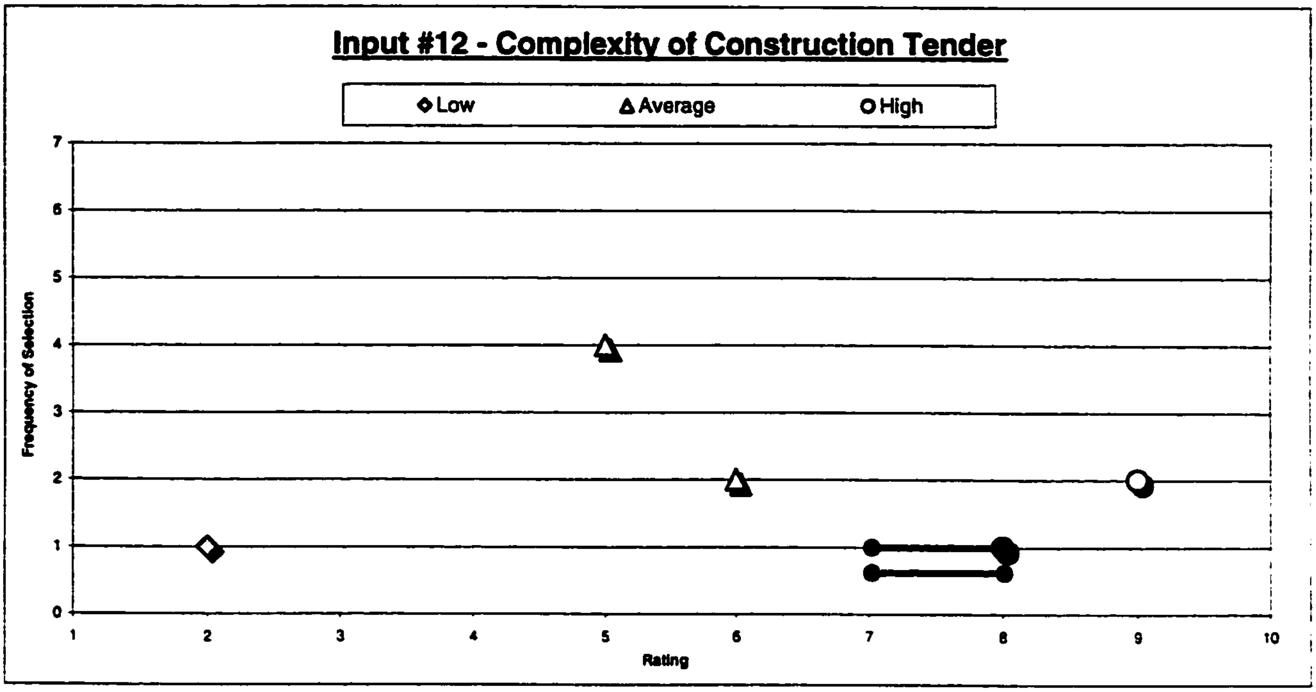

Input #11.2 - Completeness of Certified Info, Variable 'Average'
Frequency of Selection
Percent (%) Complete
80
85

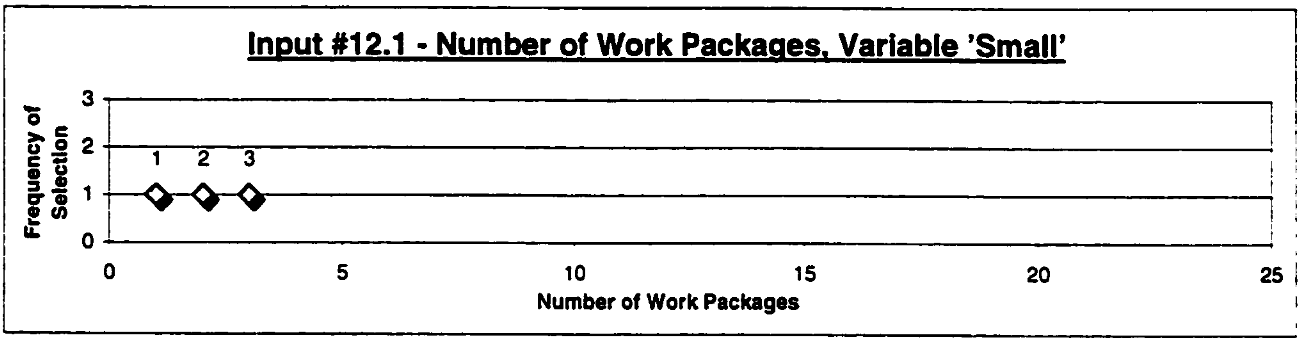

Input #11.2 - Completeness of Certified Info, Variable 'Large'
Frequency of Selection
Percent (%) Complete
50

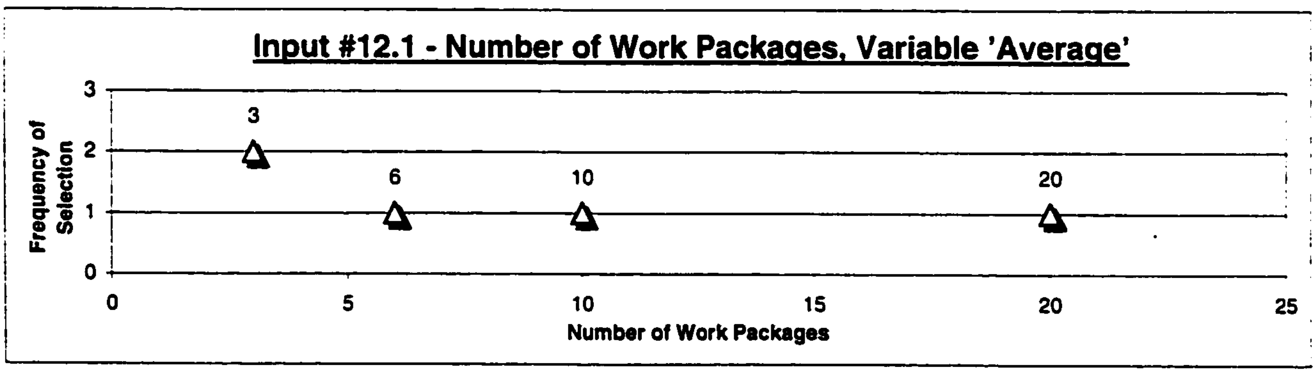

Input #11.3 - Errors in Certified Info, Variable 'Small'
Frequency of Selection
Number of Errors
0

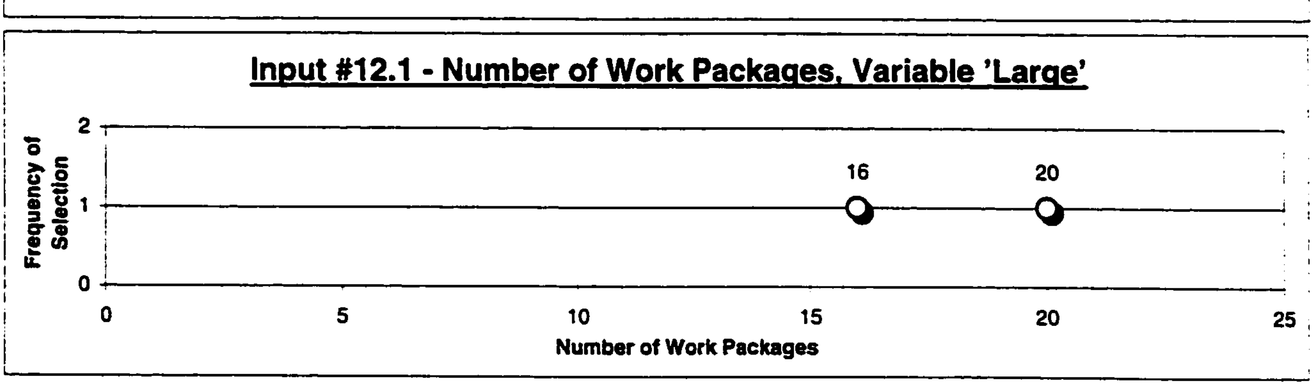

Input #11.3 - Errors in Certified Info, Variable 'Average'
Frequency of Selection
Number of Errors
5
15

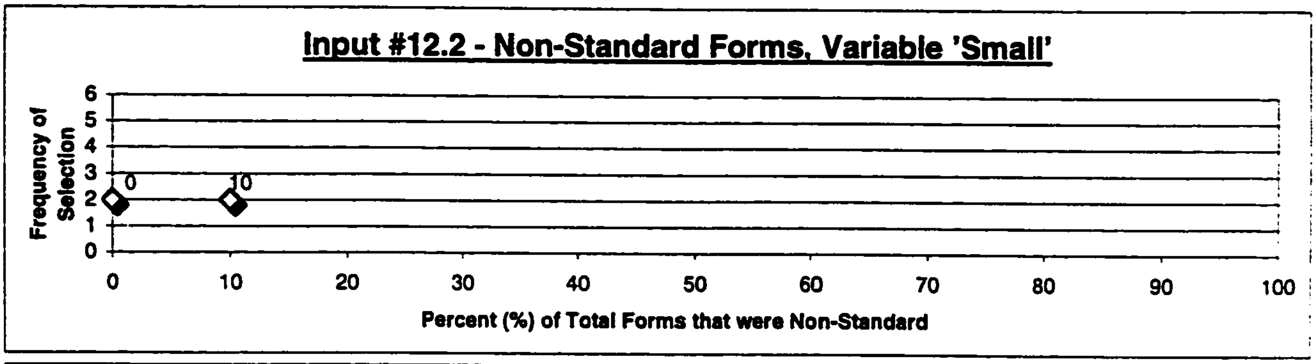

Input #11.3 - Errors in Certified Info, Variable 'Large'
Frequency of Selection
2
1
0
0
10
20
30
40
50
60
70
80
90
100
Number of Errors


Input #12 - Complexity of Construction Tender
Low
Average
High
Frequency of Selection
Rating

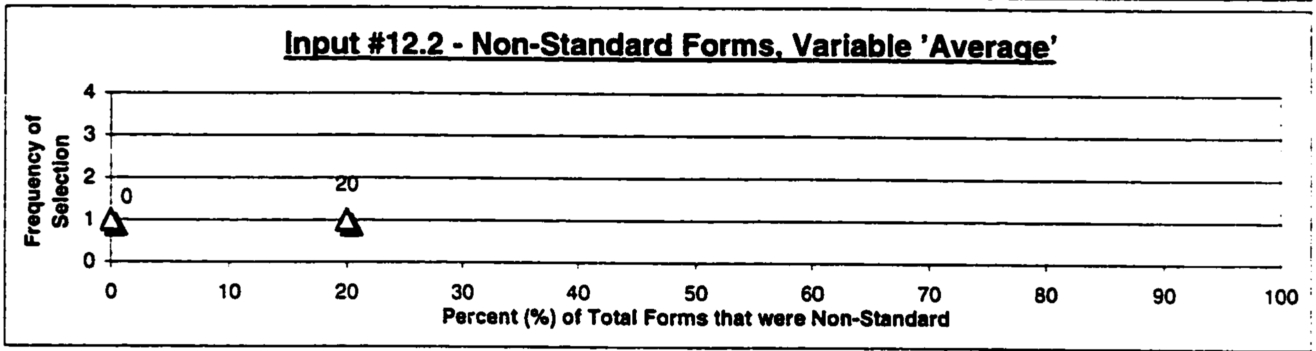

Input #12.1 - Number of Work Packages, Variable 'Small'
Frequency of Selection
1
2
3
Number of Work Packages


Input #12.1 - Number of Work Packages, Variable 'Average'
Frequency of Selection
3
6
10
20
Number of Work Packages

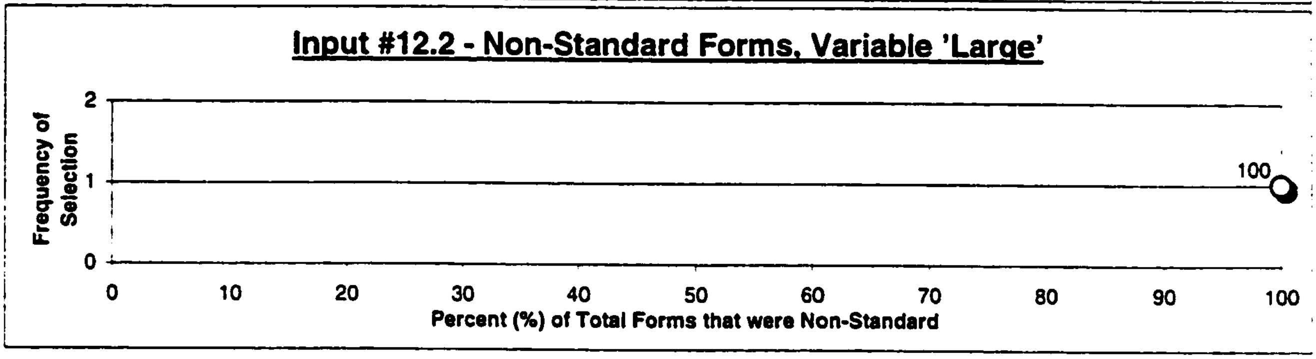

Input #12.1 - Number of Work Packages, Variable 'Large'
Frequency of Selection
16
20
Number of Work Packages



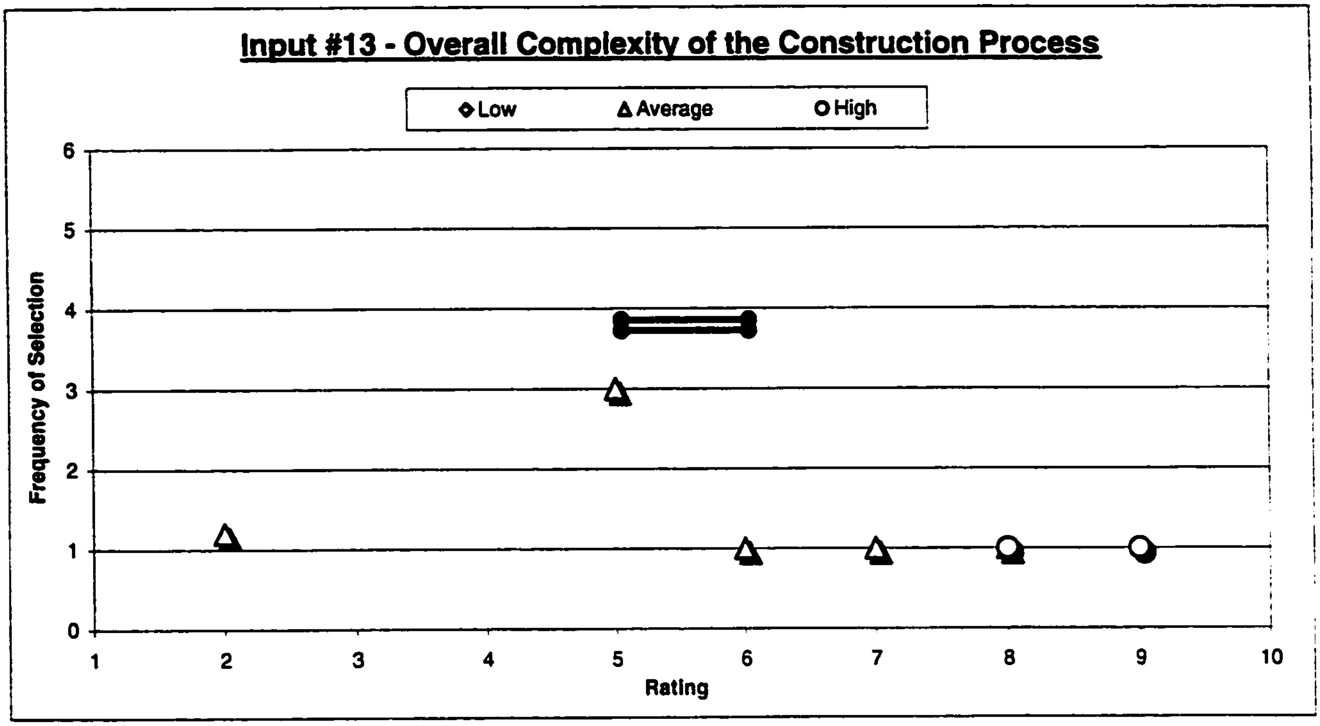
Input #12.2 - Non-Standard Forms, Variable 'Small'
Frequency of Selection
0
1
2
3
4
5
6
0
10
20
30
40
50
60
70
80
90
100
Percent (%) of Total Forms that were Non-Standard



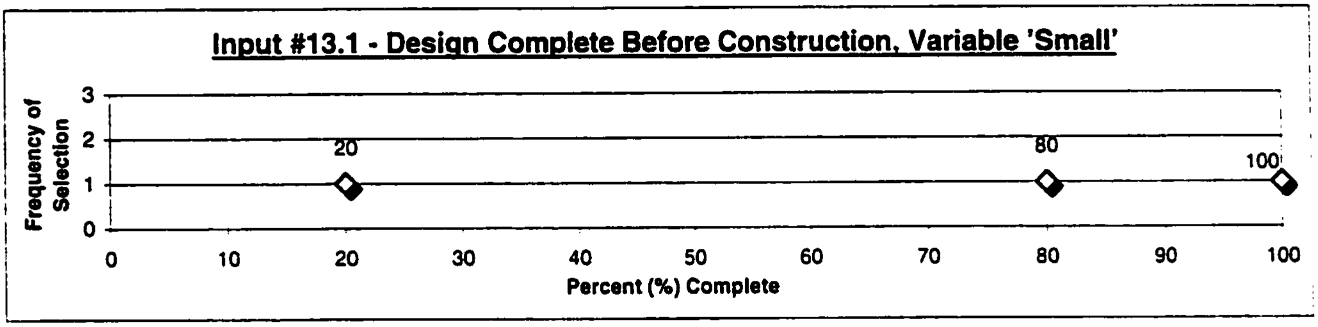
Input #12.2 - Non-Standard Forms, Variable 'Average'
Frequency of Selection
0
1
2
3
4
0
20
0
10
20
30
40
50
60
70
80
90
100
Percent (%) of Total Forms that were Non-Standard



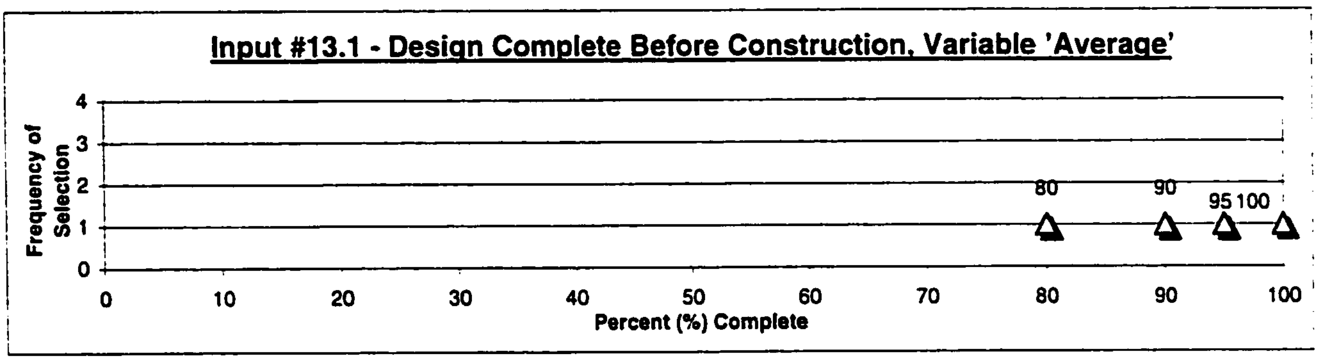
Input #12.2 - Non-Standard Forms, Variable 'Large'
Frequency of Selection
0
1
2
100
0
10
20
30
40
50
60
70
80
90
100
Percent (%) of Total Forms that were Non-Standard



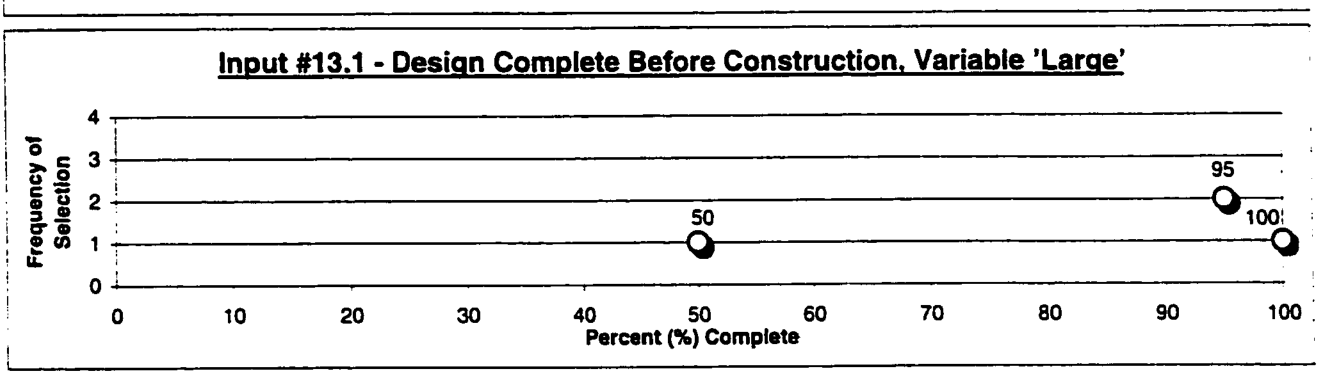
Input #13 - Overall Complexity of the Construction Process
Low
Average
High
Frequency of Selection
0
1
2
3
4
5
6
1
2
3
4
5
6
7
8
9
10
Rating



Input #13.1 - Design Complete Before Construction, Variable 'Small'
Frequency of Selection
0
1
2
3
20
80
100
0
10
20
30
40
50
60
70
80
90
100
Percent (%) Complete

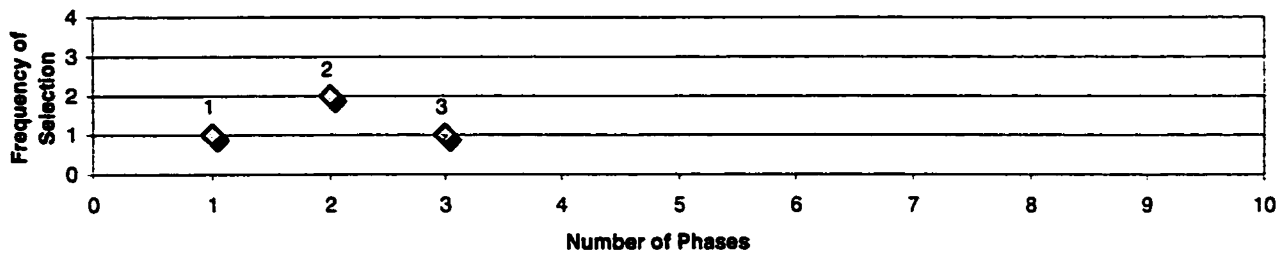


Input #13.1 - Design Complete Before Construction, Variable 'Average'
Frequency of Selection
0
1
2
3
4
80
90
95
100
0
10
20
30
40
50
60
70
80
90
100
Percent (%) Complete

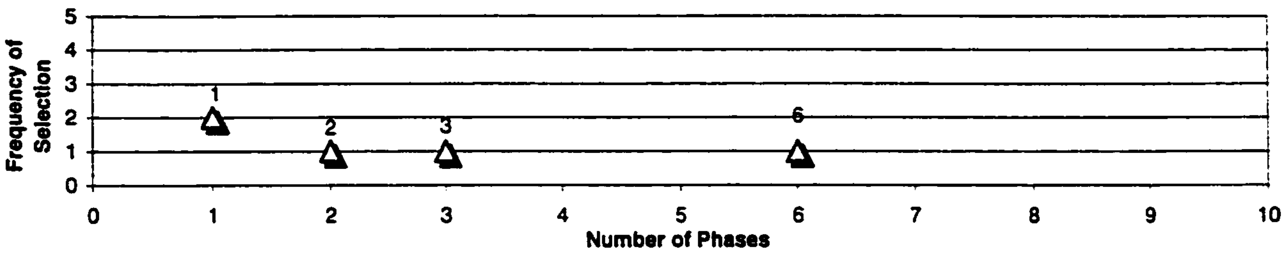


Input #13.1 - Design Complete Before Construction, Variable 'Large'
Frequency of Selection
0
1
2
3
4
50
95
100
0
10
20
30
40
50
60
70
80
90
100
Percent (%) Complete

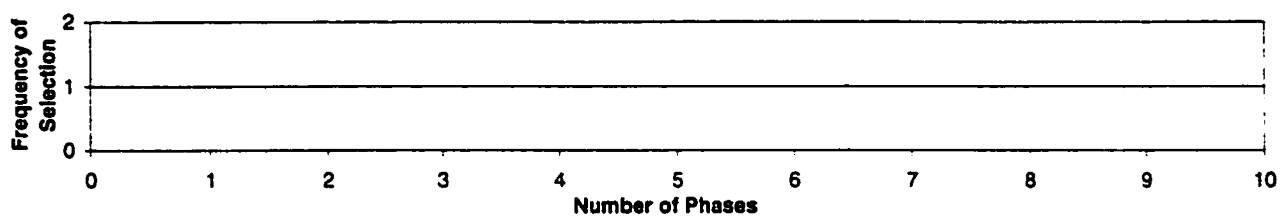
Input #13.2 - Phases of Construction, Variable 'Small'



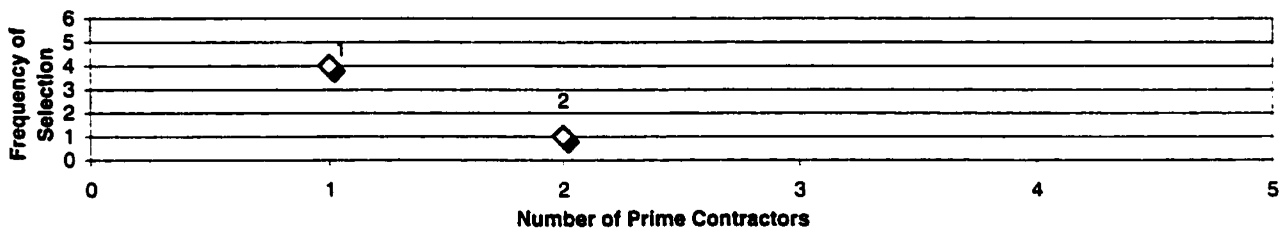
Input #13.2 - Phases of Construction, Variable 'Average'



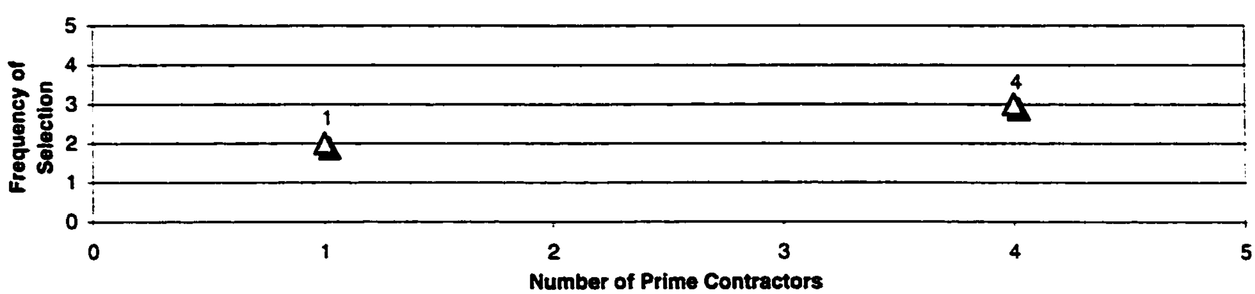
Input #13.2 - Phases of Construction, Variable 'Large'

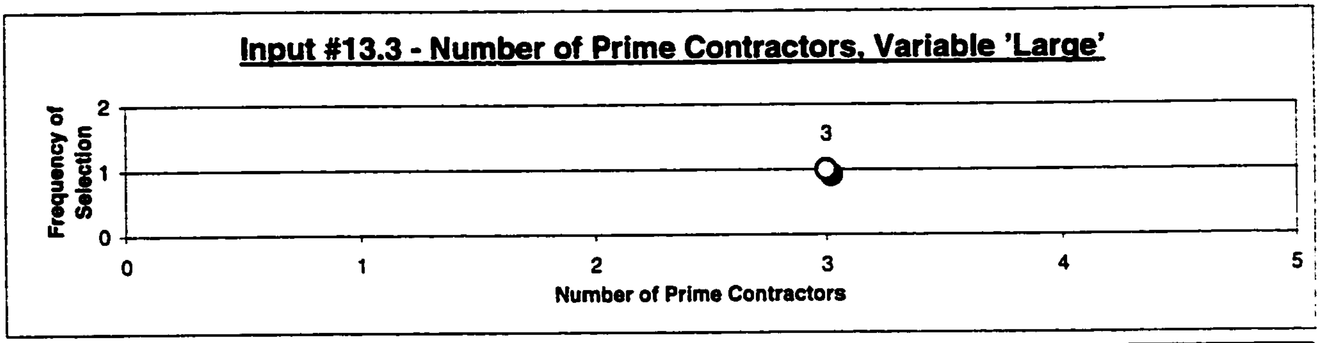


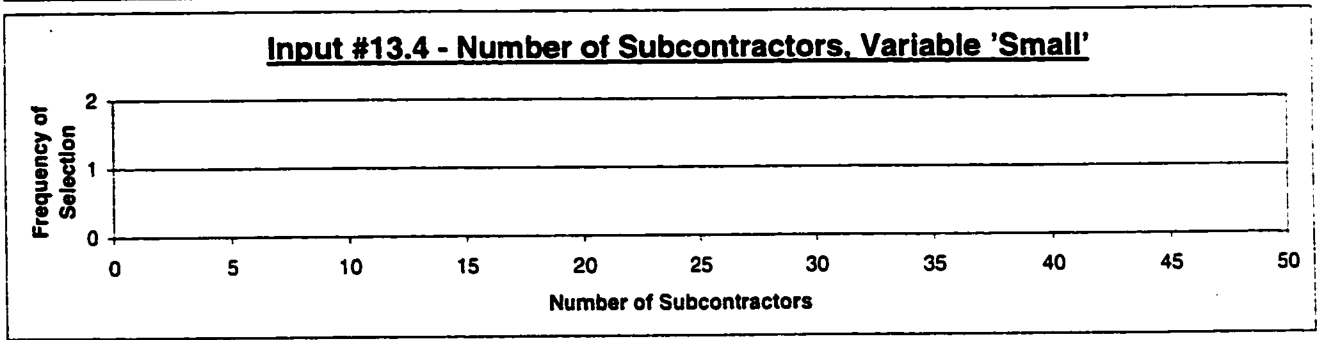
Input #13.3 - Number of Prime Contractors, Variable 'Small'

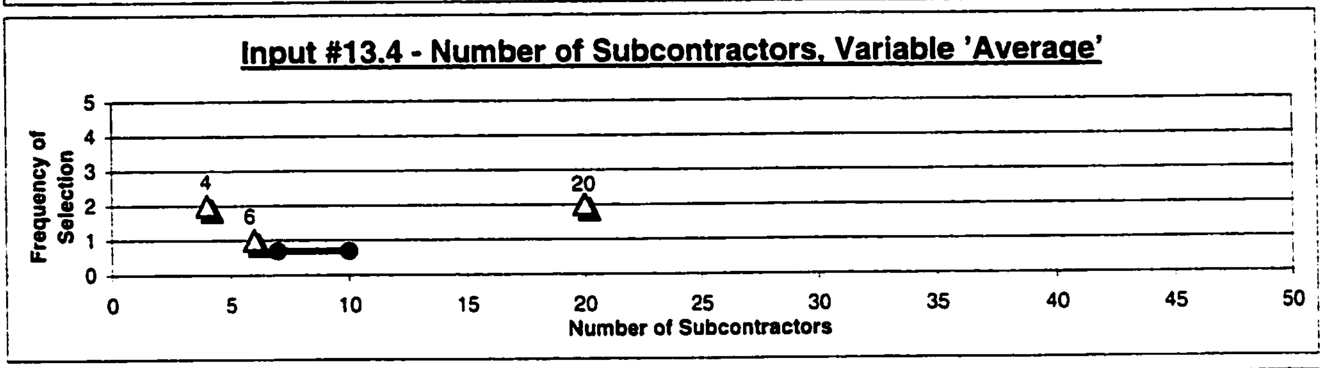


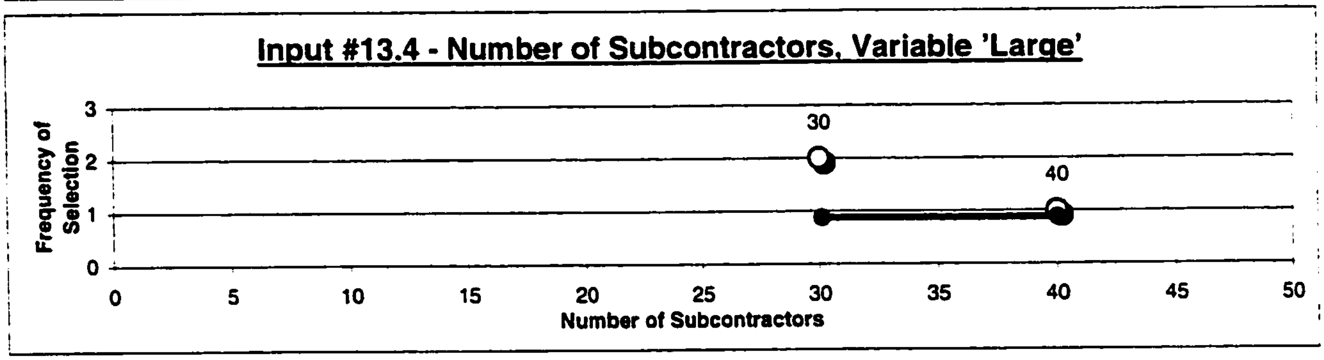
Input #13.3 - Number of Prime Contractors, Variable 'Average'

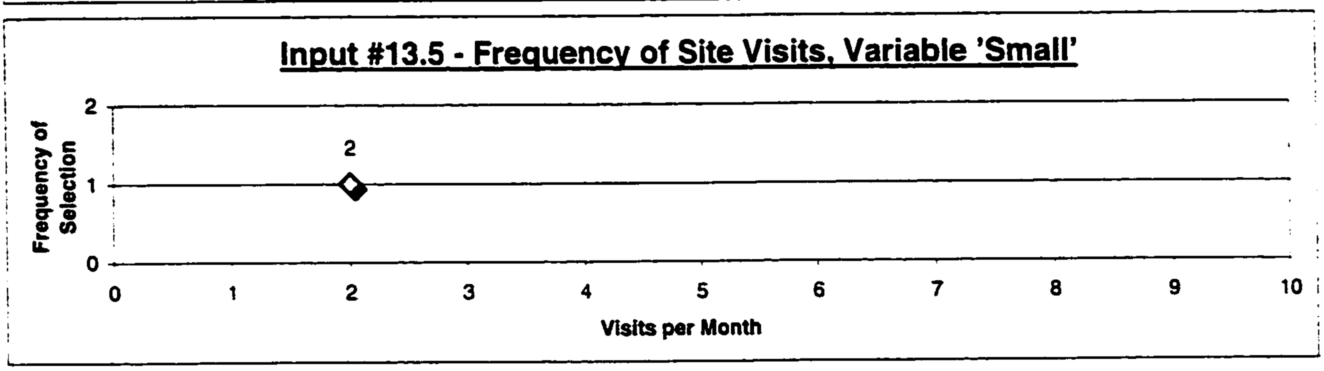


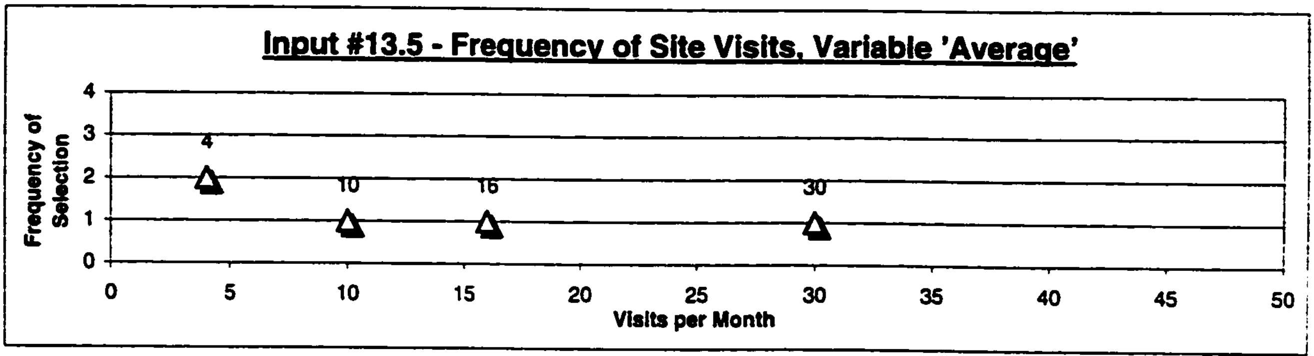

Input #13.3 - Number of Prime Contractors, Variable 'Large'
Frequency of Selection
Number of Prime Contractors
3

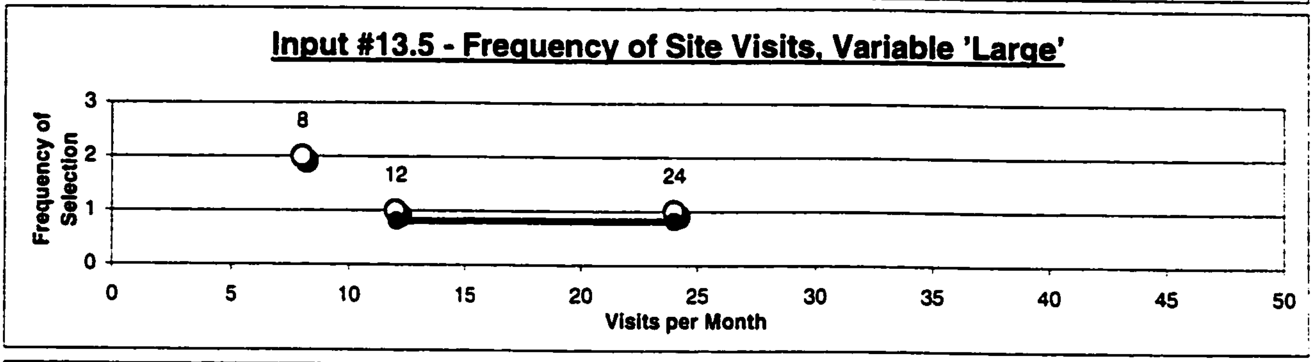

Input #13.4 - Number of Subcontractors, Variable 'Small'
Frequency of Selection
Number of Subcontractors

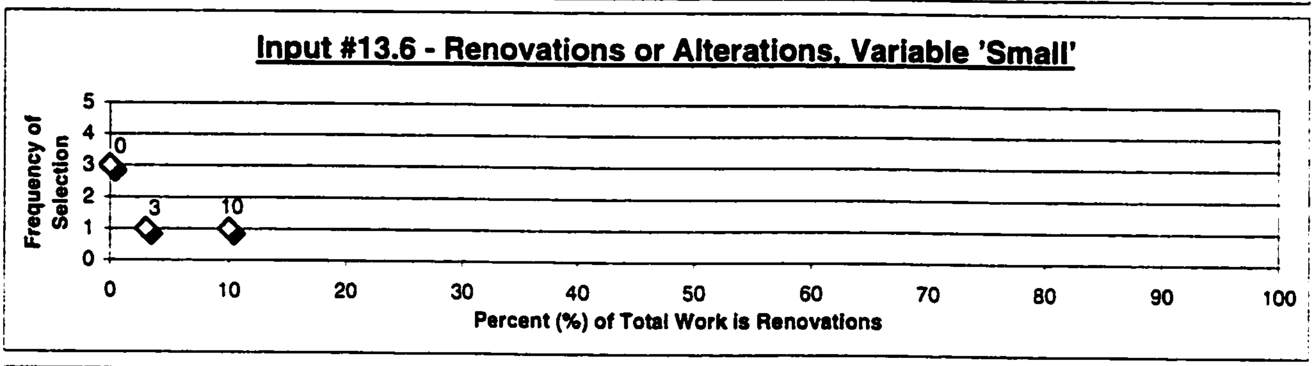

Input #13.4 - Number of Subcontractors, Variable 'Average'
Frequency of Selection
Number of Subcontractors
4
6
20

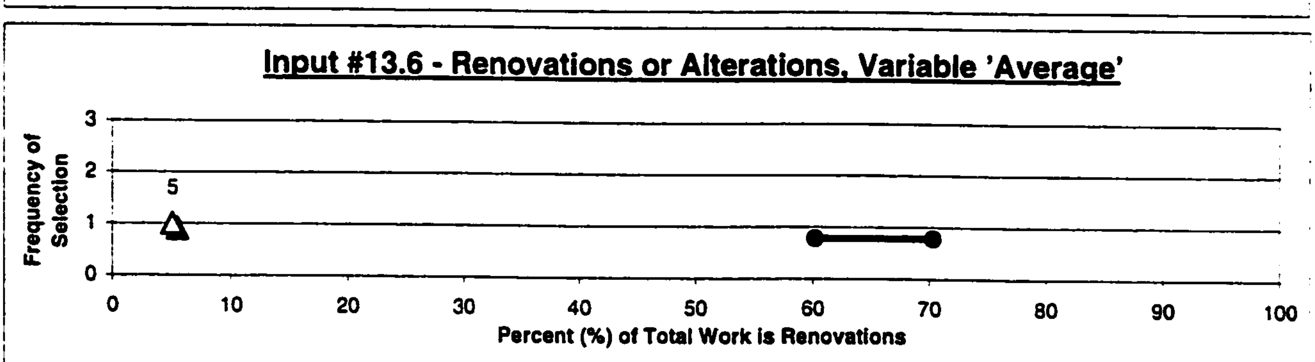

Input #13.4 - Number of Subcontractors, Variable 'Large'
Frequency of Selection
Number of Subcontractors
30
40

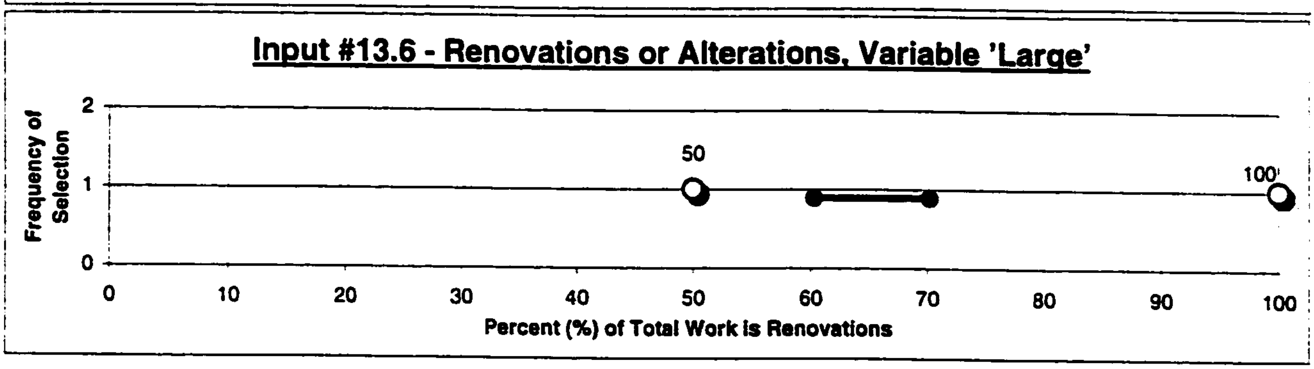

Input #13.5 - Frequency of Site Visits, Variable 'Small'
Frequency of Selection
Visits per Month
2

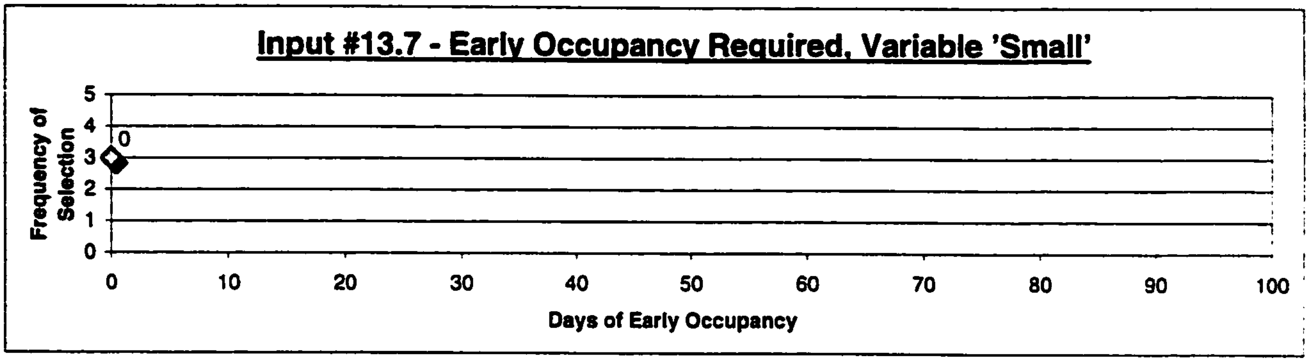

Input #13.5 - Frequency of Site Visits, Variable 'Average'
Frequency of Selection
Visits per Month
4
10
16
30

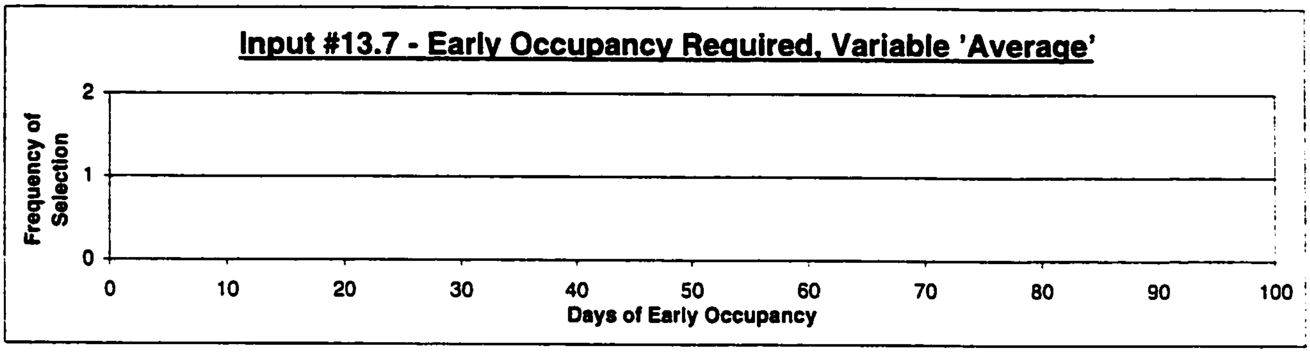

Input #13.5 - Frequency of Site Visits, Variable 'Large'
Frequency of Selection
Visits per Month
8
12
24

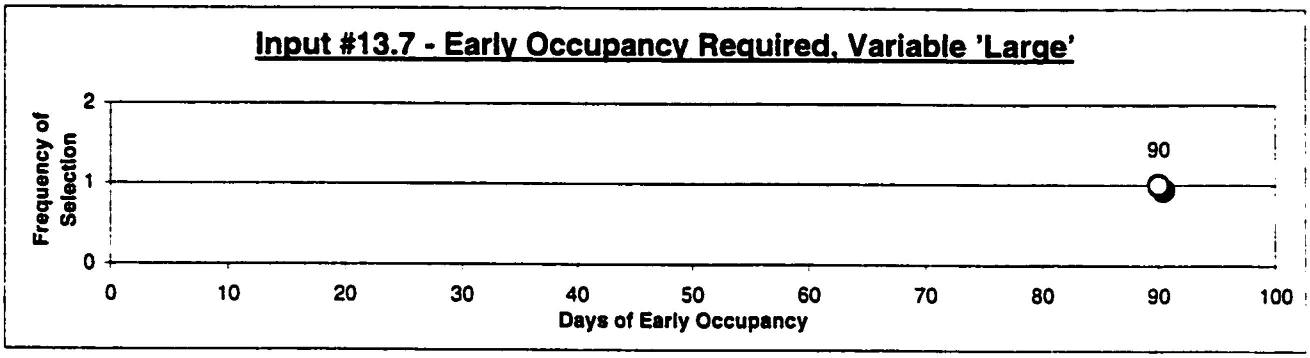

Input #13.6 - Renovations or Alterations, Variable 'Small'
Frequency of Selection
Percent (%) of Total Work is Renovations
0
3
10

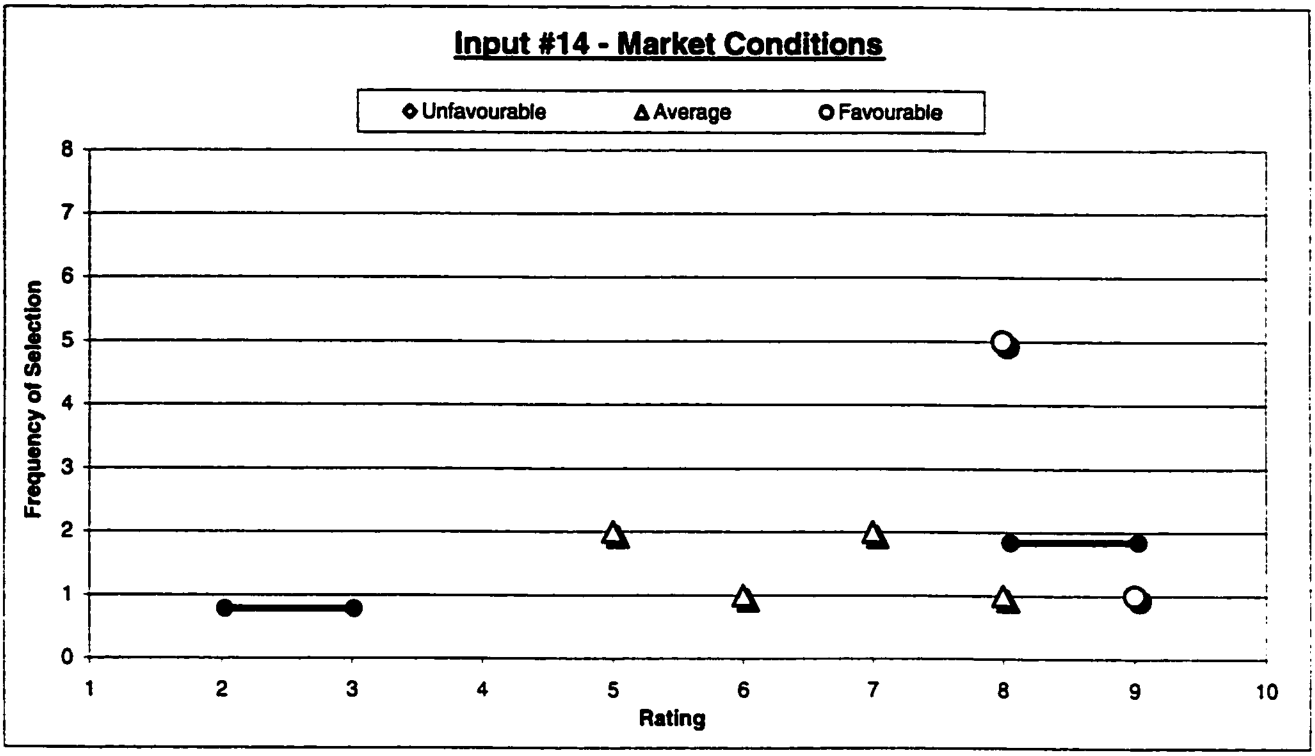

Input #13.6 - Renovations or Alterations, Variable 'Average'
Frequency of Selection
Percent (%) of Total Work is Renovations
5

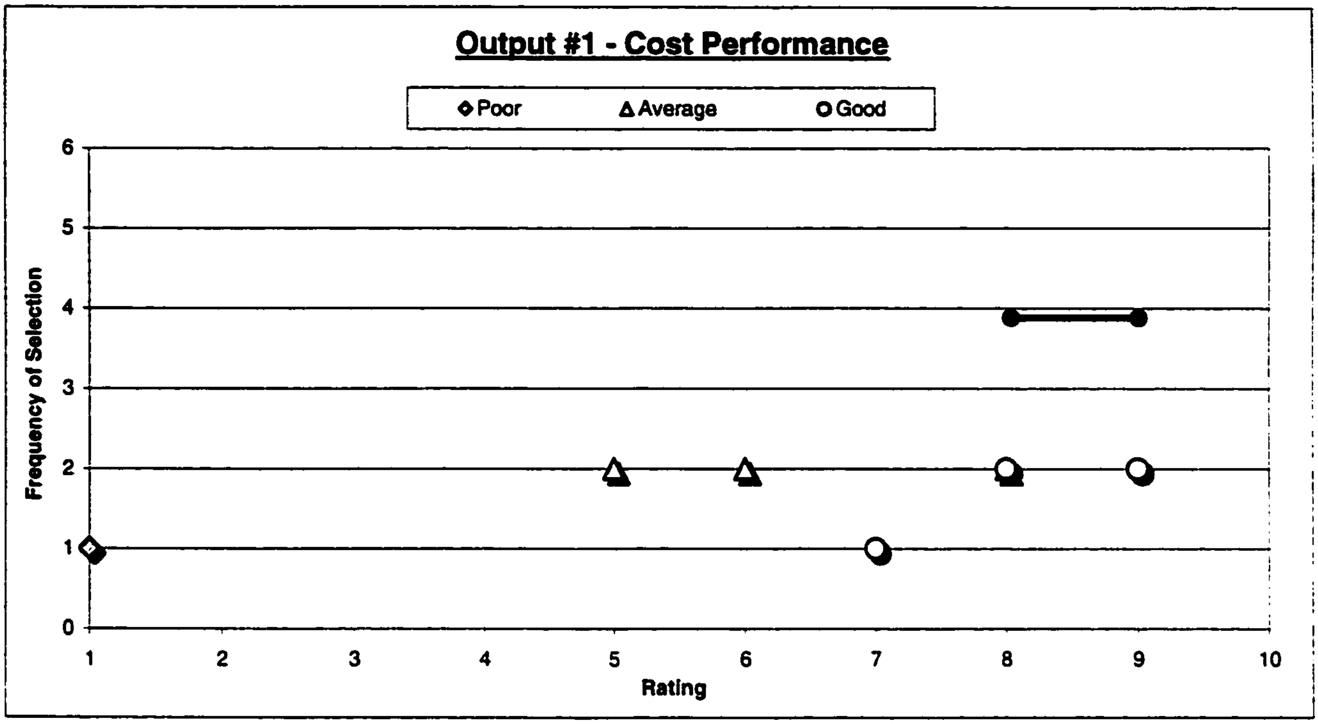

Input #13.6 - Renovations or Alterations, Variable 'Large'
Frequency of Selection
Percent (%) of Total Work is Renovations
50
100

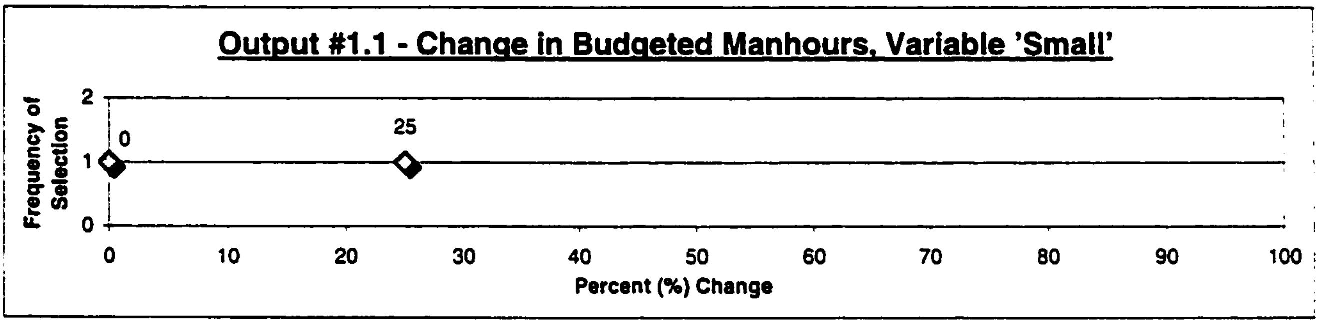

Input #13.7 - Early Occupancy Required, Variable 'Small'
Frequency of Selection
Days of Early Occupancy
0
0 10 20 30 40 50 60 70 80 90 100
0 1 2 3 4 5

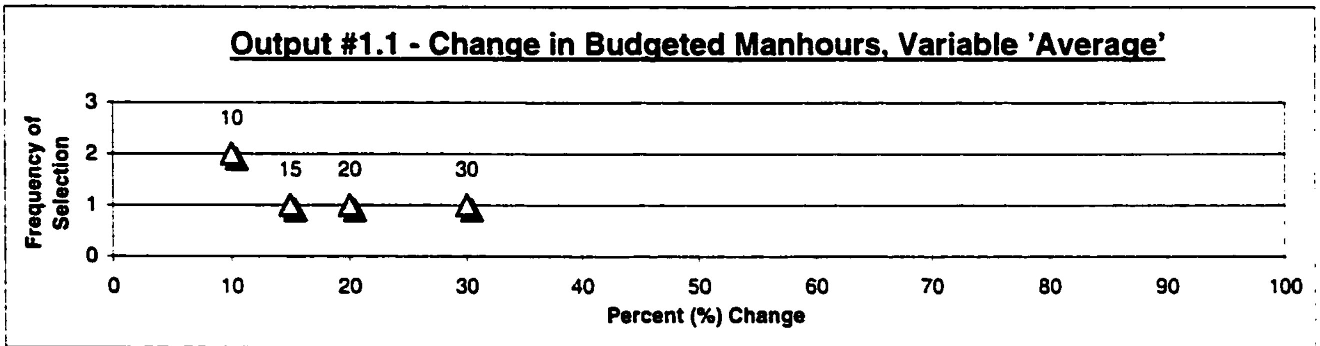

Input #13.7 - Early Occupancy Required, Variable 'Average'
Frequency of Selection
Days of Early Occupancy
0 10 20 30 40 50 60 70 80 90 100
0 1 2

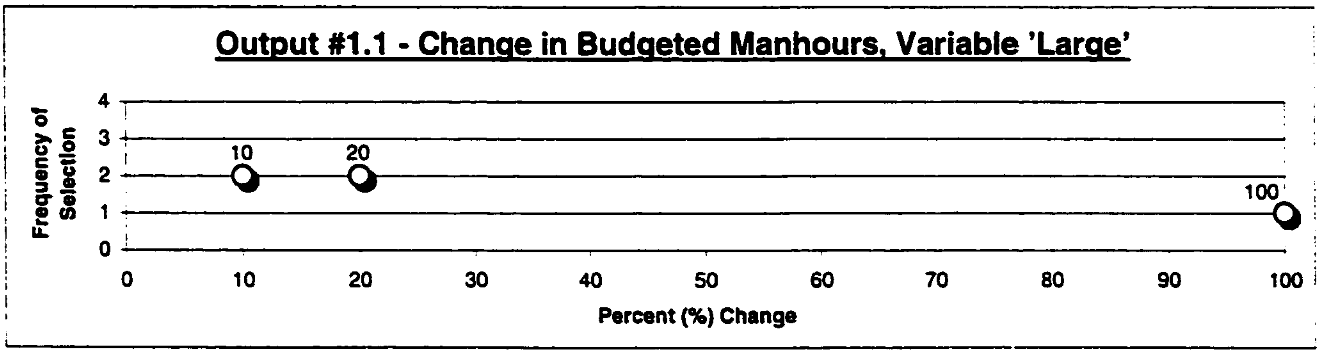

Input #13.7 - Early Occupancy Required, Variable 'Large'
Frequency of Selection
Days of Early Occupancy
90
0 10 20 30 40 50 60 70 80 90 100
0 1 2


Input #14 - Market Conditions
Unfavourable
Average
Favourable
Frequency of Selection
Rating
0
1
2
3
4
5
6
7
8
1
2
3
4
5
6
7
8
9
10

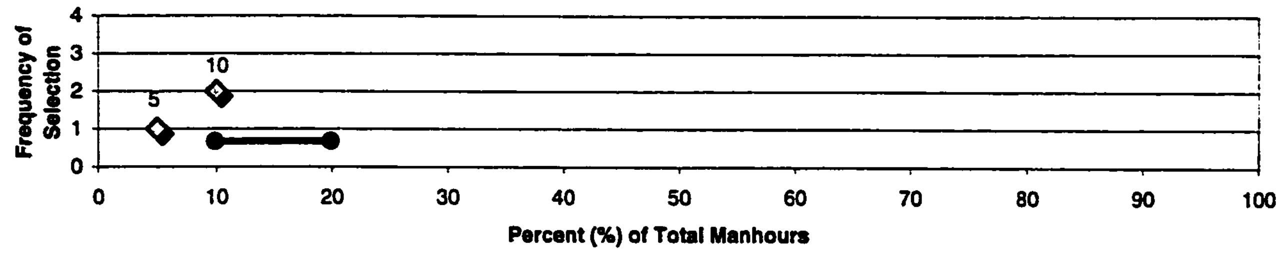

Output #1 - Cost Performance
Poor
Average
Good
Frequency of Selection
Rating
0
1
2
3
4
5
6
1
2
3
4
5
6
7
8
9
10


Output #1.1 - Change in Budgeted Manhours, Variable 'Small'
Frequency of Selection
Percent (%) Change
0
25

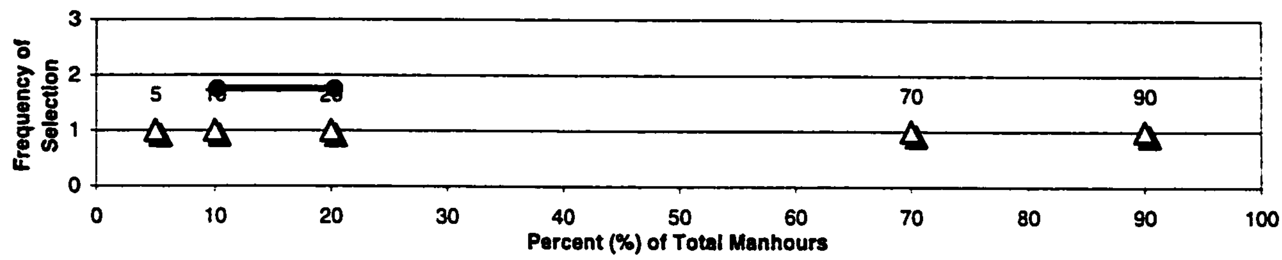

Output #1.1 - Change in Budgeted Manhours, Variable 'Average'
Frequency of Selection
Percent (%) Change
10
15
20
30


Output #1.1 - Change in Budgeted Manhours, Variable 'Large'
Frequency of Selection
Percent (%) Change
10
20
100

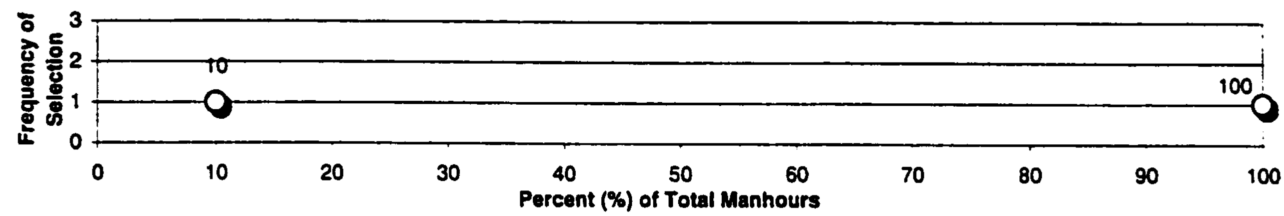
Output #1.2 - Manhours Due to Owner Changes, Variable 'Small'



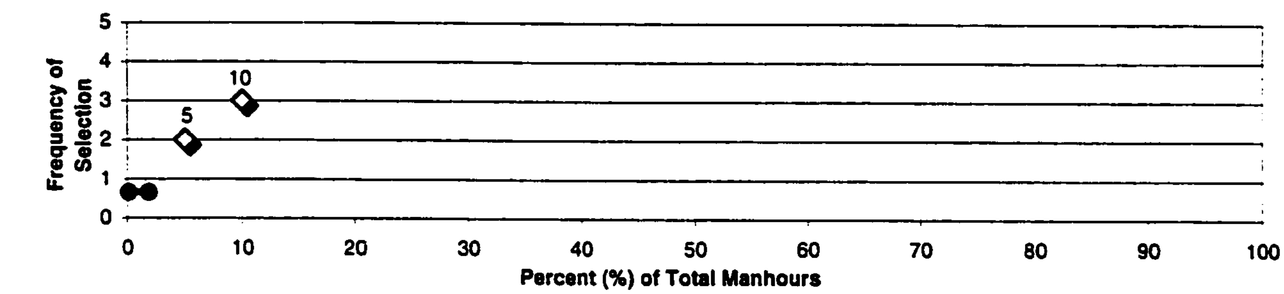
Output #1.2 - Manhours Due to Owner Changes, Variable 'Average'



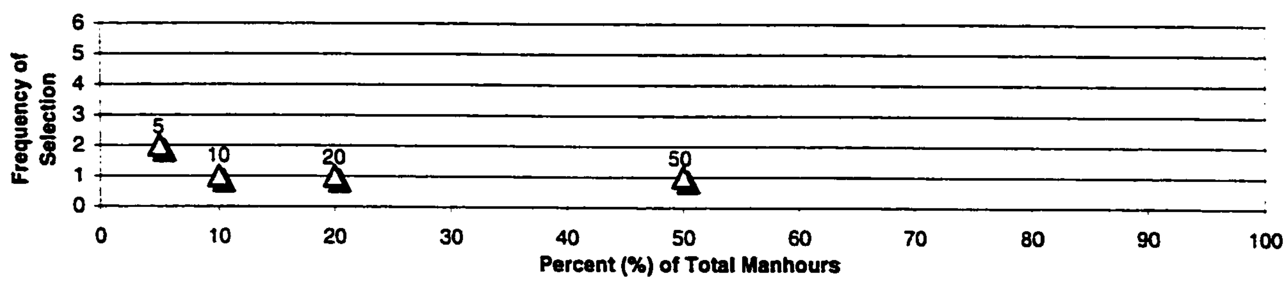
Output #1.2 - Manhours Due to Owner Changes, Variable 'Large'

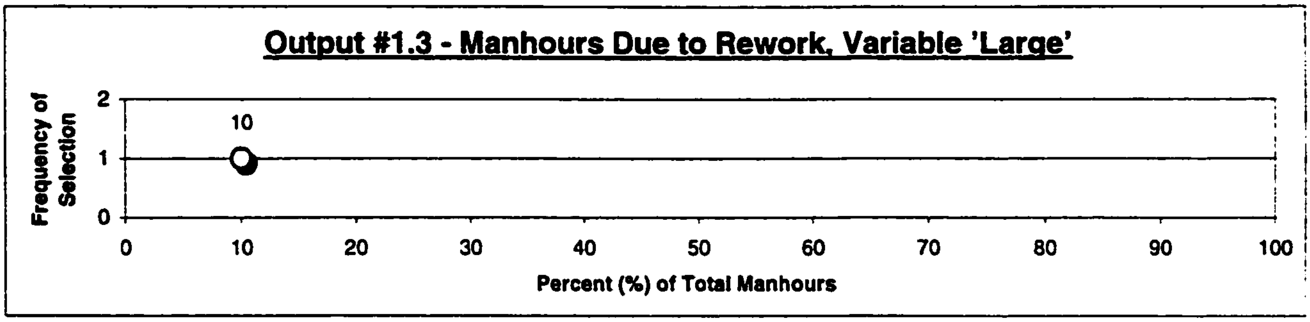


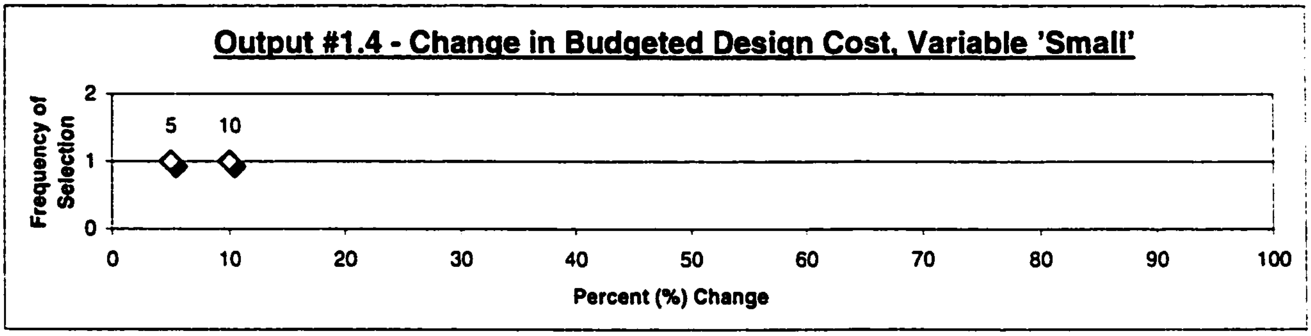
Output #1.3 - Manhours Due to Rework, Variable 'Small'

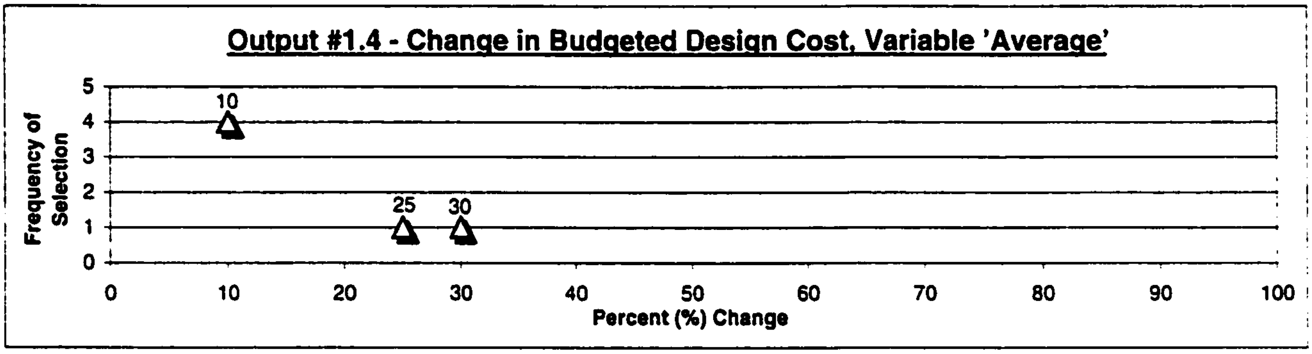


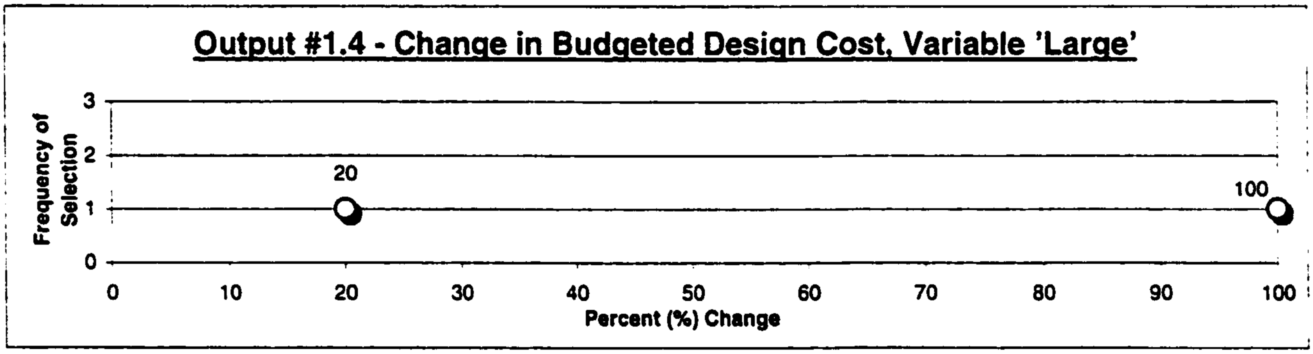
Output #1.3 - Manhours Due to Rework, Variable 'Average'

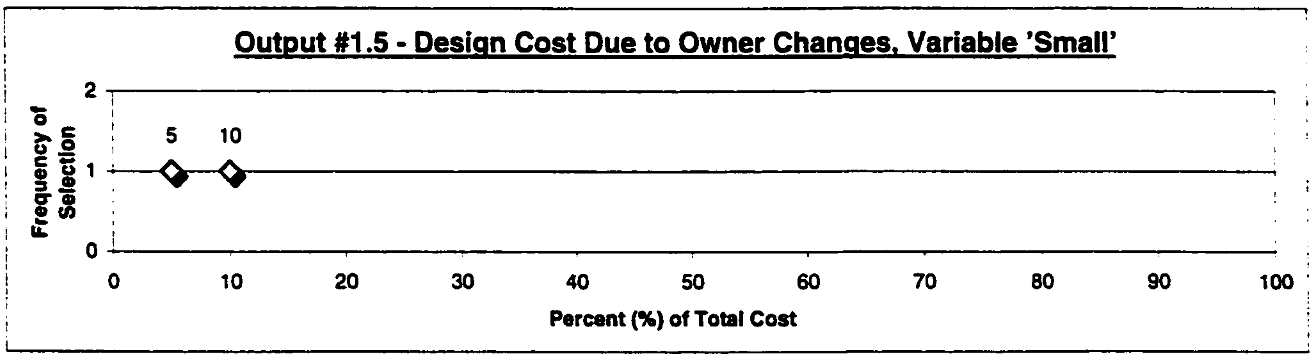


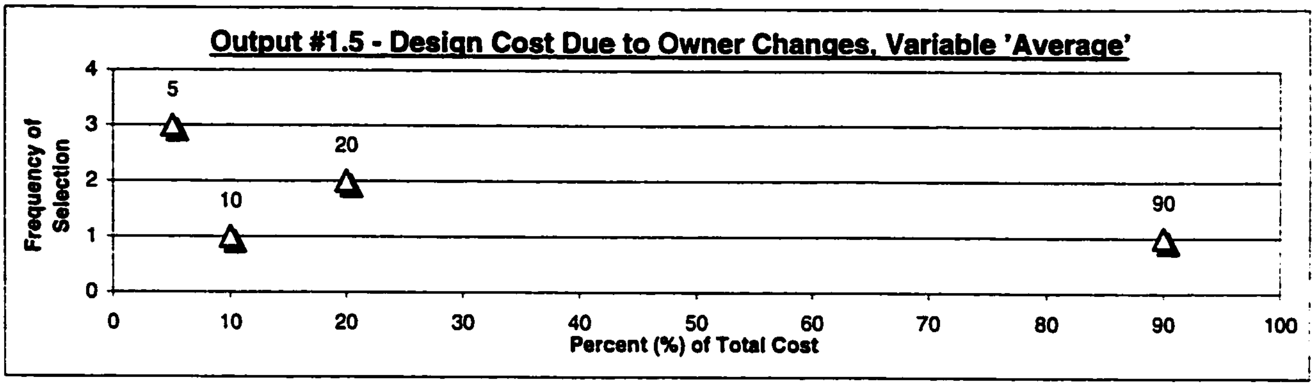

Output #1.3 - Manhours Due to Rework, Variable 'Large'
Frequency of Selection
Percent (%) of Total Manhours
10


Output #1.4 - Change in Budgeted Design Cost, Variable 'Small'
Frequency of Selection
Percent (%) Change
5
10

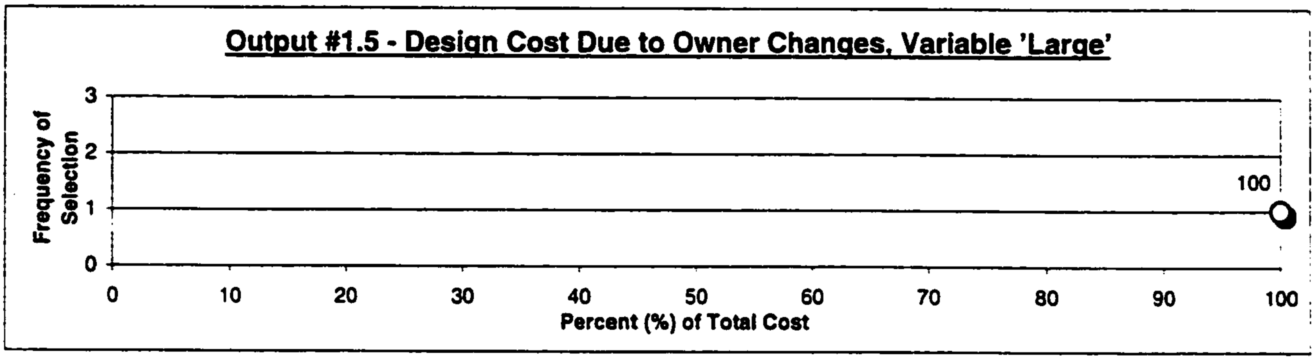

Output #1.4 - Change in Budgeted Design Cost, Variable 'Average'
Frequency of Selection
Percent (%) Change
10
25
30


Output #1.4 - Change in Budgeted Design Cost, Variable 'Large'
Frequency of Selection
Percent (%) Change
20
100

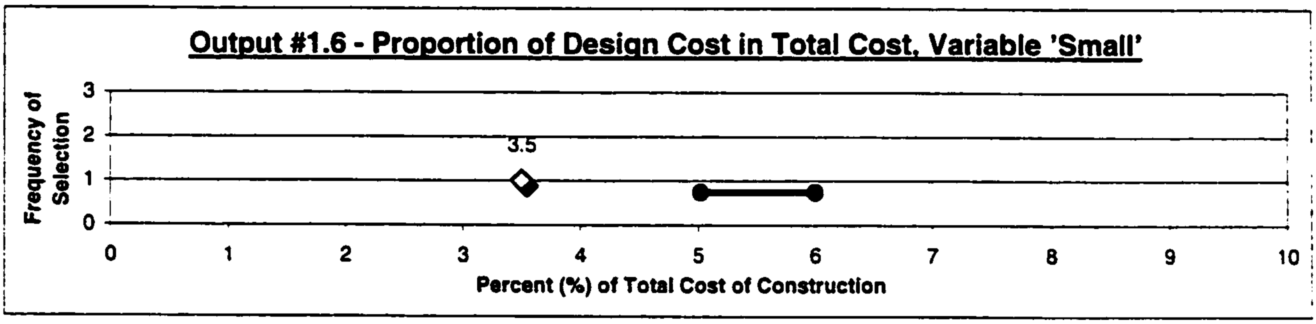

Output #1.5 - Design Cost Due to Owner Changes, Variable 'Small'
Frequency of Selection
Percent (%) of Total Cost
5
10



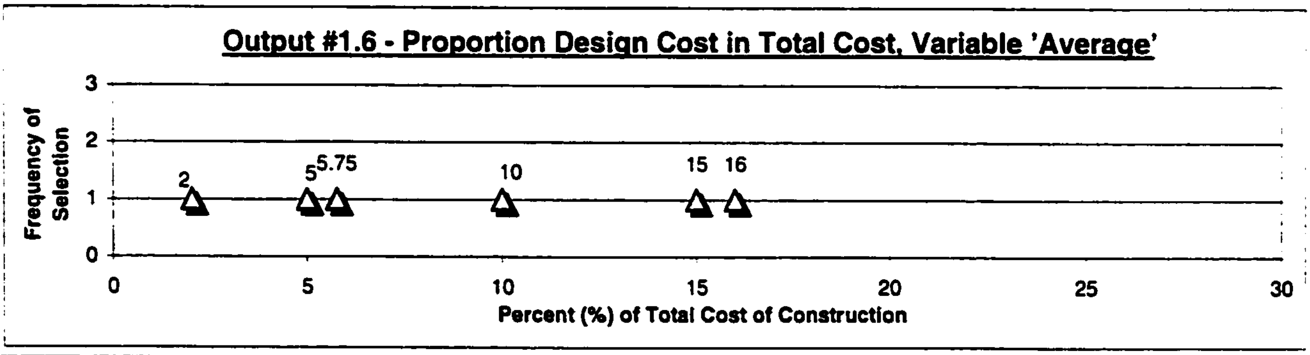
Output #1.5 - Design Cost Due to Owner Changes, Variable 'Average'
Frequency of Selection
Percent (%) of Total Cost
5
10
20
90



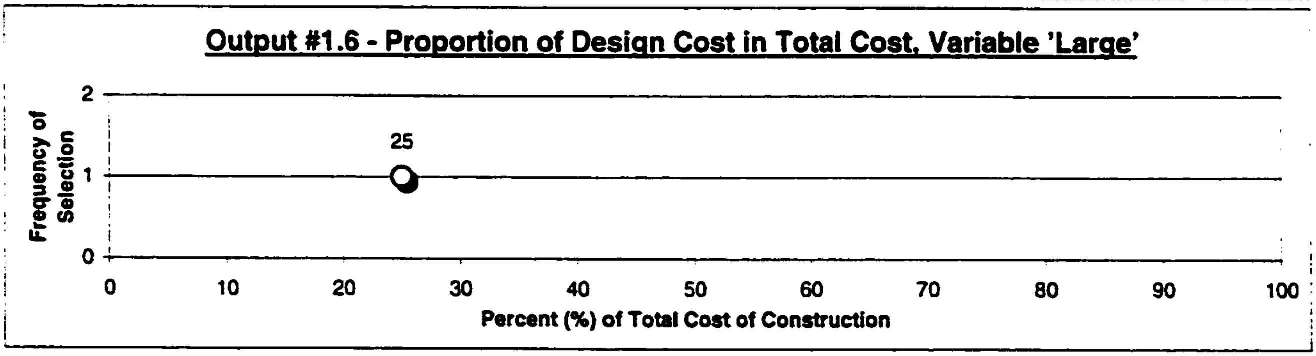
Output #1.5 - Design Cost Due to Owner Changes, Variable 'Large'
Frequency of Selection
Percent (%) of Total Cost
100



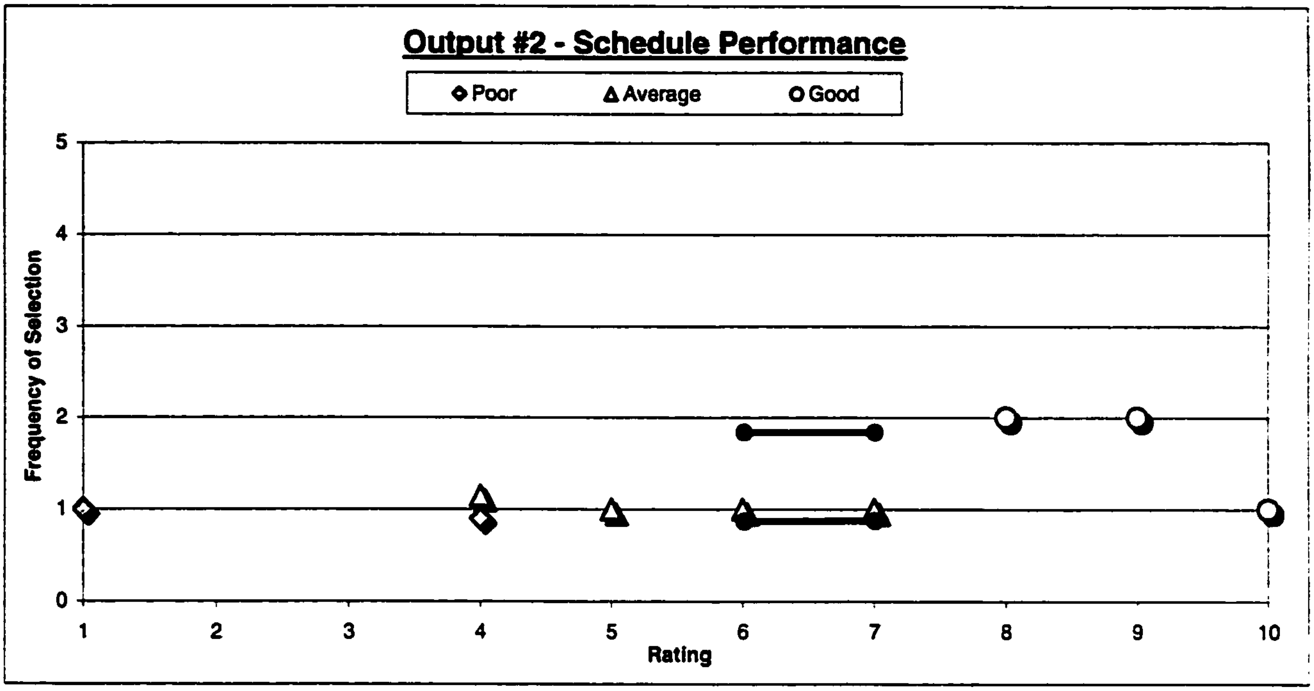
Output #1.6 - Proportion of Design Cost in Total Cost, Variable 'Small'
Frequency of Selection
Percent (%) of Total Cost of Construction
3.5



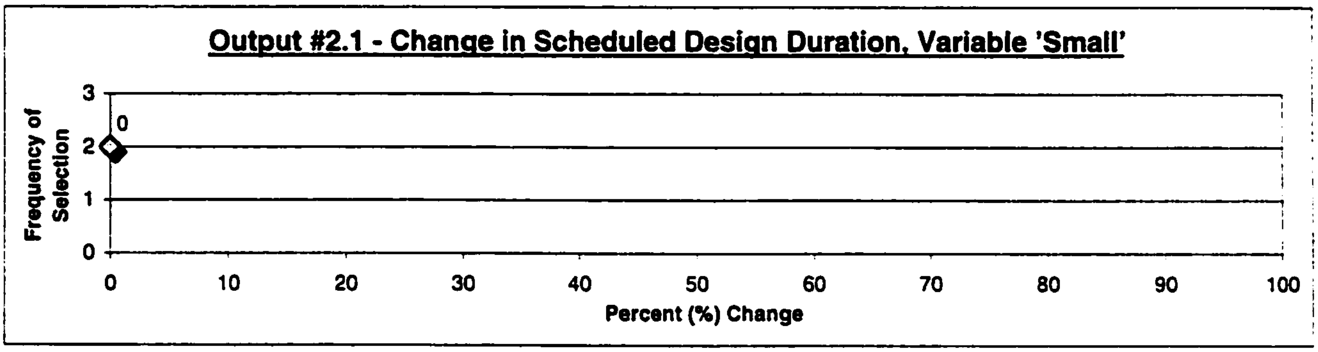
Output #1.6 - Proportion Design Cost in Total Cost, Variable 'Average'
Frequency of Selection
Percent (%) of Total Cost of Construction
2
5
5.75
10
15
16



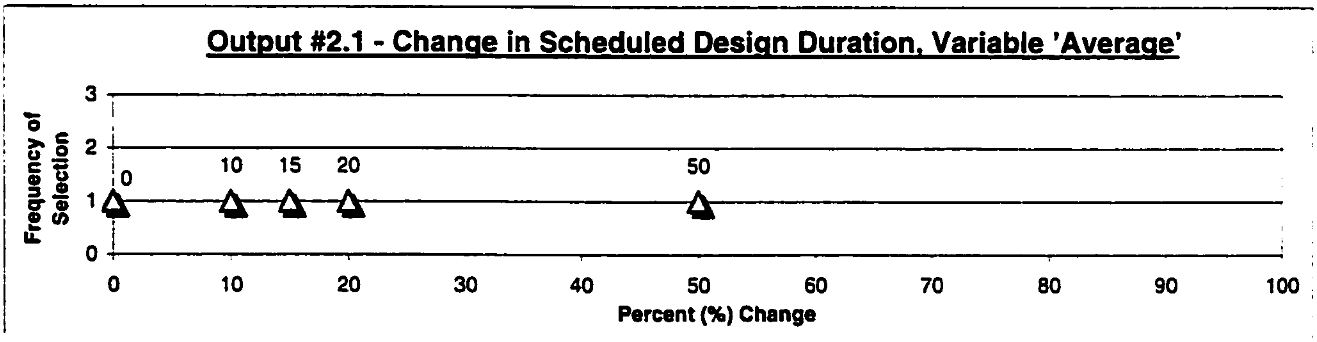
Output #1.6 - Proportion of Design Cost in Total Cost, Variable 'Large'
Frequency of Selection
Percent (%) of Total Cost of Construction
25



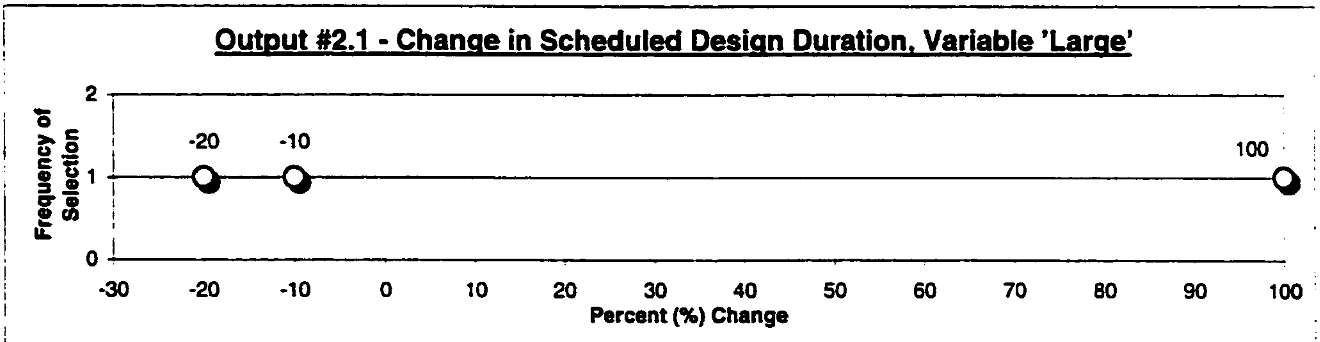
Output #2 - Schedule Performance
Poor
Average
Good
Frequency of Selection
Rating
0
1
2
3
4
5
1
2
3
4
5
6
7
8
9
10



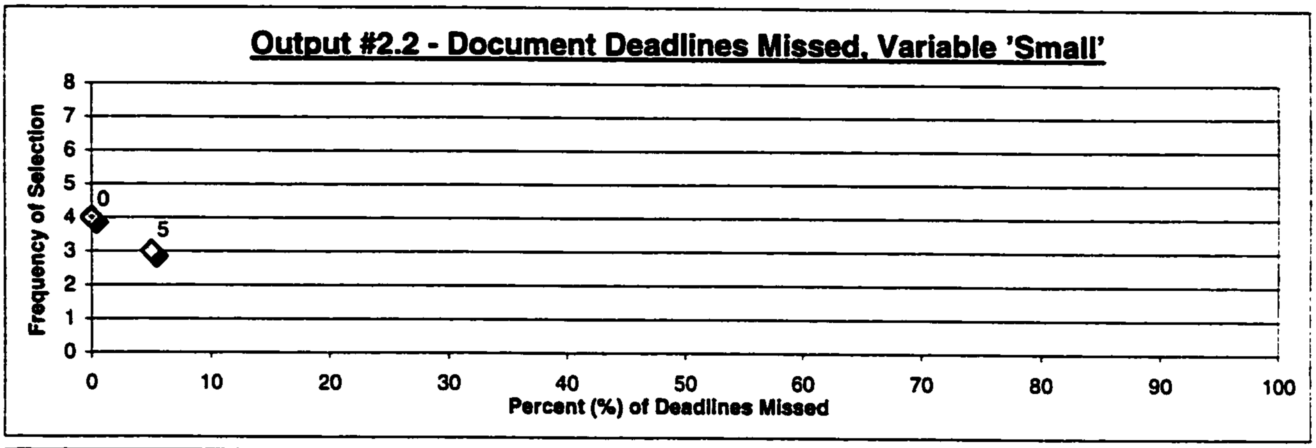
Output #2.1 - Change in Scheduled Design Duration, Variable 'Small'
Frequency of Selection
Percent (%) Change
0



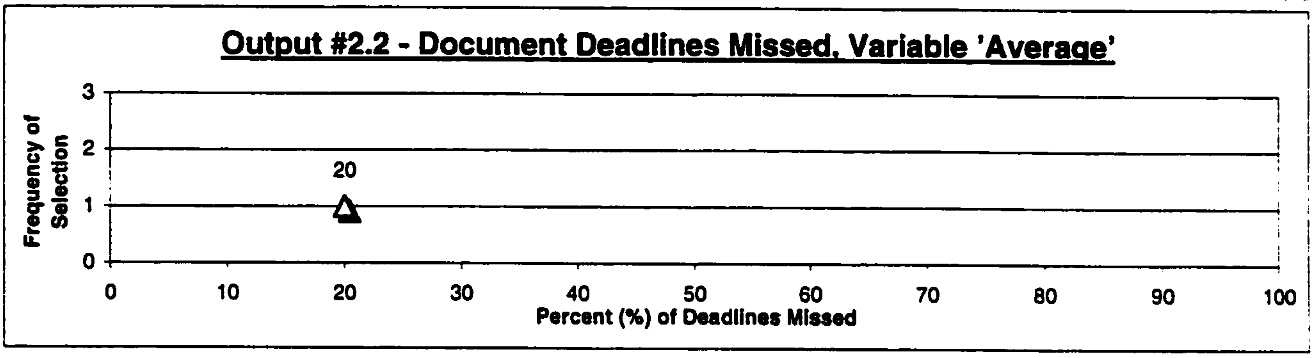
Output #2.1 - Change in Scheduled Design Duration, Variable 'Average'
Frequency of Selection
Percent (%) Change
0
10
15
20
50



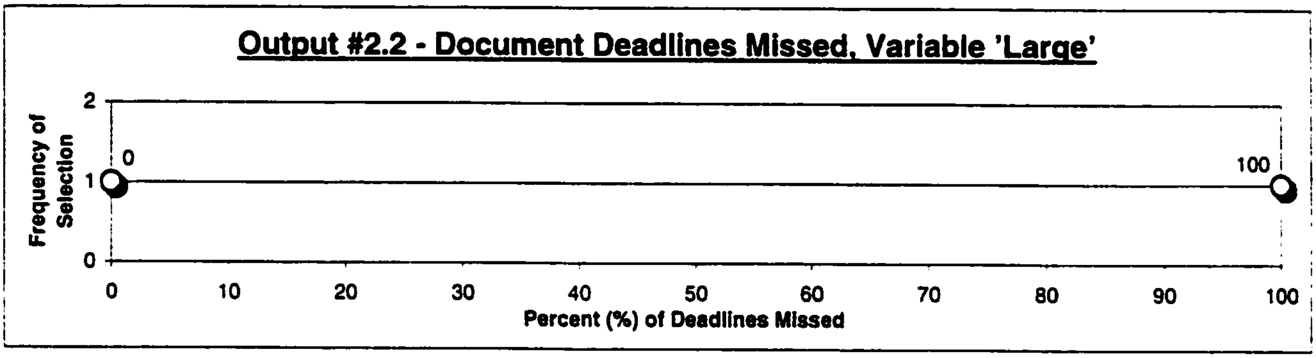
Output #2.1 - Change in Scheduled Design Duration, Variable 'Large'
Frequency of Selection
Percent (%) Change
-20
-10
100



Output #2.2 - Document Deadlines Missed, Variable 'Small'
Frequency of Selection
Percent (%) of Deadlines Missed
0
5

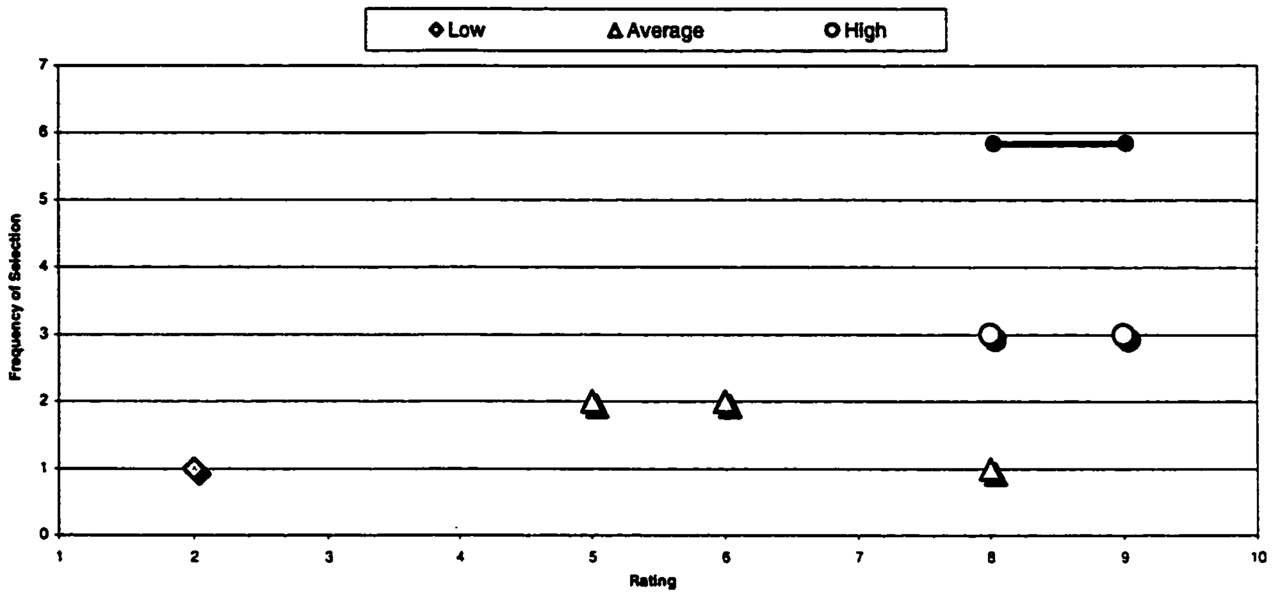


Output #2.2 - Document Deadlines Missed, Variable 'Average'
Frequency of Selection
Percent (%) of Deadlines Missed
20

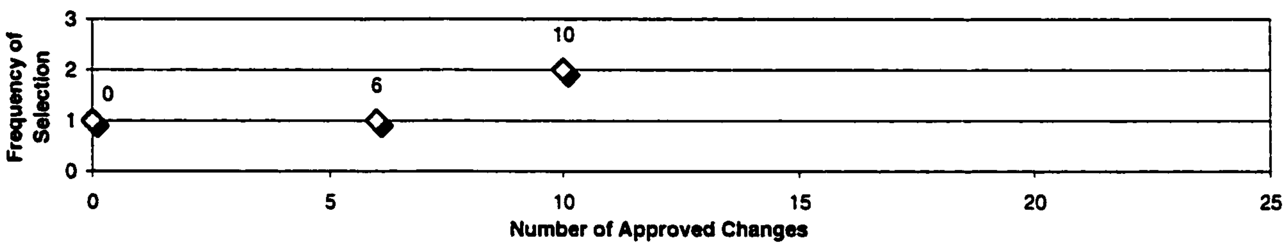


Output #2.2 - Document Deadlines Missed, Variable 'Large'
Frequency of Selection
Percent (%) of Deadlines Missed
0
100

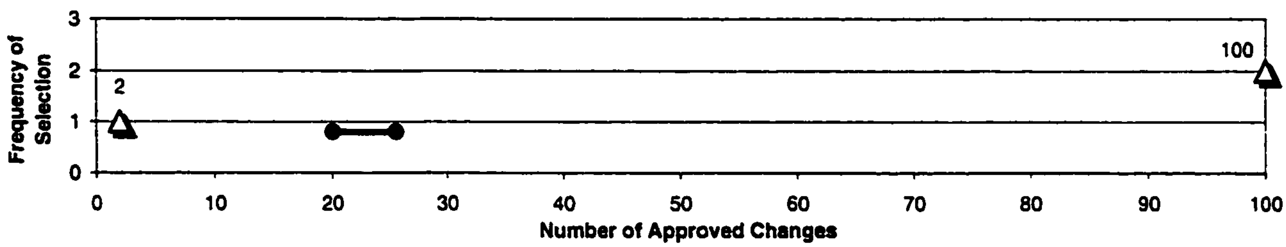
Output #3 - Overall Accuracy of Design



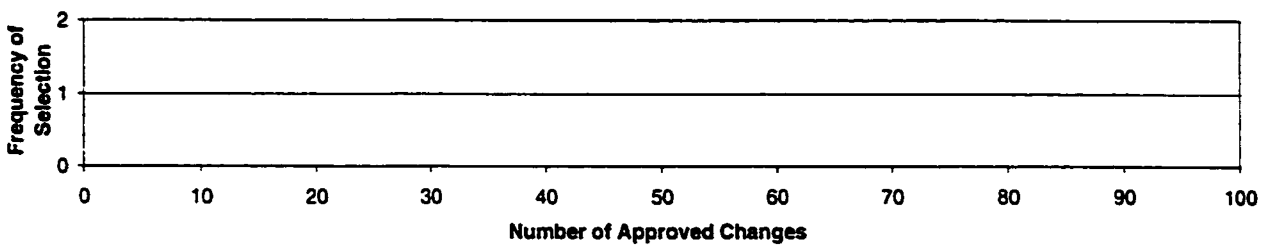
Output #3.1 - Approved Changes, Variable 'Small'



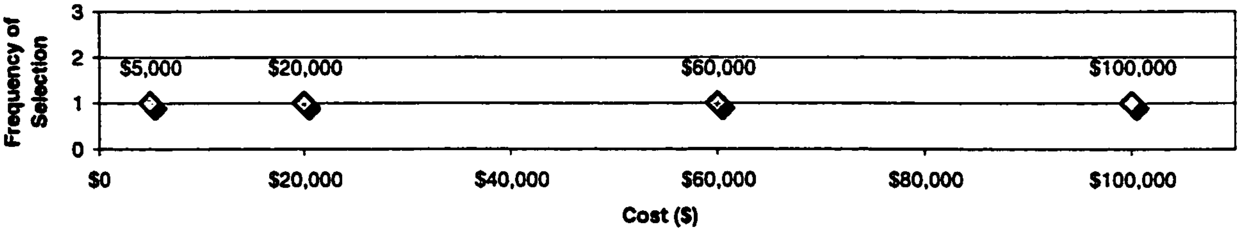
Output #3.1 - Approved Changes, Variable 'Average'



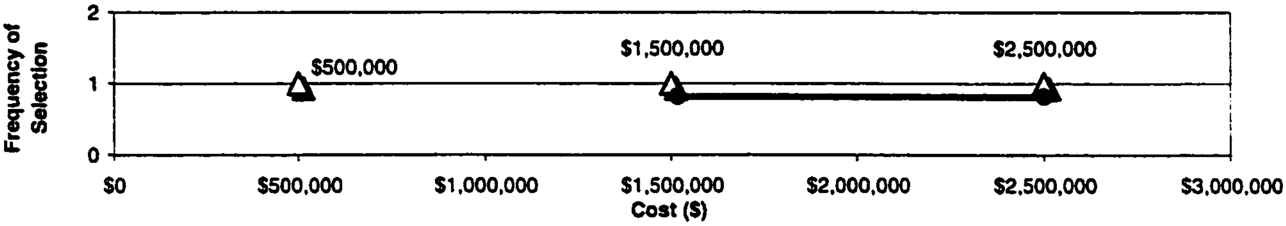
Output #3.1 - Approved Changes, Variable 'Large'



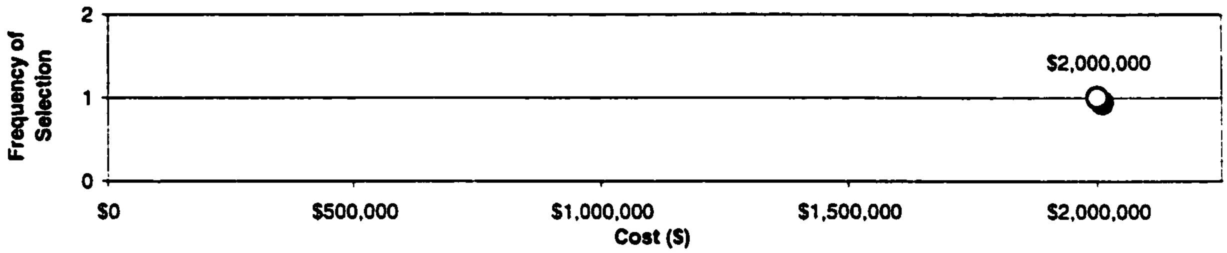
Output #3.2 - Cost of Approved Changes, Variable 'Small'



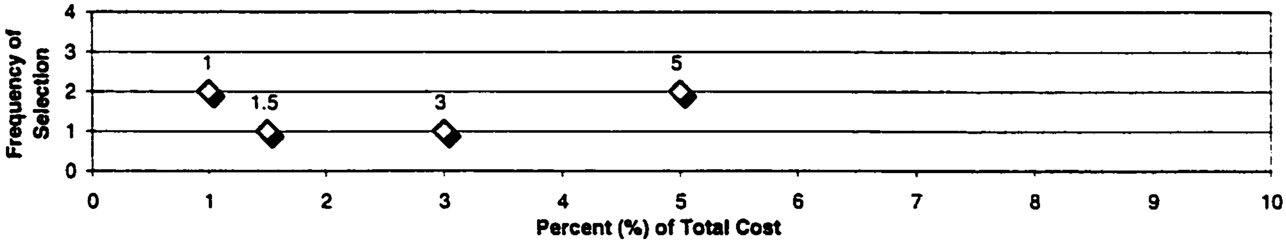
Output #3.2 - Cost of Approved Changes, Variable 'Average'



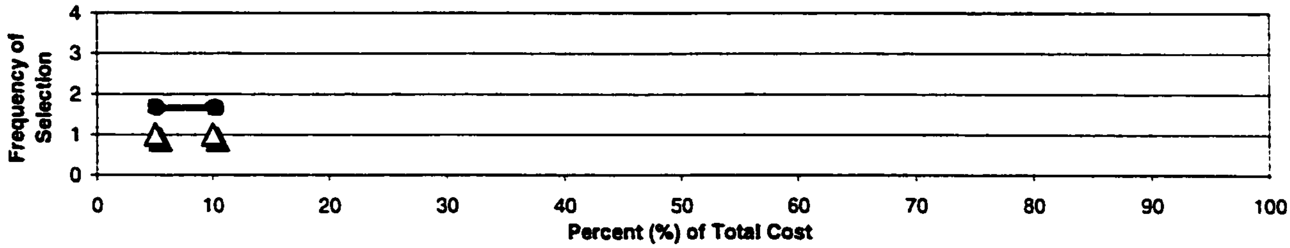
Output #3.2 - Cost of Approved Changes, Variable 'Large'

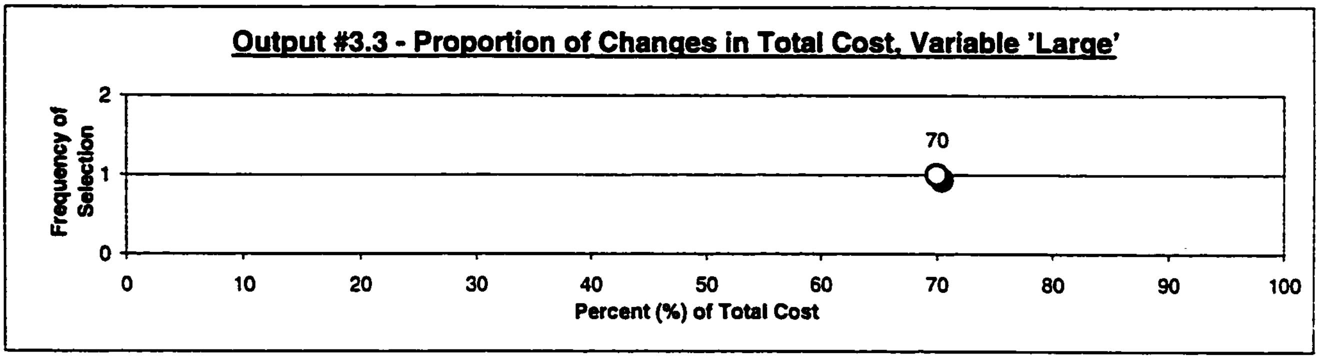


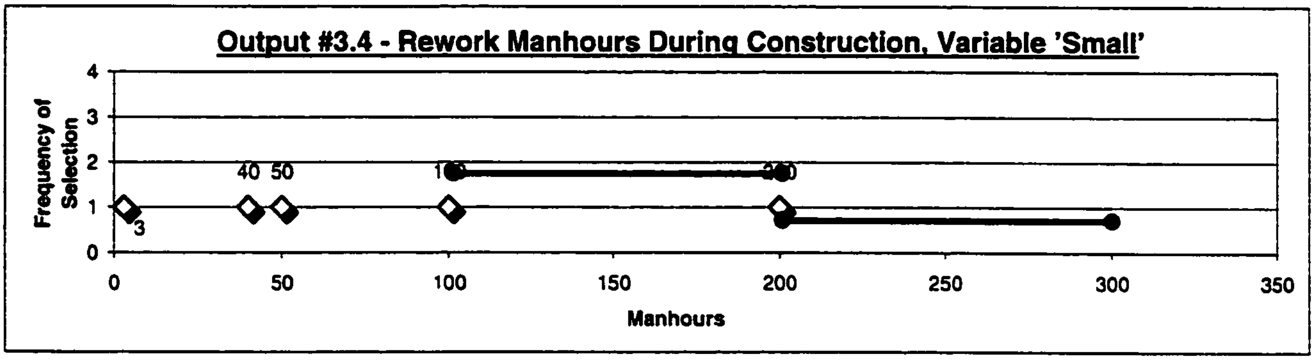
Output #3.3 - Proportion of Changes in Total Cost, Variable 'Small'

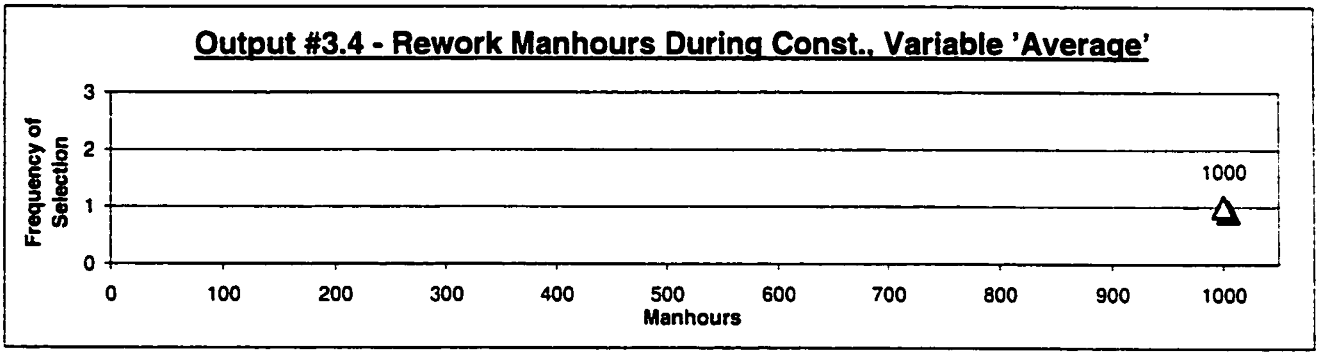


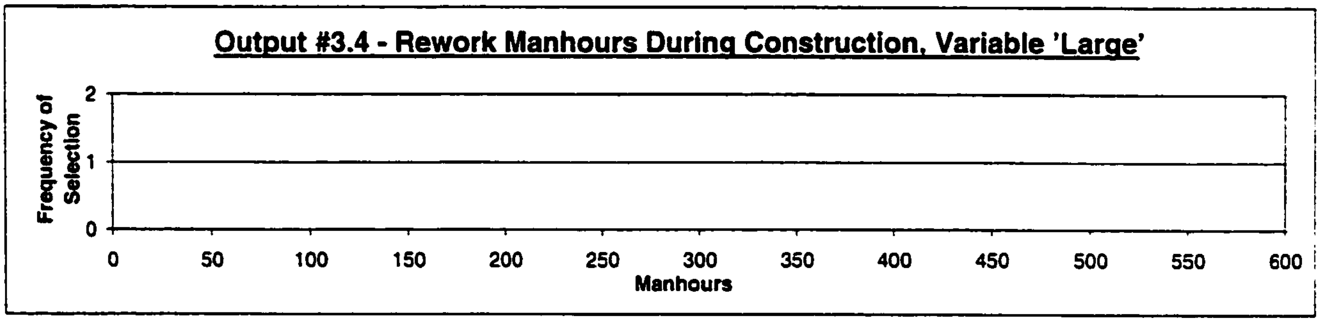
Output #3.3 - Proportion of Changes in Total Cost, Variable 'Average'

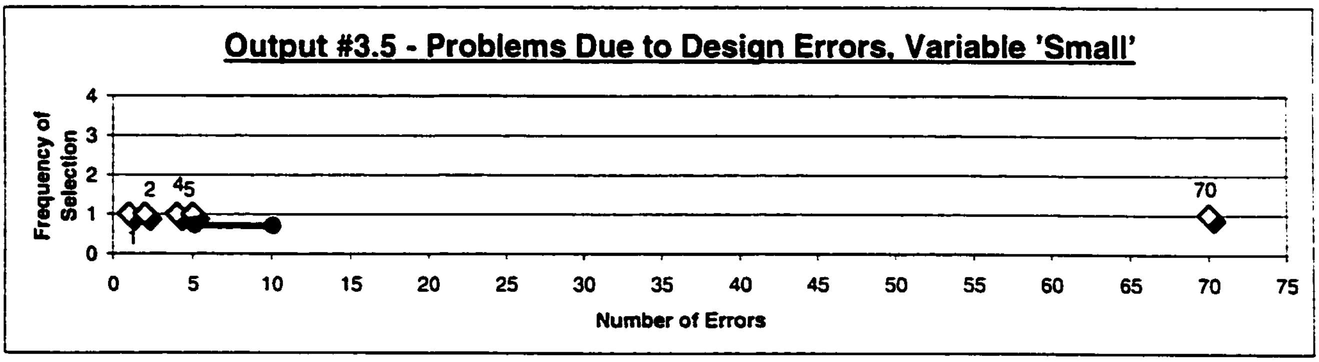


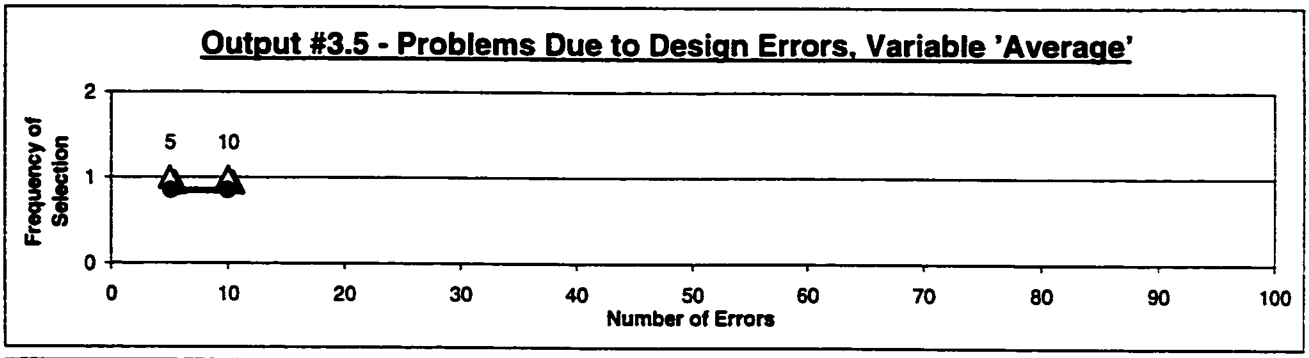

Output #3.3 - Proportion of Changes in Total Cost, Variable 'Large'
Frequency of Selection
Percent (%) of Total Cost
70


Output #3.4 - Rework Manhours During Construction, Variable 'Small'
Frequency of Selection
Manhours
3
40 50

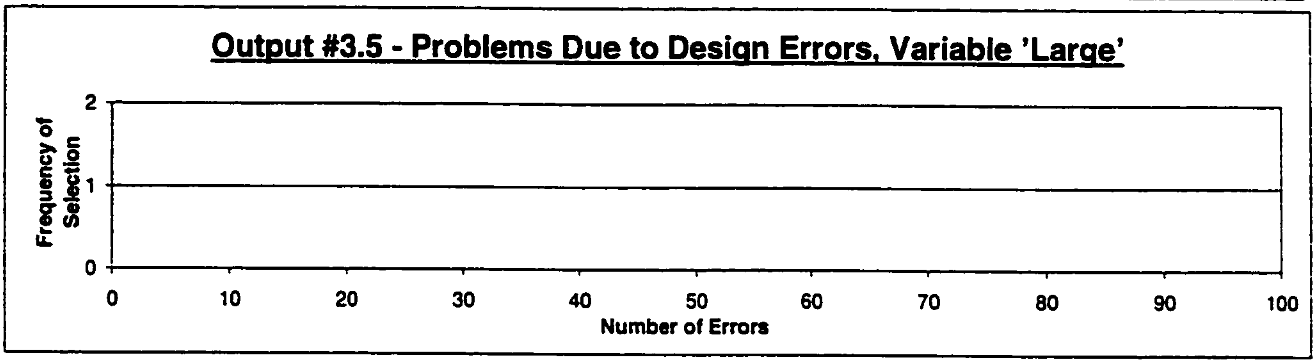

Output #3.4 - Rework Manhours During Const., Variable 'Average'
Frequency of Selection
Manhours
1000


Output #3.4 - Rework Manhours During Construction, Variable 'Large'
Frequency of Selection
Manhours


Output #3.5 - Problems Due to Design Errors, Variable 'Small'
Frequency of Selection
Number of Errors
2
70



Output #3.5 - Problems Due to Design Errors, Variable 'Average'
Frequency of Selection
Number of Errors
5
10
0
1
2
0
10
20
30
40
50
60
70
80
90
100



Output #3.5 - Problems Due to Design Errors, Variable 'Large'
Frequency of Selection
Number of Errors
0
1
2
0
10
20
30
40
50
60
70
80
90
100

Appendix 4: Parameters for the Derived Membership Functions (Trial 1)

No.	Name	Parameters
Input 1	Small-average	(0, 0, 5, 10)
	Large	(7, 8, 10, 10)
Input 1.1	Small-average	(0, 0, 40, 80)
	Large	(30, 90, 900, 900)
Input 1.2	Small-average	(0, 0, 3, 28.37)
	Large	(7.845, 23.54, 235.35, 235.35)
Input 1.3	Small- average	(0, 0, 5, 131.7)
	Large	(33.68, 101.025, 1010.2, 1010.2)
Input 2	Low	(0, 0, 3, 6)
	Average-high	(3.8, 8, 10, 10)
Input 2.1	Small	(0, 0, 2, 4)
	Average-large	(1.833, 10, 100, 100)
Input 2.2	Small-average	(0, 0, 10, 130)
	Large	(35, 105, 1050, 1050)
Input 3	Poor	(0, 0, 4, 8)
	Average	(1, 5, 11)
	Good	(6.5, 8, 10, 10)
Input 3.1	Small	(0, 0, 5.33, 10.66)
	Average-large	(6, 10, 10)
Input 3.2	Small	(0, 0, 2, 4)
	Average-large	(1.166, 6, 120, 120)
Input 3.3	Small- average	(0, 0, 50, 99.25)
	Large	(13.3, 40, 400, 400)
Input 3.4	Poor	(0, 0, 4.17, 8.34)
	Average-good	(4.5, 8, 10, 10)
Input 3.5	Small-average	(0, 0, 10, 40)
	Large	(5, 15, 150, 150)
Input 3.6	poor	(0, 0, 3.67, 7.33)
	Average-good	(4.67, 8, 10, 10)
Input 3.7	Small	(0, 0, 10.2, 20.4)
	Average-large	(5.6, 25, 250, 250)
Input 3.8	Small- average	(0, 0, 11.66)
	Large	(1, 3, 30, 30)
Input 3.9	Low	(0, 0, 2, 4)
	Average-high	(4, 10, 10)
Input 4	Small	(0, 0, 3.33, 6.67)
	Average	(3, 7, 8.33)
	Large	(6, 8, 10, 10)
Input 4.1	Small	(0, 0, 2.5, 5)
	Average-large	(-30.8, 32, 320, 320)
Input 4.2	Short-average	(0, 0, 12, 24)
	Long	(12, 36, 360, 360)

Input 4.3	Small-average	(0, 0, 3.5, 116.5)
	Large	(2, 6, 600, 600)
Input 5	Small	(0, 0, 4.083, 8.167)
	Average-large	(4.25, 8, 10, 10)
Input 5.1	Small	(0, 0, 10, 20)
	Average – large	(10.15, 40, 400, 400)
Input 5.2	Small	(0, 0, 10, 20)
	Average- large	(2.22, 40, 400, 400)
Input 6	Low	(0, 0, 2, 5)
	Average	(-1, 7, 15)
	High	(6.5, 8, 10, 10)
Input 6.1	Low	(0, 0, 2, 8)
	Average – high	(3.55, 8, 10, 10)
Input 6.2	Low	(0, 0, 2, 6)
	Average-high	(3.2, 8, 10, 10)
Input 6.3	Small	(0, 0, 10, 20)
	Average	(-6.71, 50, 94.77)
	Large	(35, 70, 100, 100)
Input 6.4	Small	(0, 0, 4, 8)
	Average-large	(-7, 8, 10, 10)
INPUT 7	Low	(0, 0, 4, 8)
	Average	(4.8, 6, 10)
	High	(4, 8, 10, 10)
Input 7.1	Small	(0, 0, 25, 50)
	Average-large	(15.07, 30, 300, 300)
Input 7.2	Small	(0, 0, 20, 40)
	Average	(0, 20, 140)
	Large	(10, 20, 200, 200)
Input 7.3	Small-average	(0, 0, 20, 100)
	Large	(50, 70, 100, 100)
Input 7.4	Small	(0, 0, 20, 40)
	Average	(15.07, 30, 90)
	Large	(20, 40, 100, 100)
Input 7.5	Small-average	(0, 0, 30, 60)
	Large	(30, 60, 100, 100)
Input 8	Low	(0, 0, 3, 5)
	Average-high	(4.5, 7, 10, 10)
Input 8.1	Small-average	(0, 0, 1, 3.5)
	Large	(0.5, 4, 50, 50)
Input 8.2	Small-average	(0, 0, 1, 2.67)
	Large	(2, 4, 40, 40)
Input 8.3	Small-average	(0, 0, 1, 5.8)
	Large	(1.5, 3, 300, 300)
Input 8.4	Small	(0, 0, 1, 3)
	Average	(-1, 3, 7)
	Large	(3, 5, 50, 50)
Input 8.5	Small	(0, 0, 2, 4)

	Average	(-10, 14, 46)
	Large	(37.5, 75, 750, 750)
Input 8.6	Small- average	(0, 0, 1, 3.5)
	Large	(0.5, 1, 100, 100)
Input 10	Poor	(0, 0, 4, 8)
	Average	(1.03, 7, 12.97)
	Good	(5.33, 8, 10, 10)
Input 10.1	Short	(0, 0, 1, 6)
	Average-large	(2.5, 5, 50, 50)
Input 10.2	Small	(0, 0, 2, 8)
	Average	(4, 6, 8)
	Large	(10, 20, 200, 200)
Input 10.3	Small	(0, 0, 3)
	Average	(-2, 2, 6)
	Large	(1.5, 3, 10, 10)
Input 10.4	Small	(0, 0, 6)
	Average	(7.54, 15, 37.38)
	Large	(10, 30, 300, 300)
Input 10.5	Risk ave	(0, 0, 5.33)
	Average	(0, 5, 9)
	Risk pro	(5, 9, 10, 10)
Input 11	Poor	(0, 0, 5, 10)
	Average	(3.5, 5, 9.5)
	Good	(6.46, 8, 10, 10)
Input 11.1	Short	(0, 0, 3.33, 6.67)
	Average	(-10, 20, 60)
	Long	(-10, 90, 900, 900)
Input 11.2	Small	(0, 0, 50, 70)
	Average-large	(40, 80, 100, 100)
Input 11.3	Small-average	(0, 0, 20, 40)
	Large	(20, 40, 400, 400)
Input 12	Low	(0, 0, 2, 4)
	Average	(0, 5, 9)
	High	(3.5, 7, 10,10)
Input 12.1	Small-average	(0, 0, 10, 30)
	Large	(4, 8, 80, 80)
Input 12.2	Small-average	(0, 0, 24)
	Large	(30, 60, 100, 100)
Input 13	Low	(0, 0, 2, 4)
	Average	(1, 5, 6, 8.67)
	High	(6.66, 8, 10, 10)
Input 13.1	Small	(0, 0, 80, 120)
	Average-large	(0, 50, 100)
Input 13.2	Small-average	(0, 0, 2, 7)
	Large	(3, 6, 10, 10)
Input 3.3	Small-average	(0, 0, 1, 6.25)
	Large	(1.5, 3, 10, 10)

Input 13.4	Small	(0, 0, 3, 6)
	Average	(2, 4, 6, 34)
	Large	(15, 30, 100, 100)
Input 13.5	Small	(0, 0, 1, 2)
	Average-large	(-11.8, 30, 45)
Input 13.6	Small	(0, 0, 20)
	Average	(-25, 5, 35)
	Large	(15, 30, 100, 100)
Input 13.7	Small	(0, 0, 45)
	Average	(0, 45, 90)
	Large	(45, 90, 900, 900)
Input 14	Unfavor	(0, 0, 3, 6)
	Average	(1, 5, 9)
	Favor	(4.4, 8, 10, 10)
Output 1	Poor	(0, 0, 3, 6)
	Average	(3, 6, 8, 11)
	Good	(5, 9, 10, 10)
Output 1.1	Small	(0, 0, 25, 50)
	Average	(2.5, 5, 10, 40)
	Large	(5, 10, 100, 100)
Output 1.2	Small-average	(0, 0, 10, 30)
	Large	(12.5, 25, 100, 100)
Output 1.3	Small-average	(0, 0, 5, 57.5)
	Large	(20.83, 41.67, 100, 100)
Output 1.4	Small	(0, 0, 5, 10)
	Average	(2.54, 10, 32.4)
	Large	(10, 20, 100, 100)
Output 1.5	Small-average	(0, 0, 5, 34.85)
	Large	(12.5, 25, 100, 100)
Output 1.6	Small	(0, 0, 5, 7)
	Average	(-1, 5, 10, 40)
	Large	(12.5, 25, 100, 100)
Output 2	Poor	
	Average	(4, 7, 8.5)
	good	(4.52, 9, 10, 10)
Output 2.1 (+)	Small-average	(0, 0, 10, 70)
	Large	(26.67, 54, 100, 100)
Output 2.1 (-)		
Output 2.2	Small	(0, 0, 12.5)
	Average	(4.17, 20, 35.83)
	Large	(26, 30, 100, 100)
Output 3	Low	(0, 0, 7, 14)
	Average-high	(4.57, 8, 10, 10)
Output 3.1	Small	(0, 0, 10, 20)
	Average	(-16, 20, 30)
	Large	(10, 40, 100, 100)
Output 3.2	Small	(0, 0, 2, 18)

	Average-large	(-20, 40, 400, 400)
Output 3.3	Small	(0, 0, 1, 9)
	Average	(-24.85, 5, 34.85)
	Large	(20, 40, 100, 100)
Output 3.4	Small	(0, 0, 200, 350)
	Average-large	(40, 80, 2000, 2000)
Output 3.5	Small-average	(0, 0, 5, 15)
	Large	(0, 20, 100, 100)
Input 9	Low	(0, 0, 3, 4.45)
	Average	(1.03, 7, 10)
	High	(6, 9, 10, 10)
Input 9.1	Small	(0, 0, 0.3, 0.6)
	Average-large	(0.3, 0.6, 100, 100)
Input 9.2	Small	(0, 0, 6)
	Average	(1.25, 2.5, 5, 6.25)
	Large	(4, 8, 100, 100)
Input 9.3	Small	(0, 0, 0,3, 0.6)
	Average	(1.2, 2.4, 3.6)
	Large	(5, 10, 100, 100)
Input 9.4	Small-average	(0, 0, 6)
	Large	(3, 6, 60, 60)
Input 9.5	Small-average	(0, 0, 0.53)
	Large	(5, 10, 100, 100)
Input 9.6	Small	(0, 0, 0.4, 0.8)
	Average-large	(4, 8, 80, 80)
Input 9.7	Small	(0, 0, 0.5, 1)
	Average-large	(3.6, 7.2, 72, 72)
Input 9.8	Small-average	(0, 0, 0.457)
	Large	(0.2285, 0.457, 50, 50)

Appendix 4: Parameters for the Derived Membership Functions (Trial 2)

No.	Name	Parameter a	Parameter b	Parameter c	Parameter d
Input 1	small	0	0	3	6
	average	4	5	5	7
	large	4	8	10	10
Input 1.1	small	0	0	100	200
	av-large	150	650	6500	6500
Input 1.2	small	0	0	3	17
	av-large	-96	100	1000	1000
Input 1.3	small	0	0	4	16
	average	-88	100	100	288
	large	100	288	2880	2880
Input 2	low	0	0	4	8
	average	-2.5	5	5	12.5
	high	4	8	10	10
Input 2.1	small	0	0	2	4
	av-large	1.05	10	100	100
Input 2.2	small	0	0	10	90
	average	10	20	200	210
	large	200	210	2100	2100
Input 3	poor	0	0	3	6
	average	4	5	5	7.67
	good	4.67	8	10	10
Input 3.1	low	0	0	2	4
	average	1	5	6	10
	high	4	18	10	10
Input 3.2	small	0	0	10	20
	av-large	-5	15	150	150
Input 3.3	small-av	0	0	50	95
	large	20	40	100	100
Input 3.4	poor	0	0	2.33	4.67
	average	-1	8	8	17
	good	7	9	10	10
Input 3.5	small	0	0	5	10
	av-large	5	10	100	100
Input 3.6	poor	0	1	1	3
	average	-1	5	5	11
	good	6	9	10	10
Input 3.7	small	0	0	14.33	25

	av-large	14.33	25	100	100
Input 3.8	small-av	0	0	0	7.5
	large	1	2	100	100
Input 3.9	small	0	0	2	7.5
	av-large	5	10	10	10
Input 4	small-av	0	0	4	13
	large	4	8	10	10
Input 4.1	small	0	0	2	4
	av-large	-2.5	7.5	100	100
Input 4.2	short-av	0	0	12	15
	long	4.5	9	100	100
Input 4.3	small	0	0	3.6	7.2
	av-large	-4	10	2000	2000
Input 5	small	0	0	1.33	2.67
	average	-1	5	5	11
	large	4	8	10	10
Input 5.1	small	0	0	10	20
	average	-20.6	40	40	70.3
	large	15	30	300	300
Input 5.2	small-av	0	0	400	800
	large	10	20	2000	2000
Input 6	low	0	0	2	4
	average	-1	3	3	7
	high	2	8	10	10
Input 6.1	low	0	0	2	6
	av-high	0.54	8	10	10
Input 6.2	low	0	0	2	4
	av-high	2	8	10	10
Input 6.3	small-av	0	0	20	120
	large	65	75	100	100
Input 6.4	small	0	0	3	7
	av-large	2	8	10	10
Input 7	low	0	0	2	8
	average	4	6	6	8.67
	high	5.67	9	10	10
Input 7.1	small-av	0	0	10	63.33
	large	13.33	60	100	100
Input 7.2	small	0	0	10	20
	av-large	5	10	100	100
Input 7.3	small-av	0	0	0	120
	large	0	100	100	100
Input 7.4	small	0	0	10	25

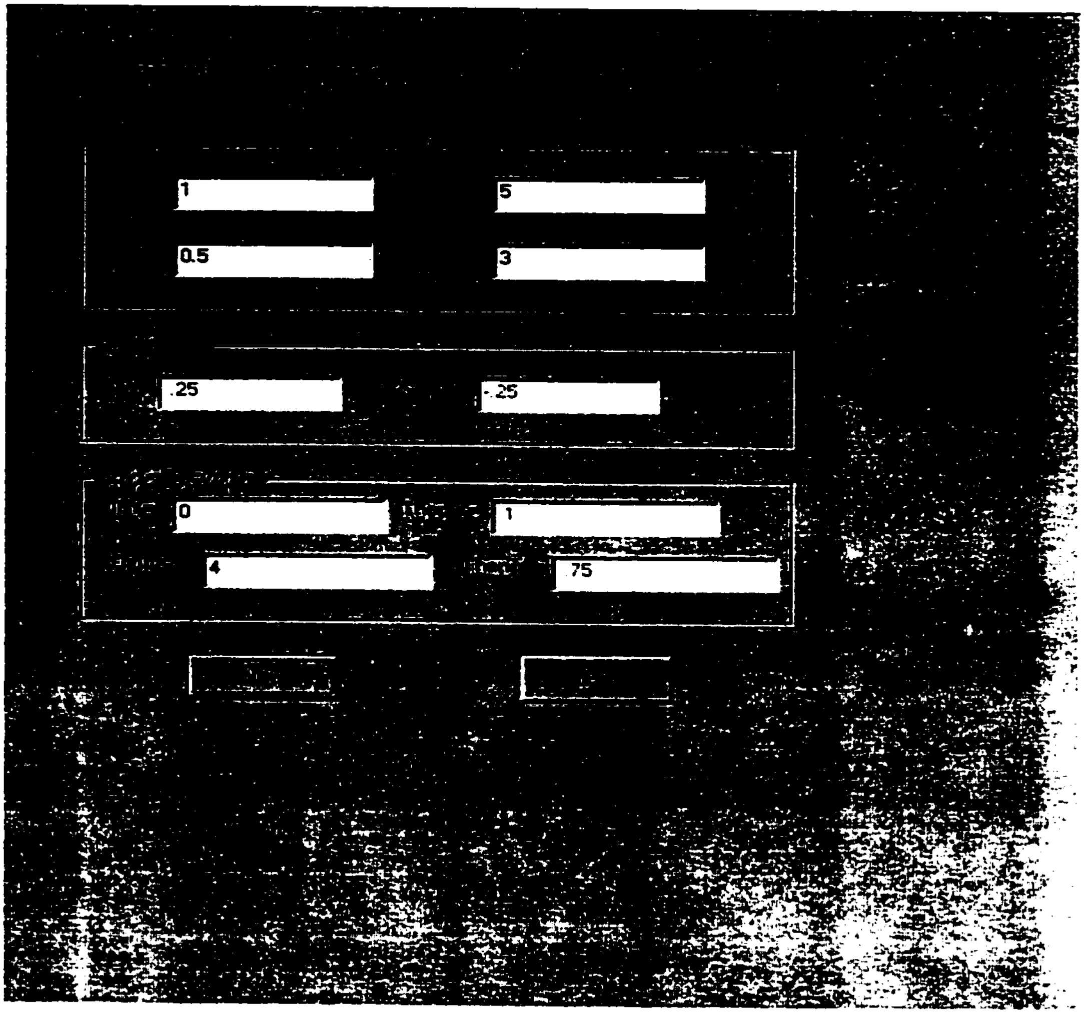
	average	10	20	30	40
	large	20	40	100	100
Input 7.5	small	0	0	10	40
	average	10	40	40	70
	large	25	50	100	100
Input 8	low	0	0	3	6
	average	4	5	5	14
	high	4	8	10	10
Input 8.1	small-av	0	0	1	4
	large	0	5	100	100
Input 8.2	small-av	0	0	1	3
	large	-4	8	100	100
Input 8.3	small-av	0	0	1	37.25
	large	3	6	100	100
Input 8.4	small-av	0	0	2	5.75
	large	2.75	5	50	50
Input 8.5	short-av	0	0	30	60
	long	7.5	15	1000	1000
Input 8.6	small-av	0	0	0	7.5
	large	0.5	1	20	20
Input 10	poor	0	0	1	9
	average	3	7	7	9
	good	7	9	10	10
Input 10.1	short	0	0	1	3.7
	average	-0.3	5	5	23
	long	5	10	1000	1000
Input 10.2	small	0	0	0	9
	average	3	6	10	13
	large	-77	100	1000	1000
Input 10.3	small	0	0	0	2.4
	average	0	2.4	2.4	4.8
	large	1.5	3	10	10
Input 10.4	small	0	0	3	20
	average	7.5	15	20	35
	large	25	30	100	100
Input 10.5	risk-ave	0	0	3	6
	average	4	5	5	6
	risk-pro	3	8	10	10
Input 11	poor-av	0	0	5	9
	good	4	10	10	10
Input 11.1	short	0	0	3.33	6.67
	average	-30	40	40	50

	long	10	20	2000	2000
Input 11.2	small	0	0	50	100
	av-large	40	70	100	100
Input 11.3	small	0	0	0	7.5
	average	2.5	5	15	17.5
	large	10.83	21.67	100	100
Input 12	low	0	0	2	6
	average	3	5	5	7
	high	3	9	10	10
Input 12.1	small-av	0	0	3	28
	large	15	16	100	100
Input 12.2	small-av	0	0	0	30
	large	-70	100	100	100
Input 13	low	0	0	0.5	5
	average	0.52	5	5	8
	high	5	8	10	10
Input 13.1	small	0	0	100	100
	av-large	28	95	100	100
Input 13.2	small-av	0	0	2	8
	large	2	8	10	10
Input 13.3	small	0	0	1	2.3
	av-large	-5.1	4	50	50
Input 13.4	small	0	0	2	4
	average	2	4	20	22
	large	12	30	100	100
Input 13.5	small	0	0	2	4
	average	2	4	4	56
	large	4	8	100	100
Input 13.6	small-av	0	0	0	15
	large	-35	50	100	100
Input 13.7	small	0	0	0	45
	average	22.5	45	45	67.5
	large	45	90	100	100
Input 14	unfavorable	0	0	3	5
	average	3	5	7	9
	favorable	4	8	10	10
Output 1	poor	0	1	1	7
	average	3	5	6	8
	good	6	8	10	10
Output 1.1	small	0	0	25	50
	av-large	5	10	100	100
Output 1.2	small	0	0	10	20

	av-large	0	10	100	100
Output 1.3	small	0	0	10	20
	av-large	2.5	5	100	100
Output 1.4	small-av	0	0	10	35
	large	10	20	100	100
Output 1.5	small-av	0	0	5	118
	large	23	100	100	100
Output 1.6	small-av	0	0	16	32
	large	23	25	100	100
Output 2	poor-av	0	0	4	10
	good	6	8	10	10
Output 2.1 (+)	small-av	0	0	0	75
	large	0	75	100	100
Output 2.1 (-)	small	-5	0	0	0
	average	-10	-5	-5	0
	large	-100	-100	-10	-5
Output 2.2	small	0	0	0	20
	average	0	20	20	40
	large	20	40	100	100
Output 3	low	0	0	2	4
	average	1	5	6	10
	high	4	8	10	10
Output 3.1	small	0	0	10	20
	av-large	10	20	100	100
Output 3.2	small	0	0	10	55
	large-av	25	50	500	500
Output 3.3	small-av	0	0	5	15
	large	-50	70	100	100
Output 3.4	small	0	0	200	400
	average	-400	1000	1000	2400
	large	1000	2400	24000	24000
Output 3.5	small-av	0	0	5	15
	large	5	15	100	100
Input 9	low	0	0	2	5
	average	2	5	5	11
	high	4	8	10	10
Input 9.1	small	0	0	0	0.9
	average	0	45	45	90
	large	45	90	900	900
Input 9.2	small-av	0	0	0	1.33
	large	0	40	400	400
Input 9.3	small-av	0	0	0	4.8

	large	0.5	1	100	100
Input 9.4	small-av	0	0	0	6.67
	large	0	6.67	667	667
Input 9.5	small-av	0	0	0	8
	large	0	8	800	800
Input 9.6	small	0	0	0	0.8
	average	0	0.8	0.8	1.6
	large	0.8	1.6	160	160
Input 9.7	small	0	0	0	0.15
	average	0	7.2	8.4	15.6
	large	8.4	15.6	156	156
Input 9.8	small-av	0	0	0	0.8
	large	0.2	0.6	600	600

Appendix 5: Visual Basic Program for Parameter Calculation



Appendix 6: Membership Functions Testing Results (Trial 1)

Input 1								
			Testing value □(x)					
	Real number	Linguistic term	small	average	high	Highest	Match Y/N	Total accuracy (%)
Case 1	3	Small	1		0	1	Y	100
Case 2	5	Average	1		0	1	Y	
Case 3	8	Large	0.4		1	1	Y	
Case 4	3	Small	1		0	1	Y	
Case 5	9	Large	0.2		1	1	Y	

Input 1.1								
			Testing value □(x)					
	Real number	Linguistic term	small	average	high	Highest	Match Y/N	Total accuracy (%)
Case 1	2	Small	1		0	1	Y	100
Case 2	900	Average	0		1	1	Y	
Case 3	60	Large	0.5		0.5	0.5	Y	
Case 4	12	Small	1		0	1	Y	
Case 5	1000	Large	0		1	1	Y	

Input 1.2								
			Testing value □(x)					
	Real number	Linguistic term	small	average	high	Highest	Match Y/N	Total accuracy (%)
Case 1	0.3	Small	1		0	1	Y	75
Case 2	100	Average	0		1	1	N	
Case 3	100	Large	0		1	1	Y	
Case 4	1.1	Small	1		0	1	Y	
Case 5	N/A	N/A	N/A		N/A	0	N/A	

Input 1.3								
			Testing value □(x)					
	Real number	Linguistic term	small	average	high	Highest	Match Y/N	Total accuracy (%)
Case 1	4	Small	1		0	1	Y	80
Case 2	10	Small	0.96		0.039	0.961	Y	
Case 3	5	Small	1		0	1	Y	
Case 4	4	Small	1		0	1	Y	
Case 5	100	Average	0.25		0.75	0.75	N	

Input 2								
			Testing value □(x)					
	Real number	Linguistic term	low	average	average-large	Highest	Match Y/N	Total accuracy (%)
Case 1	8	high	0		1	1	Y	80
Case 2	8	high	0		1	1	Y	
Case 3	10	Average	0		1	1	Y	
Case 4	4	low	0.67		0.05	0.67	Y	
Case 5	5	Average	0.33		0.29	0.33	N	

Input 2.1									
			Testing value □(x)						
	Real number	Linguistic term	small	average	average-large	Highest	Match Y/N	Total accuracy (%)	
Case 1	20	Average	0		1	1	Y	100	
Case 2	5	Average	0		0.388	0.388	Y		
Case 3	4	large	0		0.27	0.27	Y		
Case 4	10	Average	0		1	1	Y		
Case 5	10	Average	0		1	1	Y		

Input 2.2									
			Testing value □(x)						
	Real number	Linguistic term	small	average	large	Highest	Match Y/N	Total accuracy (%)	
Case 1	50	small	0.7		0.2	0.7	Y	80	
Case 2	20	Average	0.92		0	0.917	Y		
Case 3	20	small	0.92		0	0.92	Y		
Case 4	50	Average	0.67		0.21	0.67	Y		
Case 5	200	Average	0		1	1	N		

Input 3									
			Testing value □(x)						
	Real number	Linguistic term	Poor	Average	Good	Highest	Match Y/N	Total accuracy (%)	
Case 1	5	Average	0.8	1	0	1	Y	100	
Case 2	7	Average	0.25	0.667	0.333	0.667	Y		
Case 3	5	Average	0.75	1	0	1	Y		
Case 4	9	good	0	0.33	1	1	Y		
Case 5	3	poor	1	0.5	0	1	Y		

Input 3.1									
			Testing value □(x)						
	Real number	Linguistic term	small	average	average-large	Highest	Match Y/N	Total accuracy (%)	
Case 1	5	Average	1		0	1	N	40	
Case 2	8	Average	0.5		0.5	0.5	Y		
Case 3	8	large	0.5		0.5	0.5	Y		
Case 4	5	Average	1		0	1	N		
Case 5	6	Average	1		0	1	N		

Input3.2									
			Testing value □(x)						
	Real number	Linguistic term	small	average	average-large	Highest	Match Y/N	Total accuracy (%)	
Case 1	2	Small	1		0.2	1	Y	75	
Case 2	10	Small	0		1	1	N		
Case 3	15	Average	0		1	1	Y		
Case 4	5	Average	0		0.79	0.79	Y		
Case 5	N/A	N/A	N/A				N/A		

Input 3.3								
			Testing value □(x)					
	Real number	Linguistic term	small-	average	Large	Highest	Match Y/N	Total accuracy (%)
Case 1	50	Small	1		1	1	Y	100
Case 2	16.7	Average	1		0.127	1	Y	note: cannot
Case 3	50	Average	1		1	1	Y	distinguish these two
Case 4	80	Average	0.39		1	1	n	
Case 5	50	Large	1		1	1	Y	

Input 3.4								
			Testing value □(x)					
	Real number	Linguistic term	Poor	average	average-good	Highest	Match Y/N	Total accuracy (%)
Case 1	5	poor	0.8		0.1	0.8	Y	100
Case 2	8	Average	0.08		1	1	Y	
Case 3	9	good	0		1	1	Y	
Case 4	9	good	0		1	1	Y	
Case 5	8	Average	0.08		1	1	Y	

Input 3.5								
			Testing value □(x)					
	Real number	Linguistic term	small-	average	large	Highest	Match Y/N	Total accuracy (%)
Case 1	20	Average	0.7		1	1	N	60
Case 2	20	Large	0.7		1	1	Y	
Case 3	10	Average	1		0.5	1	Y	
Case 4	24	Large	0.53		1	1	Y	
Case 5	20	Average	0.67		1	1	N	

Input 3.6								
			Testing value □(x)					
	Real number	Linguistic term	poor	average	average-good	Highest	Match Y/N	Total accuracy (%)
Case 1	5	average	0.9		0.1	0.9	n	60
Case 2	5	average	0.9		0.1	0.9	n	
Case 3	9	good	0		1	1	y	
Case 4	9	good	0		1	1	y	
Case 5	1	poor	1		0	1	y	

Input 3.7								
			Testing value □(x)					
	Real number	Linguistic term	small	average	average-large	Highest	Match Y/N	Total accuracy (%)
Case 1	25	large	0		1	1	Y	100
Case 2	28	large	0		1	1	Y	
Case 3	30	large	0		1	1	Y	
Case 4	25	large	0		1	1	Y	
Case 5	30	Average	0		1	1	Y	

Input 3.8									
			Testing value □(x)						
	Real number	Linguistic term	small	average	large	Highest	Match Y/N	Total accuracy (%)	
Case 1	0	small	1		0	1	Y	80	
Case 2	1	small	1		0	1	Y		
Case 3	2	large	1		0.33	1	n		
Case 4	0	small	1		0	1	Y		
Case 5	100	large	0		1	1	Y		

Input 3.9									
			Testing value □(x)						
	Real number	Linguistic term	low	average	average-high	Highest	Match Y/N	Total accuracy (%)	
Case 1	8	Average	0		0.7	0.7	Y	100	
Case 2	10	high	0		1	1	Y		
Case 3	10	Average	0		1	1	Y		
Case 4	9	Average	0		0.83	0.83	Y		
Case 5	2	low	1		0	1	y		

Input 4									
			Testing value □(x)						
	Real number	Linguistic term	small	Average	large	Highest	Match Y/N	Total accuracy (%)	
Case 1	4	small	0.8	0.3	0	0.8	Y	80	
Case 2	2	small	1	0	0	1	Y		
Case 3	8	large	0	0.25	1	1	Y		
Case 4	4	Average	0.75	0.25	0	0.75	n		
Case 5	8	large	0	0.25	1	1	Y		

Input 4.1									
			Testing value □(x)						
	Real number	Linguistic term	small	average	average-large	Highest	Match Y/N	Total accuracy (%)	
Case 1	0.6	Average	1		0.5	1	N	80	
Case 2	2	small	1		0.522	1	Y		
Case 3	14	large	0		0.71	0.71	Y		
Case 4	75	Average	0		1	1	y		
Case 5	80	Average	0		1	1	y		

Input4.2									
			Testing value □(x)						
	Real number	Linguistic term	short	average	average-long	Highest	Match Y/N	Total accuracy (%)	
Case 1	6	Average	1		1	1	Y	100	
Case 2	3	short	1		0	1	y		
Case 3	12	long	1		1	1	Y		
Case 4	12	Average	1		1	1	Y		
Case 5	24	long	0		1	1	y		

Input 4.3								
			Testing value □(x)					
	Real number	Linguistic term	small	average	large	Highest	Match Y/N	Total accuracy (%)
Case 1	0.9	Small	1		0	1	Y	60
Case 2	3.6	Small	1		0.2	1	Y	
Case 3	25	Average	0.81		1	1	N	
Case 4	10	Average	0.94		1	1	N	
Case 5	100	Large	0.15		1	1	Y	

Input 5								
			Testing value □(x)					
	Real number	Linguistic term	small	average	average-large	Highest	Match Y/N	Total accuracy (%)
Case 1	7	Average	0.3		0.7	0.7	Y	40
Case 2	5	Average	0.8		0.2	0.8	n	
Case 3	8	Average	0		1	1	Y	
Case 4	5	Average	0.8		0.2	0.8	n	
Case 5	6	Average	0.53		0.47	0.53	n	

Input 5.1								
			Testing value □(x)					
	Real number	Linguistic term	small	average	average-large	Highest	Match Y/N	Total accuracy (%)
Case 1	30	Average	0.3		0.7	0.7	Y	80
Case 2	40	Average	0.01		1	1	Y	
Case 3	45	Average	0		1	1	Y	
Case 4	20	Average	0.67		0.33	0.67	N	
Case 5	45	Average	0		1	1	Y	

Input 5.2								
			Testing value □(x)					
	Real number	Linguistic term	small	average	average-large	Highest	Match Y/N	Total accuracy (%)
Case 1	35	Average	0		0.9	0.9	Y	75
Case 2	300	Small	0		1	1	N	
Case 3	N/A	N/A	0		0	0	N/A	
Case 4	190	Average	0		1	1	Y	
Case 5	1750	Average	0		1	1	Y	

Input 6								
			Testing value □(x)					
	Real number	Linguistic term	low	average	high	Highest	Match Y/N	Total accuracy (%)
Case 1	5	Average	0	0.8	0	0.8	Y	40
Case 2	3	Average	0.67	0.5	0	0.667	n	
Case 3	3	low	0.67	0.5	0	0.67	Y	
Case 4	8	Average	0	0.88	1	1	n	
Case 5	3	Average	0.67	0.5	0	0.67	N	

Input 6.1								
			Testing value □(x)					
	Real number	Linguistic term	low	average	average-high	Highest	Match Y/N	Total accuracy (%)
Case 1	5	Average	0.5		0.3	0.5	n	60
Case 2	7	Average	0.17		0.775	0.775	Y	
Case 3	4	low	0.67		0.1	0.67	Y	
Case 4	9	high	0		1	1	Y	
Case 5	3	Average	0.83		0	0.83	n	

Input 6.2								
			Testing value □(x)					
	Real number	Linguistic term	low	average	average-high	Highest	Match Y/N	Total accuracy (%)
Case 1	5	Average	0.3		0.4	0.4	Y	100
Case 2	8	Average	0		1	1	Y	
Case 3	8	Average	0		1	1	y	
Case 4	7	Average	0		0.79	0.79	y	
Case 5	5	Average	0.25		0.38	0.38	y	

Input 6.3								
			Testing value □(x)					
	Real number	Linguistic term	small	Average	large	Highest	Match Y/N	Total accuracy (%)
Case 1	70	Average	0	0.6	1	1	N	80
Case 2	N/A	N/A				0	N/A	
Case 3	20	Small	0.57	0.47	0	0.57	Y	
Case 4	50	Average	0	1	0.43	1	Y	
Case 5	5	Small	1	0.21	0	1	Y	

Input 6.4								
			Testing value □(x)					
	Real number	Linguistic term	small	average	average-large	Highest	Match Y/N	Total accuracy (%)
Case 1	6	Average	0.5		0.9	0.9	y	100
Case 2	5	Average	0.75		0.8	0.8	Y	
Case 3	3	small	1		0.67	1	Y	
Case 4	5	Average	0.75		0.8	0.8	y	
Case 5	8	Average	0		1	1	y	

Input 7								
			Testing value □(x)					
	Real number	Linguistic term	low	average	high	Highest	Match Y/N	Total accuracy (%)
Case 1	6	Average	0.3	1	0.5	1	Y	60
Case 2	5	Average	0.64	0.167	0.25	0.643	N	
Case 3	9	high	0	0.25	1	1	Y	
Case 4	5	Average	0.64	0.17	0.25	0.64	n	
Case 5	2	low	1	0	0	1	y	

Input 7.1								
			Testing value □(x)					
	Real number	Linguistic term	small	average	average-large	Highest	Match Y/N	Total accuracy (%)
Case 1	50	Average	0		1	1	Y	50
Case 2	10	Average	1		0	1	n	
Case 3	10	Average	1		0	1	n	
Case 4	20	Small	1		0.33	1	Y	
Case 5	N/A							

Input 7.2								
			Testing value □(x)					
	Real number	Linguistic term	small	average	large	Highest	Match Y/N	Total accuracy (%)
Case 1	10	Small	1	0.5	0	1	Y	80
Case 2	10	Average	1	0.5	0	1	n	
Case 3	60	Large	0	0.67	1	1	Y	
Case 4	70	Large	0	0.58	1	1	Y	
Case 5	0	Small	1	0	0	1	Y	

Input 7.3								
			Testing value □(x)					
	Real number	Linguistic term	small	average	large	Highest	Match Y/N	Total accuracy (%)
Case 1	10	Small	1		0	1	Y	80
Case 2	100	Large	0		1	1	y	
Case 3	0	Small	1		0	1	Y	
Case 4	50	Large	0.63		0	0.63	n	
Case 5	0	Average	1		0	1	y	

Input 7.4								
			Testing value □(x)					
	Real number	Linguistic term	small	average	large	Highest	Match Y/N	Total accuracy (%)
Case 1	20	Small	1	0.3	0	1	Y	80
Case 2	10	Small	1	0	0	1	Y	
Case 3	10	Small	1	0	0	1	Y	
Case 4	25	average	0.75	0.67	0.25	0.75	n	
Case 5	0	Small	1	0	0	1	y	

Input 7.5								
			Testing value □(x)					
	Real number	Linguistic term	small	average	large	Highest	Match Y/N	Total accuracy (%)
Case 1	10	small	1		0	1	Y	100
Case 2	10	small	1		0	1	Y	
Case 3	10	small	1		0	1	Y	
Case 4	50	large	0.33		0.67	0.67	Y	
Case 5	N/A							

Input 8									
			Testing value □(x)						
	Real number	Linguistic term	low	average	average-high	Highest	Match Y/N	Total accuracy (%)	
Case 1	5	Average	0		0.2	0.2	Y	100	
Case 2	5	Average	0		0.2	0.2	Y		
Case 3	8	Average	0		1	1	Y		
Case 4	N/A								
Case 5	2	low	1		0	1	Y		

Input 8.1									
			Testing value □(x)						
	Real number	Linguistic term	small	average	large	Highest	Match Y/N	Total accuracy (%)	
Case 1	3	small	0.2		0.7	0.7	n	60	
Case 2	1	small	1		0.143	1	Y		
Case 3	1	Average	1		0.143	1	n		
Case 4	1	small	1		0.143	1	Y		
Case 5	5	large	0		1	1	y		

Input 8.2									
			Testing value □(x)						
	Real number	Linguistic term	small	average	large	Highest	Match Y/N	Total accuracy (%)	
Case 1	2	small	0.4		0	0.4	Y	100	
Case 2	1	Average	1		0	1	Y		
Case 3	1	Average	1		0	1	Y		
Case 4	2	small	0.4		0	0.4	Y		
Case 5	n/a								

Input 8.3									
			Testing value □(x)						
	Real number	Linguistic term	small	average	large	Highest	Match Y/N	Total accuracy (%)	
Case 1	1	Average	1		0	1	y	75	
Case 2	1	small	1		0	1	Y		
Case 3	1	Average	1		0	1	Y		
Case 4	8	Average	0		1	1	N		
Case 5	n/a								

Input 8.4									
			Testing value □(x)						
	Real number	Linguistic term	small	average	large	Highest	Match Y/N	Total accuracy (%)	
Case 1	2	Average	0.5	0.8	0	0.8	Y	80	
Case 2	2	Average	0.5	0.8	0	0.8	y		
Case 3	2	Average	0.5	0.8	0	0.8	Y		
Case 4	2	Average	0.5	0.8	0	0.8	Y		
Case 5	2	Small	0.5	0.8	0	0.8	n		

Input 8.5								
			Testing value □(x)					
	Real number	Linguistic term	small	average	large	Highest	Match Y/N	Total accuracy (%)
Case 1	20	Small	0.6	0.8	0	0.8	n	80
Case 2	14	Average	1	1	0	1	Y	
Case 3	10	Small	1	0.83	0	1	Y	
Case 4	30	Average	0	0.5	0	0.5	Y	
Case 5	30	Average	0	0.5	0	0.5	Y	

Input 8.6								
			Testing value □(x)					
	Real number	Linguistic term	small	average	large	Highest	Match Y/N	Total accuracy (%)
Case 1	3	Small	0.2		1	1	n	50
Case 2	3	Average	0.2		1	1	n	
Case 3	0	Average	1		0	1	Y	
Case 4	1	Small	1		1	1	Y	
Case 5	n/a							

Input 10								
			Testing value □(x)					
	Real number	Linguistic term	poor	average	good	Highest	Match Y/N	Total accuracy (%)
Case 1	5	Average	0.8	0.7	0	0.8	n	60
Case 2	8	Average	0	0.832	1	1	N	
Case 3	9	good	0	0.66	1	1	Y	
Case 4	9	good	0	0.66	1	1	Y	
Case 5	1	poor	1	0	0	1	y	

Input 10.1								
			Testing value □(x)					
	Real number	Linguistic term	short	average	average-large	Highest	Match Y/N	Total accuracy (%)
Case 1	5	average	0.2		1	1	Y	100
Case 2	14	average	0		1	1	Y	
Case 3	1	short	1		0	1	Y	
Case 4	10	average	0		1	1	Y	
Case 5	365	large	0		1	1	y	

Input 10.2								
			Testing value □(x)					
	Real number	Linguistic term	small	average	large	Highest	Match Y/N	Total accuracy (%)
Case 1	10	Average	0	0	0.3	0.3	n	75
Case 2	6	Average	0.33	1	0	0.33	Y	
Case 3	n/a							
Case 4	10	Average	0	0	0.3	0.3	y	
Case 5	100	large	0	0	1	1	y	

Input 10.3									
			Testing value □(x)						
	Real number	Linguistic term	small	average	large	Highest	Match Y/N	Total accuracy (%)	
Case 1	0	small	1	0.5	0	1	Y	80	
Case 2	0	small	1	0.5	0	1	Y		
Case 3	0	small	1	0.5	0	1	Y		
Case 4	1	small	0.67	0.75	0	0.75	n		
Case 5	0	small	1	0.5	0	1	Y		

Input 10.4									
			Testing value □(x)						
	Real number	Linguistic term	small	average	large	Highest	Match Y/N	Total accuracy (%)	
Case 1	20	Average	0	0.8	0.5	0.8	Y	100	
Case 2	20	Average	0	0.8	0.5	0.8	Y		
Case 3	30	large	0	0.33	1	1	Y		
Case 4	15	Average	0	1	0.25	1	Y		
Case 5	20	Average	0	0.78	0.5	0.78	y		

Input 10.5									
			Testing value □(x)						
	Real number	Linguistic term	risk a	Average	risk pro	Highest	Match Y/N	Total accuracy (%)	
Case 1	5	Average	0.1	1	0	1	Y	100	
Case 2	5	Average	0.1	1	0	1	Y		
Case 3	8	R/P	0	0.25	0.75	0.75	Y		
Case 4	2	R/A	0.62	0.4	0	0.62	Y		
Case 5	5	Average	0.1	1	0	1	Y		

Input 11									
			Testing value □(x)						
	Real number	Linguistic term	poor-	average	good	Highest	Match Y/N	Total accuracy (%)	
Case 1	3	poor	1		0	1	y	100	
Case 2	5	Average	1		0	1	Y		
Case 3	6	poor	0.75		0	0.75	Y		
Case 4	5	Average	1		0	1	y		
Case 5	6	Average	0.75		0	0.75	y		

Input 11.1									
			Testing value □(x)						
	Real number	Linguistic term	short	average	long	Highest	Match Y/N	Total accuracy (%)	
Case 1	40	Average	0	0.5	0.5	0.5	Y	60	
Case 2	21	Average	0	0.975	0.31	0.975	y		
Case 3	20	long	0	1	0.3	1	n		
Case 4	45	Average	0	0.38	0.55	0.55	n		
Case 5	38	long	0	0	1	1	y		

Input 11.2								
			Testing value □(x)					
	Real number	Linguistic term	small	average	average-large	Highest	Match Y/N	Total accuracy (%)
Case 1	70	Average	0		0.8	0.8	Y	100
Case 2	70	Average	0		0.75	0.75	Y	
Case 3	20	Small	1		0	1	Y	
Case 4	85	Average	0		1	1	Y	
Case 5	80	Average	0		1	1	Y	

Input 11.3								
			Testing value □(x)					
	Real number	Linguistic term	small	average	large	Highest	Match Y/N	Total accuracy (%)
Case 1	15	Average	1		0	1	Y	100
Case 2	0	Small	1		0	1	Y	
Case 3	n/a							
Case 4	5	Average	1		0	1	Y	
Case 5	n/a							

Input 12								
			Testing value □(x)					
	Real number	Linguistic term	low	average	high	Highest	Match Y/N	Total accuracy (%)
Case 1	5	average	0	1	0.4	1	Y	100
Case 2	n/a					0		
Case 3	5	average	0	1	0.4	1	Y	
Case 4	n/a					0		
Case 5	9	high	0	0	1	1	y	

Input 12.1								
			Testing value □(x)					
	Real number	Linguistic term	small	average	large	Highest	Match Y/N	Total accuracy (%)
Case 1	3	Small	1		0	1	Y	100
Case 2	n/a							
Case 3	3	average	1		0	1	Y	
Case 4	1	Small	1		0	1	Y	
Case 5	16	large	0.7		1	1	y	

Input 12.2								
			Testing value □(x)					
	Real number	Linguistic term	small	average	large	Highest	Match Y/N	Total accuracy (%)
Case 1	10	small	0.6		0	0.6	Y	100
Case 2	n/a							
Case 3	20	Average	0.17		0	0.17	Y	
Case 4	10	small	0.58		0	0.58	Y	
Case 5	n/a							

Input 13								
			Testing value □(x)					
	Real number	Linguistic term	low	average	high	Highest	Match Y/N	Total accuracy (%)
Case 1	5	Average	0	1	0	1	Y	33
Case 2	n/a					0		
Case 3	8	Average	0	0.25	1	1	n	
Case 4	n/a					0		
Case 5	2	Average	1	0.25	0	1	n	

Input 13.1								
			Testing value □(x)					
	Real number	Linguistic term	small	average	average-large	Highest	Match Y/N	Total accuracy (%)
Case 1	90	Average	0.8		1	1	Y	100
Case 2	n/a							
Case 3	95	large	0.63		1	1	Y	
Case 4	100	large	0.5		1	1	Y	
Case 5	95	Average	0.63		1	1	y	

Input 13.2								
			Testing value □(x)					
	Real number	Linguistic term	small	Average	large	Highest	Match Y/N	Total accuracy (%)
Case 1	3	small	0.8		0	0.8	Y	100
Case 2	n/a							
Case 3	3	Average	0.8		0	0.8	Y	
Case 4	1	small	1		0	1	Y	
Case 5	1	Average	1		0	1	Y	

Input 13.3								
			Testing value □(x)					
	Real number	Linguistic term	small	Average	large	Highest	Match Y/N	Total accuracy (%)
Case 1	1	small	1		0	1	y	75
Case 2	n/a							
Case 3	3	Average	0.62		1	1	Y	
Case 4	1	small	1		0	1	y	
Case 5	4	Average	0.43		1	1	N	

Input 13.4								
			Testing value □(x)					
	Real number	Linguistic term	small	average	large	Highest	Match Y/N	Total accuracy (%)
Case 1	4	Average	0.7	1	0	1	Y	100
Case 2	n/a							
Case 3	30	large	0	0.14	1	1	Y	
Case 4	4	Average	0.67	1	0	1	Y	
Case 5	N/A	N/A	N/A				N/A	

Input 13.5								
			Testing value □(x)					
	Real number	Linguistic term	small	average	average-large	Highest	Match Y/N	Total accuracy (%)
Case 1	2	Small	0		0.3	0.3	n	75
Case 2	n/a							
Case 3	8	Large	0		0.47	0.47	Y	
Case 4	16	Average	0		0.67	0.67	Y	
Case 5	4	Average	0		0.38	0.38	Y	

Input 13.6								
			Testing value □(x)					
	Real number	Linguistic term	small	average	large	Highest	Match Y/N	Total accuracy (%)
Case 1	10	Small	0.5	0.8	0	0.8	n	75
Case 2	n/a							
Case 3	0	Small	1	0.83	0	1	Y	
Case 4	50	Large	0	0	1	1	Y	
Case 5	0	Small	1	0.83	0	1	Y	

Input 13.7								
			Testing value □(x)					
	Real number	Linguistic term	small	average	large	Highest	Match Y/N	Total accuracy (%)
Case 1	0	Small	1	0	0	1	Y	100
Case 2	n/a							
Case 3	n/a							
Case 4	0	Small	1	0	0	1	Y	
Case 5	N/A	N/A	N/A		N/A	0	N/A	

Input 14								
			Testing value □(x)					
	Real number	Linguistic term	unfav	average	favor	Highest	Match Y/N	Total accuracy (%)
Case 1	5	Average	0.3	1	0.2	1	Y	60
Case 2	8	Average	0	0.25	1	1	n	
Case 3	8	favor	0	0.25	1	1	Y	
Case 4	7	Average	0	0.5	0.72	0.72	n	
Case 5	6	Average	0	0.75	0.44	0.75	y	

Output 1								
			Testing value □(x)					
	Real number	Linguistic term	poor	average	good	Highest	Match Y/N	Total accuracy (%)
Case 1	5	average	0.3	0.7	0	0.7	Y	80
Case 2	5	average	0.33	0.667	0	0.667	Y	
Case 3	9	good	0	0.67	1	1	Y	
Case 4	8	good	0	1	0.75	1	n	
Case 5	1	poor	1	0	0	1	y	

Output 1.1									
			Testing value □(x)						
	Real number	Linguistic term	small	average	large	Highest	Match Y/N	Total accuracy (%)	
Case 1	15	Average	1		1	1	n	40	
Case 2	30	Average	0.8		1	1	n		
Case 3	20	large	1		1	1	Y		
Case 4	20	large	1		1	1	Y		
Case 5	20	Average	1		1	1	n		

Output 1.2									
			Testing value □(x)						
	Real number	Linguistic term	small	average	large	Highest	Match Y/N	Total accuracy (%)	
Case 1	10	small	1		0.9	1	Y	20	
Case 2	90	Average	0		1	1	n		
Case 3	10	large	1		0	1	n		
Case 4	20	Average	0.5		0.6	0.6	n		
Case 5	70	Average	0		1	1	n		

Output 1.3									
			Testing value □(x)						
	Real number	Linguistic term	small	average	large	Highest	Match Y/N	Total accuracy (%)	
Case 1	10	small	0.9		0	0.9	Y	80	
Case 2	10	small	0.9		0	0.9	Y		
Case 3	10	large	0.9		0	0.9	n		
Case 4	10	small	0.9		0	0.9	Y		
Case 5	20	Average	0.71		0	0.71	Y		

Output 1.4									
			Testing value □(x)						
	Real number	Linguistic term	small	average	large	Highest	Match Y/N	Total accuracy (%)	
Case 1	10	small	0	1	0	1	N	50	
Case 2	30	Average	0	0.107	1	1	n		
Case 3	10	Average	0	1	0	1	Y		
Case 4	20	large	0	0.55	1	1	y		
Case 5	n/a								

Output 1.5									
			Testing value □(x)						
	Real number	Linguistic term	small	average	large	Highest	Match Y/N	Total accuracy (%)	
Case 1	10	Small	0.8		0	0.8	Y	25	
Case 2	90	Average	0		1	1	N		
Case 3	5	Average	1		0	1	Y		
Case 4	20	Average	0.5		0.6	0.6	n		
Case 5	N/A	N/A	N/A				N/A		

Output 1.6								
			Testing value □(x)					
	Real number	Linguistic term	small	average	large	Highest	Match Y/N	Total accuracy (%)
Case 1	15	Average	0	0.8	0.2	0.8	Y	100
Case 2	n/a							
Case 3	16	Average	0	0.8	0.28	0.8	Y	
Case 4	n/a							
Case 5	3.5	Small	1	0.75	0	1	Y	

Output 2								
			Testing value □(x)					
	Real number	Linguistic term	poor	average	good	Highest	Match Y/N	Total accuracy (%)
Case 1	4	poor	0.9	0	0	0.9	Y	80
Case 2	4	Average	0.91	0	0	0.91	n	
Case 3	7	Average	0.09	1	0.55	1	Y	
Case 4	8	good	0	0.33	0.78	0.78	Y	
Case 5	1	poor	1	0	0	1	Y	

Output 2.1								
			Testing value □(x)					
	Real number	Linguistic term	small	average	large	Highest	Match Y/N	Total accuracy (%)
Case 1	20	Average	0.8		0	0.8	Y	100
Case 2	15	Average	0.92		0	0.917	y	
Case 3	0	Average	1		0	1	Y	
Case 4	0	Small	1		0	1	Y	
Case 5	100	Large	0		1	1	y	

Output 2.2								
			Testing value □(x)					
	Real number	Linguistic term	small	average	large	Highest	Match Y/N	Total accuracy (%)
Case 1	5	Small	0.6	0.1	0	0.6	Y	100
Case 2	20	average	0	1	0	1	Y	
Case 3	0	Small	1	0	0	1	Y	
Case 4	0	Small	1	0	0	1	Y	
Case 5	100	large	0	0	1	1	y	

Output 3								
			Testing value □(x)					
	Real number	Linguistic term	low	average	average-high	Highest	Match Y/N	Total accuracy (%)
Case 1	6	average	1		0.4	1	n	60
Case 2	5	average	1		0.125	1	n	
Case 3	9	high	0.71		1	1	Y	
Case 4	9	high	0.71		1	1	Y	
Case 5	2	low	1		0	1	y	

Output 3.1									
			Testing value □(x)						
	Real number	Linguistic term	small	average	large	Highest	Match Y/N	Total accuracy (%)	
Case 1	6	Average	1	0.6	0	1	n	33	
Case 2	n/a								
Case 3	100	large	0	0	1	1	Y		
Case 4	10	Average	1	0.72	0	1	n		
Case 5	n/a								

Output 3.2									
			Testing value □(x)						
	Real number	Linguistic term	small	average	average-large	Highest	Match Y/N	Total accuracy (%)	
Case 1	5	small	0.8		0.4	0.8	Y	33	
Case 2	n/a								
Case 3	50	small	0		1	1	n		
Case 4	6	Average	0.75		0.43	0.75	n		
Case 5	n/a								

Output 3.3									
			Testing value □(x)						
	Real number	Linguistic term	small	Average	large	Highest	Match Y/N	Total accuracy (%)	
Case 1	5	small	0.5	1	0	1	n	33	
Case 2	n/a								
Case 3	5	small	0.5	1	0	1	n		
Case 4	1	small	1	0.83	0	1	Y		
Case 5	n/a								

Output 3.4									
			Testing value □(x)						
	Real number	Linguistic term	small	average	average-large	Highest	Match Y/N	Total accuracy (%)	
Case 1	40	small	1		0	1	y	100	
Case 2	n/a								
Case 3	n/a								
Case 4	50	small	1		0.25	1	y		
Case 5	n/a								

Output 3.5									
			Testing value □(x)						
	Real number	Linguistic term	small	average	large	Highest	Match Y/N	Total accuracy (%)	
Case 1	4	Small	1		0.2	1	Y	100	
Case 2	n/a								
Case 3	n/a								
Case 4	2	Average	1		0.1	1	Y		
Case 5	N/A	N/A	N/A				N/A		

Input 9								
			Testing value □(x)					
	Real number	Linguistic term	low	average	high	Highest	Match Y/N	Total accuracy (%)
Case 1	5	Average	0	0.7	0	0.7	Y	75
Case 2	5	Average	0	0.7	0	0.7	Y	
Case 3	6	Average	0	0.83	0	0.83	Y	
Case 4	n/a							
Case 5	3	Average	1	0.33	0	1	n	

Input 9.1								
			Testing value □(x)					
	Real number	Linguistic term	small	average	average-large	Highest	Match Y/N	Total accuracy (%)
Case 1	0.3	Small	1		0	1	Y	100
Case 2	0	Small	1		0	1	Y	
Case 3	0	Small	1		0	1	Y	
Case 4	90	Large	0		1	1	Y	
Case 5	0	Small	1		0	1	Y	

Input 9.2								
			Testing value □(x)					
	Real number	Linguistic term	small	average	large	Highest	Match Y/N	Total accuracy (%)
Case 1	0.2	Small	1	0	0	1	Y	40
Case 2	0	Small	1	0	0	1	N	
Case 3	0	average	1	0	0	1	n	
Case 4	40	large	0	0	1	1	Y	
Case 5	1	average	0.83	0	0	0.83	n	

Input 9.3								
			Testing value □(x)					
	Real number	Linguistic term	small	average	large	Highest	Match Y/N	Total accuracy (%)
Case 1	0.6	Small	0.9	0	0	0.9	y	50
Case 2	0	Small	1	0	0	1	Y	
Case 3	1	large	0.67	0	0	0.67	n	
Case 4	0.8	average	0.8	0	0	0.8	n	
Case 5	n/a							

Input 9.4								
			Testing value □(x)					
	Real number	Linguistic term	small-	average	large	Highest	Match Y/N	Total accuracy (%)
Case 1	1.5	Average	0.8		0	0.8	Y	75
Case 2	0	small	1		0	1	Y	
Case 3	0	small	1		0	1	Y	
Case 4	5	Average	0.17		0.67	0.67	n	
Case 5	n/a							

Input 9.5									
			Testing value □(x)						
	Real number	Linguistic term	small	average	large	Highest	Match Y/N	Total accuracy (%)	
Case 1	2.5	average	0.8		0	0.8	y	100	
Case 2	0	small	1		0	1	Y		
Case 3	0	small	1		0	1	Y		
Case 4	4	average	0.6		0	0.6	y		
Case 5	n/a								

Input 9.6									
			Testing value □(x)						
	Real number	Linguistic term	small	average	average-large	Highest	Match Y/N	Total accuracy (%)	
Case 1	0.2	small	1		0	1	Y	100	
Case 2	0	small	1		0	1	Y		
Case 3	0	small	1		0	1	Y		
Case 4	0	small	1		0	1	Y		
Case 5	n/a								

Input 9.7									
			Testing value □(x)						
	Real number	Linguistic term	small	Average	average-large	Highest	Match Y/N	Total accuracy (%)	
Case 1	0.1	small	1		0	1	Y	100	
Case 2	0	small	1		0	1	Y		
Case 3	0	small	1		0	1	Y		
Case 4	1	small	0.9		0	0.9	y		
Case 5	n/a								

Input 9.8									
			Testing value □(x)						
	Real number	Linguistic term	small	average	large	Highest	Match Y/N	Total accuracy (%)	
Case 1	0.1	small	0.8		0	0.8	y	100	
Case 2	0	small	1		0	1	Y		
Case 3	0.6	large	0		1	1	Y		
Case 4	40	large	0		1	1	y		
Case 5	n/a								

Appendix 6-2: Membership Function Testing Results (Trial 2)

No.	Name	test 1	f (test1)	actual term	match (y/n)	test 2	f (test2)	actual term	match (y/n)
Input 1	small					5	0.3333	average	y
	average					5	1		
	large					5	0.25		
Input 1.1	small	12	1	small-av	y	14.5	1	average	n
	av-large	12	0			14.5	0		
Input 1.2	small					17.5	0	average	y
	av-large					17.5	0.5791		
Input 1.3	small	7	0.75	average	n	5	0.9167	average	n
	average	7	0.5053			5	0.4947		
	large	7	0			5	0		
Input 2	low	3	1	low	y	8	0	high	y
	average	3	0.7333			8	0.6		
	high	3	0			8	1		
Input 2.1	small	10	0	average	y	10	0	average	y
	av-large	10	1			10	1		
Input 2.2	small					20	0.875	average	y
	average					20	1		
	large					20	0		
Input 3	poor	8	0	average-goo	y	7.5	0	good	y
	average	8	0			7.5	0.0637		
	good	8	1			7.5	0.8498		
Input 3.1	small					8.5	0	large	n
	average					8.5	0.375		
	large					8.5	0.3214		
Input 3.2	small	6	1	average	n	25	0	average	y
	av-large	6	0.55			25	1		
Input 3.3	small-av	83	0.2667	average	n	67	0.6222	small	n
	large	83	1			67	1		
Input 3.4	poor	8	0	average-good		8.5	0	good	n
	average	8	1		y	8.5	0.9444		
	good	8	0.5			8.5	0.75		
Input 3.5	small	25	0	small-av	y	15	0	large	y
	av-large	25	1			15	1		
Input 3.6	poor	7	0	average		6	0	average	y
	average	7	0.6667		y	6	0.8333		
	good	7	0.3333			6	0		
Input 3.7	small	30	0	average	y	20	0.4686	large	y
	av-large	30	1			20	0.5314		
Input 3.8	small-av	0	1	average	y	10	0	small	n
	large	0	0			10	1		
Input 3.9	low	7	0.0909	average		7.5	0	high	y
	av-high	7	0.4		n	7.5	0.5		
Input 4	small-av	5	0.8889	average	y	6.5	0.7222	average	y
	large	5	0.25			6.5	0.625		
Input 4.1	small					45	0	average	y
	av-large					45	1		

Input 4.2	short-av	6	1	long	n	18	0	average	n
	long	6	0.3333			18	1		
Input 4.3	small	1200	0	average	y	60	0	average	y
	av-large	1200	1			60	1		
Input 5	small	6	0	average		7.5	0	average-lar	y
	average	6	0.8333		y	7.5	0.5833		
	large	6	0.5			7.5	0.875		
Input 5.1	small	20	0	average		40	0	average	y
	average	20	0.67		y	40	1		
	large	20	0.3333			40	1		
Input 5.2	small-av					45	1	average	y
	large					45	1		
Input 6	low	8	0	average		7.5	0	average-hi	y
	average	8	0			7.5	0		
	high	8	1		n	7.5	0.9167		
Input 6.1	low	8	0	average		7.5	0	average-hi	y
	av-high	8	1		y	7.5	0.933		
Input 6.2	low	8	0	average		7.5	0	average-hi	y
	av-high	8	1		y	7.5	0.9167		
Input 6.3	small-av	50	0.7	average	y	80	0.4	average-lar	y
	large	50	0			80	1		
Input 6.4	small	4	0.75	small	y	5.5	0.375	average	y
	av-large	4	0.3333			5.5	0.5833		
Input 7	low	8	0	average		3.5	0.75	low	y
	average	8	0.2509			3.5	0		
	high	8	0.6997		n	3.5	0		
Input 7.1	small-av	50	0.25	average		60	0.0624	average	n
	large	50	0.7857		n	60	1		
Input 7.2	small	50	0	large	y	20	0	small-av	y
	av-large	50	1			20	1		
Input 7.3	small-av	5	0.9583	small-av	y	20	0.8333	average	y
	large	5	0.05			20	0.2		
Input 7.4	small					20	0.3333	small	n
	average					20	1		
	large					20	0		
Input 7.5	small	10	1	small-av	y	20	0.6667	small	y
	average	10	0			20	0.3333		
	large	10	0			20	0		
Input 8	low	8	0	average-high		3.5	0.8333	low	y
	average	8	0.6667			3.5	0		
	high	8	1		y	3.5	0		
Input 8.1	small-av	1	1	average	y	2	0.6667	small	y
	large	1	0.2			2	0.4		
Input 8.2	small-av	1	1	average	y	1	1	small	y
	large	1	0.4167			1	0.4167		
Input 8.3	small-av	1	1	average	y	3	0.9448	average	y
	large	1	0			3	0		
Input 8.4	small-av	3	0.7333	average	y	2	1	small	y
	large	3	0.1111			2	0		
Input 8.5	short-av	30	1	average	y	14	1	average	y
	long	30	1			14	0.8667		

Input 8.6	small-av	0	1	average	y	1	0.8667	small	n
	large	0	0			1	1		
Input 10	poor	8	0.125	good		6.5	0.3125	average-goo	y
	average	8	0.5		y	6.5	0.875		
	good	8	0.5			6.5	0		
Input 10.1	short	1	1	average	n	15	0	average	n
	average	1	0.2453			15	0.4444		
	long	1	0			15	1		
Input 10.2	small	0	1	average	n				
	average	0	0						
	large	0	0.435						
Input 10.3	small	0	1	average	n	4	0	large	y
	average	0	0			4	0.3333		
	large	0	0			4	1		
Input 10.4	small	30	0	large	y	20	0	large	n
	average	30	0.3333			20	1		
	large	30	1			20	0		
Input 10.5	risk-ave	7	0	average	n	5.5	0.1667	average	y
	average	7	0			5.5	0.5		
	risk-pro	7	0.8			5.5	0.5		
Input 11	poor-av	7	0.5	average	y	5.5	0.875	average	y
	good	7	0.5			5.5	0.25		
Input 11.1	short					30	0	average	n
	average					30	0.8571		
	long					30	1		
Input 11.2	small	1	1	average	n	50	1	small	y
	av-large	1	0			50	0.3333		
Input 11.3	small	1	0.8667	average	n				
	average	1	0						
	large	1	0						
Input 12	low	1	1	average	n	7.5	0	average-hi	y
	average	1	0			7.5	0		
	high	1	0			7.5	0.75		
Input 12.1	small-av	1	1	average	y	10	0.72	average	y
	large	1	0			10	0		
Input 12.2	small-av	0	1	average	y	60	0	large	y
	large	0	0.4118			60	0.7647		
Input 13	low	7	0	average	n	5.5	0	average	y
	average	7	0.3333			5.5	0.8333		
	high	7	0.6667			5.5	0.1667		
Input 13.1	small	100	1	average	y				
	av-large	100	1						
Input 13.2	small-av	1	1	average	y	2	1	average	y
	large	1	0			2	0		
Input 13.3	small	1	1	average	n	4	0	average	y
	av-large	1	0.6703			4	1		
Input 13.4	small	6	0	average	y				
	average	6	1						
	large	6	0						
Input 13.5	small	2	1	average	n	4	0	average	y
	average	2	0			4	1		

	large	2	0			4	0		
Input 13.6	small-av	5	0.6667	average	y	20	0	average	n
	large	5	0.4706			20	0.6471		
Input 13.7	small								
	average								
	large								
Input 14	unfavorabl	5	0	average	y	2.5	1	unfavorable	y
	average	5	1			2.5	0		
	favorable	5	0.25			2.5	0		
Output 1	poor	7	0	average	y	8.5	0	good	y
	average	7	0.5			8.5	0		
	good	7	0.5			8.5	1		
Output 1.1	small	5	1	average	n	5	1	average	n
	av-large	5	0			5	0		
Output 1.2	small					98	0	average	y
	av-large					98	1		
Output 1.3	small	5	1	average	y	2	1	small	y
	av-large	5	1			2	0		
Output 1.4	small-av	5	1	average	y				
	large	5	0						
Output 1.5	small-av					98	0.177	average	n
	large					98	0.974		
Output 1.6	small-av	10	1	average	y	5	1	small-av	y
	large	10	0			5	0		
Output 2	poor-av	7	0.5	average	y	6.5	0.5833	average-goo	y
	good	7	0.5			6.5	0.25		
Output 2.1	small-av					5	0.9333	small-averag	y
	large					5	0.0667		
Output 2.1	small								
	average								
	large								
Output 2.2	small					5	0.75	small	y
	average					5	0.25		
	large					5	0		
Output 3	low	7	0	average	y	8.5	0	high	y
	average	7	0.75			8.5	0.375		
	high	7	0.75			8.5	1		
Output 3.1	small	20	0	average	y	80	0	large	y
	av-large	20	1			80	1		
Output 3.2	small	10	1	average	n				
	large-av	10	0						
Output 3.3	small-av	25	0	average	n	5	1	small-av	y
	large	25	0.625			5	0.4583		
Output 3.4	small	500	0	average	y	250	0.75	small	y
	average	500	0.6429			250	0.4643		
	large	500	0			250	0		
Output 3.5	small-av					5	1	small	y
	large					5	0		
Input 9	low					3.5	0.5	low-average	y
	average					3.5	0.5		
	high					3.5	0		

Input 9.1	small	0.15	0.8333	small	y	0.25	0.7222	small	y
	average	0.15	0.0033			0.25	0.0056		
	large	0.15	0			0.25	0		
Input 9.2	small-av	0.2	0.8496	small	y	0.175	0.8684	small-av	y
	large	0.2	0.005			0.175	0.0044		
Input 9.3	small-av					0	1	small	y
	large					0	0		
Input 9.4	small-av					2.75	0.5877	average	y
	large					2.75	0.4123		
Input 9.5	small-av					0.3	0.9625	small	y
	large					0.3	0.0375		
Input 9.6	small					0.35	0.5625	small	y
	average					0.35	0.4375		
	large					0.35	0		
Input 9.7	small					0.45	0	small	n
	average					0.45	0.0625		
	large					0.45	0		
Input 9.8	small-av	0	1	average	y	0	1	small	y
	large	0	0			0	0		

No.	Name	test 3	f (test3)	actual term	match (y/n)	test 4	f (test4)	actual term	match
Input 1	small	5	0.3333	small	n	1	1	small	y
	average	5	1			1	0		
	large	5	0.25			1	0		
Input 1.1	small	40	1	average	n				
	av-large	40	0						
Input 1.2	small	15	0.1429	small	n				
	av-large	15	0.5663						
Input 1.3	small	10	0.5	small	y	25	0	small	n
	average	10	0.5213			25	0.6011		
	large	10	0			25	0		
Input 2	low	7	0.25	average	y	8	0	av-high	y
	average	7	0.7333			8	0.6		
	high	7	0.75			8	1		
Input 2.1	small	10	0	average	y	9	0	average	y
	av-large	10	1			9	0.8883		
Input 2.2	small	100	0	average	y				
	average	100	1						
	large	100	0						
Input 3	poor	3	1	poor	y	8.5	0	good	y
	average	3	0			8.5	0		
	good	3	0			8.5	1		
Input 3.1	small	10	0	large	n	10	0	large	n
	average	10	0			10	0		
	large	10	0.4286			10	0.4286		
Input 3.2	small	4	1	average	n	2	1	small	y
	av-large	4	0.45			2	0.35		
Input 3.3	small-av	50	1	average	y				
	large	50	1						
Input 3.4	poor	7	0	average	y	9.5	0	good	y
	average	7	0.8889			9.5	0.8333		
	good	7	0			9.5	1		
Input 3.5	small	8	0.4	average	y	20	0	large	y
	av-large	8	0.6			20	1		
Input 3.6	poor	7	0	average	y	8	0	good	y
	average	7	0.6667			8	0.5		
	good	7	0.3333			8	0.6667		
Input 3.7	small	12	1	average	n	30	0	large	y
	av-large	12	0			30	1		
Input 3.8	small-av	1	0.8667	small	y	0	1	small	y
	large	1	0			0	0		
Input 3.9	low	10	0	high	y	5	0.4545	averagae	n
	av-high	10	1			5	0		
Input 4	small-av	9	0.4444	large	y	7	0.6667	av-large	y
	large	9	1			7	0.75		
Input 4.1	small	22	0	average	y	0.65	1	average	n
	av-large	22	1			0.65	0.315		

Input 4.2	short-av	9	1	short	y	10	1	ave-long	y
	long	9	1			10	1		
Input 4.3	small	14.5	0	average	y				
	av-large	14.5	1						
Input 5	small	8	0	average	n	8	0	ave-large	y
	average	8	0.5			8	0.5		
	large	8	1			8	1		
Input 5.1	small	50	0	large	y	7.5	1	small	y
	average	50	0.67			7.5	0.4637		
	large	50	1			7.5	0		
Input 5.2	small-av	40	1	average	y	10	1	small	y
	large	40	1			10	0		
Input 6	low	2	1	low	y	7	0	average	n
	average	2	0.75			7	0		
	high	2	0			7	0.8333		
Input 6.1	low	8	0	average	y	8	0	ave-high	y
	av-high	8	1			8	1		
Input 6.2	low	3	0.5	low	y	7	0	average	y
	av-high	3	0.1667			7	0.8333		
Input 6.3	small-av	5	1	small	y	50	0.7	average	y
	large	5	0			50	0		
Input 6.4	small	1	1	small	y	3	1	small	y
	av-large	1	0			3	0.1667		
Input 7	low	8	0	av-high	y	6	0.3333	average	y
	average	8	0.2509			6	1		
	high	8	0.6997			6	0.0991		
Input 7.1	small-av	30	0.625	large	n	80	0	average	n
	large	30	0.3572			80	1		
Input 7.2	small	40	0	large	y	20	0	large	y
	av-large	40	1			20	1		
Input 7.3	small-av	20	0.8333	small-av	y	0	1	average	y
	large	20	0.2			0	0		
Input 7.4	small	30	0	average	y	20	0.3333	small	n
	average	30	1			20	1		
	large	30	0.5			20	0		
Input 7.5	small	30	0.3333	average	y				
	average	30	0.6667						
	large	30	0.2						
Input 8	low	9	0	av-high	y	8	0	ave-high	y
	average	9	0.5556			8	0.6667		
	high	9	1			8	1		
Input 8.1	small-av	4	0	large	y	2	0.6667	small	y
	large	4	0.8			2	0.4		
Input 8.2	small-av	4	0	large	y	4	0	large	y
	large	4	0.6667			4	0.6667		
Input 8.3	small-av	1	1	average	y	1	1	average	y
	large	1	0			1	0		
Input 8.4	small-av	4	0.4667	large	y	7	0	large	y
	large	4	0.5556			7	1		
Input 8.5	short-av	1	1	short	y				
	long	1	0						

Input 8.6	small-av	1	0.8667	small	n	1	0.8667	average	y
	large	1	1			1	1		
Input 10	poor	4	0.625	poor-av	y	6	0.375	average	y
	average	4	0.25			6	0.75		
	good	4	0			6	0		
Input 10.1	short	1	1	short	y	5	0	short	n
	average	1	0.2453			5	1		
	long	1	0			5	0		
Input 10.2	small	20	0	large	y	2	0.7778	small	y
	average	20	0			2	0		
	large	20	0.548			2	0.4463		
Input 10.3	small	2	0.1667	average	y	0	1	small	y
	average	2	0.8333			0	0		
	large	2	0.3333			0	0		
Input 10.4	small	20	0	average	y	0	1	small	y
	average	20	1			0	0		
	large	20	0			0	0		
Input 10.5	risk-ave	1	1	risk-ave	y	1	1	risk-ave	y
	average	1	0			1	0		
	risk-pro	1	0			1	0		
Input 11	poor-av	8	0.25	average	n	5	1	average	y
	good	8	0.6667			5	0.1667		
Input 11.1	short	20	0	average	n	20	0	average	y
	average	20	0.7143			20	0.7143		
	long	20	1			20	1		
Input 11.2	small	80	0.4	average	y	100	0	large	y
	av-large	80	1			100	1		
Input 11.3	small	5	0.3333	average	y				
	average	5	1						
	large	5	0						
Input 12	low	6	0	average	y	5	0.25	average	y
	average	6	0.5			5	1		
	high	6	0.5			5	0.3333		
Input 12.1	small-av	2	1	average	y	8	0.8	ave-large	y
	large	2	0			8	0		
Input 12.2	small-av	10	0.6667	small	y	0	1	small-ave	y
	large	10	0.4706			0	0.4118		
Input 13	low	8	0	av-high	y	5.5	0	average	y
	average	8	0			5.5	0.8333		
	high	8	1			5.5	0.1667		
Input 13.1	small	80	1	small-av	y	95	1	ave-large	y
	av-large	80	0.7761			95	1		
Input 13.2	small-av	1	1	average	y				
	large	1	0						
Input 13.3	small	2	0.2308	average	y	1	1	average	n
	av-large	2	0.7802			1	0.6703		
Input 13.4	small	3	0.5	average	y	4	0	average	y
	average	3	0.5			4	1		
	large	3	0			4	0		
Input 13.5	small	40	0	large	y	24	0	large	y
	average	40	0.3077			24	0.6154		

	large	40	1			24	1		
Input 13.6	small-av	10	0.3333	small-av	n	30	0	large	y
	large	10	0.5294			30	0.7647		
Input 13.7	small								
	average								
	large								
Input 14	unfavorable	8	0	favorable	y	8	0	favorable	y
	average	8	0.5			8	0.5		
	favorable	8	1			8	1		
Output 1	poor	10	0	good	y	9	0	good	y
	average	10	0			9	0		
	good	10	1			9	1		
Output 1.1	small	25	1	large	y	30	0.8	large	y
	av-large	25	1			30	1		
Output 1.2	small	95	0	large	y	25	0	large	y
	av-large	95	1			25	1		
Output 1.3	small	5	1	average	y	5	1	average	y
	av-large	5	1			5	1		
Output 1.4	small-av	60	0	large	y				
	large	60	1						
Output 1.5	small-av	95	0.2035	large	y	25	0.823	large	n
	large	95	0.9351			25	0.026		
Output 1.6	small-av	11	1	average	y	1	1	small	y
	large	11	0			1	0		
Output 2	poor-av	10	0	good	y	10	0	good	y
	good	10	1			10	1		
Output 2.1	small-av	10	0.8667	average	y				
	large	10	0.1333						
Output 2.1	small								
	average								
	large								
Output 2.2	small	30	0	large	y	0	1	small	y
	average	30	0.5			0	0		
	large	30	0.5			0	0		
Output 3	low	7	0	low-av	y	8	0	ave-high	y
	average	7	0.75			8	0.5		
	high	7	0.75			8	1		
Output 3.1	small	40	0	large	y	10	1	average	n
	av-large	40	1			10	0		
Output 3.2	small	40	0.3333	large	y	50	0.1111	large	y
	large-av	40	0.6			50	1		
Output 3.3	small-av	40	0	large	y	5	1	average	y
	large	40	0.75			5	0.4583		
Output 3.4	small	300	0.5	average	y	80	1	average	n
	average	300	0.5			80	0.3429		
	large	300	0			80	0		
Output 3.5	small-av	20	0	large	y	0	1	small	y
	large	20	1			0	0		
Input 9	low	9	0	high	y				
	average	9	0.3333						
	high	9	1						

Input 9.1	small	0	1	small	y	0	1	small	y
	average	0	0			0	0		
	large	0	0			0	0		
Input 9.2	small-av	0.4	0.6992	small	y	3.5	0	small	n
	large	0.4	0.01			3.5	0.0875		
Input 9.3	small-av	10	0	large	y	0.3	0.9375	small	y
	large	10	1			0.3	0		
Input 9.4	small-av	0.6	0.91	small	y	4	0.4003	average	n
	large	0.6	0.09			4	0.5997		
Input 9.5	small-av	0	1	small	y	10	0	average	n
	large	0	0			10	1		
Input 9.6	small	0	1	small	y	0	1	small	y
	average	0	0			0	0		
	large	0	0			0	0		
Input 9.7	small	0	1	small	y	0	1	small	y
	average	0	0			0	0		
	large	0	0			0	0		
Input 9.8	small-av	0	1	small	y	0	1	small	y
	large	0	0			0	0		

No.	Name	test 5	f (test5)	actual term	match (y/n)	test 6	f (test6)	actual term	match (y/n)
Input 1	small	3	1	small	y	3	1	average	n
	average	3	0			3	0		
	large	3	0			3	0		
Input 1.1	small	20	1	small	y	6	1	average	n
	av-large	20	0			6	0		
Input 1.2	small	2	1	average	n				
	av-large	2	0.5						
Input 1.3	small	5	0.9167	average	n	5	0.9167	small	y
	average	5	0.4947			5	0.4947		
	large	5	0			5	0		
Input 2	low	8	0	high	y	8	0	high	y
	average	8	0.6			8	0.6		
	high	8	1			8	1		
Input 2.1	small	10	0	large	y	4	0	average	n
	av-large	10	1			4	0.3296		
Input 2.2	small	10	1	average	n				
	average	10	0						
	large	10	0						
Input 3	poor	9	0	good	y	7	0	ave-good	y
	average	9	0			7	0.2509		
	good	9	1			7	0.6997		
Input 3.1	small	9	0	average	n	8	0	average	y
	average	9	0.25			8	0.5		
	large	9	0.3571			8	0.2857		
Input 3.2	small	7	1	average	n	5	1	average	n
	av-large	7	0.6			5	0.5		
Input 3.3	small-av	0.33	1	average	y	80	0.3333	large	y
	large	0.33	0			80	1		
Input 3.4	poor	8	0	good	n	8	0	average	y
	average	8	1			8	1		
	good	8	0.5			8	0.5		
Input 3.5	small	12	0	average	y	10	0	average	y
	av-large	12	1			10	1		
Input 3.6	poor	9	0	good	y	8	0	average	n
	average	9	0.3333			8	0.5		
	good	9	1			8	0.6667		
Input 3.7	small	20	0.4686	large	y	25	0	large	y
	av-large	20	0.5314			25	1		
Input 3.8	small-av	1	0.8667	small	y	0	1	average	y
	large	1	0			0	0		
Input 3.9	low	9	0	high	y	8	0	average	y
	av-high	9	0.8			8	0.6		
Input 4	small-av	6	0.7778	average	y	8	0.5556	ave-large	y
	large	6	0.5			8	1		
Input 4.1	small	1.1	1	average	n	2.5	0.75	small	y
	av-large	1.1	0.36			2.5	0.5		

Input 4.2	short-av	4	1	average	y		10	1	average	y
	long	4	0				10	1		
Input 4.3	small	2	1	average	n		3.5	1	small-ave	y
	av-large	2	0.4286				3.5	0.5357		
Input 5	small	9	0	large	y		8	0	ave-large	y
	average	9	0.3333				8	0.5		
	large	9	1				8	1		
Input 5.1	small	40	0	average	y		40	0	ave-large	y
	average	40	1				40	1		
	large	40	1				40	1		
Input 5.2	small-av	200	1	average	y		150	1	average	n
	large	200	1				150	1		
Input 6	low	6	0	average	n		3	0.5	low-ave	y
	average	6	0.25				3	1		
	high	6	0.6667				3	0.1667		
Input 6.1	low	8	0	high	y		4	0.5	low-ave	y
	av-high	8	1				4	0.4638		
Input 6.2	low	8	0	high	y		4	0	low-ave	n
	av-high	8	1				4	0.3333		
Input 6.3	small-av	70	0.5	large	y		55	0.65	average	y
	large	70	0.5				55	0		
Input 6.4	small	7	0	large	y		5	0.5	ave-large	y
	av-large	7	0.8333				5	0.5		
Input 7	low	6	0.3333	average	y		5	0.5	average	y
	average	6	1				5	0.5		
	high	6	0.0991				5	0		
Input 7.1	small-av	30	0.625	average	y		25	0.7187	average	y
	large	30	0.3572				25	0.2501		
Input 7.2	small	20	0	small	n		65	0	ave-large	y
	av-large	20	1				65	1		
Input 7.3	small-av	70	0.4167	large	y		65	0.4583	average	n
	large	70	0.7				65	0.65		
Input 7.4	small	50	0	average	n		65	0	ave-large	n
	average	50	0				65	0		
	large	50	1				65	1		
Input 7.5	small	10	1	small	y		30	0.3333	small	n
	average	10	0				30	0.6667		
	large	10	0				30	0.2		
Input 8	low	5	0.3333	average	y		8	0	ave-high	y
	average	5	1				8	0.6667		
	high	5	0.25				8	1		
Input 8.1	small-av	1	1	small	y		9	0	large	y
	large	1	0.2				9	1		
Input 8.2	small-av	1	1	small	y		6	0	large	y
	large	1	0.4167				6	0.8333		
Input 8.3	small-av	1	1	small	y		3	0.9448	large	n
	large	1	0				3	0		
Input 8.4	small-av	1	1	small	y		2	1	average	y
	large	1	0				2	0		
Input 8.5	short-av	2	1	short	y					
	long	2	0							

Input 8.6	small-av					1	0.8667	small	n
	large					1	1		
Input 10	poor	10	0	good	y	8	0.125	good	y
	average	10	0			8	0.5		
	good	10	1			8	0.5		
Input 10.1	short	6	0	average	y	2.5	0.4444	short	n
	average	6	0.9444			2.5	0.5283		
	long	6	0.2			2.5	0		
Input 10.2	small	5	0.4444	small	n	2.5	0.7222	small	y
	average	5	0.6667			2.5	0		
	large	5	0.4633			2.5	0.4492		
Input 10.3	small	3	0	large	y	2	0.1667	average	y
	average	3	0.75			2	0.8333		
	large	3	1			2	0.3333		
Input 10.4	small	15	0.2941	average	y				
	average	15	1						
	large	15	0						
Input 10.5	risk-ave	1	1	risk-ave	y	4	0.6667	risk-ave	y
	average	1	0			4	0		
	risk-pro	1	0			4	0.2		
Input 11	poor-av	8	0.25	good	y	4	1	poor-ave	y
	good	8	0.6667			4	0		
Input 11.1	short					4.5	0.6497	long	n
	average					4.5	0.4929		
	long					4.5	0		
Input 11.2	small					75	0.5	ave-large	y
	av-large					75	1		
Input 11.3	small					15	0	small-ave	y
	average					15	1		
	large					15	0.3847		
Input 12	low	2	1	low	y	7	0	ave-high	y
	average	2	0			7	0		
	high	2	0			7	0.6667		
Input 12.1	small-av	3	1	small	y	1	1	average	y
	large	3	0			1	0		
Input 12.2	small-av	10	0.6667	small	y	20	0.3333	average	n
	large	10	0.4706			20	0.5294		
Input 13	low	2	0.6667	low	y	6	0	average	y
	average	2	0.3304			6	0.6667		
	high	2	0			6	0.3333		
Input 13.1	small	100	1	large	y	50	1	small-ave	y
	av-large	100	1			50	0.3284		
Input 13.2	small-av	2	1	small	y	2	1	average	y
	large	2	0			2	0		
Input 13.3	small	1	1	small	y	1	1	average	n
	av-large	1	0.6703			1	0.6703		
Input 13.4	small	3	0.5	small	y	8.5	0	average	y
	average	3	0.5			8.5	1		
	large	3	0			8.5	0		
Input 13.5	small	1	1	small	y	0.5	1	small	y
	average	1	0			0.5	0		

	large	1	0			0.5	0		
Input 13.6	small-av	90	0	large	y	65	0	ave-large	y
	large	90	1			65	1		
Input 13.7	small								
	average								
	large								
Input 14	unfavorabl	5	0	favorable	n	7	0	favorable	n
	average	5	1			7	1		
	favorable	5	0.25			7	0.75		
Output 1	poor	9	0	good	y	7	0	ave-good	y
	average	9	0			7	0.5		
	good	9	1			7	0.5		
Output 1.1	small	25	1	average	y	12	1	average	y
	av-large	25	1			12	1		
Output 1.2	small	90	0	large	y	15	0.5	small-ave	y
	av-large	90	1			15	1		
Output 1.3	small	10	1	small	y	1	1	small	y
	av-large	10	1			1	0		
Output 1.4	small-av					25	0.4	large	y
	large					25	1		
Output 1.5	small-av	25	0.823	small	y	20	0.8673	average	y
	large	25	0.026			20	0		
Output 1.6	small-av	25	0.4375	average	n	5.5	1	small	y
	large	25	1			5.5	0		
Output 2	poor-av	9	0.1667	good	y	8	0.3333	ave-good	y
	good	9	1			8	1		
Output 2.1	small-av	10	0.8667	small	y	0	1	average	y
	large	10	0.1333			0	0		
Output 2.1	small								
	average								
	large								
Output 2.2	small	1	0.95	small	y	0	1	small	y
	average	1	0.05			0	0		
	large	1	0			0	0		
Output 3	low	9	0	high	y	8	0	high	y
	average	9	0.25			8	0.5		
	high	9	1			8	1		
Output 3.1	small	7	1	average	n	22.5	0	average	y
	av-large	7	0			22.5	1		
Output 3.2	small					40	0.3333	average	y
	large-av					40	0.6		
Output 3.3	small-av					10	0.5	average	y
	large					10	0.5		
Output 3.4	small					150	1	small	y
	average					150	0.3929		
	large					150	0		
Output 3.5	small-av					7.5	0.75	small	y
	large					7.5	0.25		
Input 9	low	3	0.6667	low	y	6	0	average	y
	average	3	0.3333			6	0.8333		
	high	3	0			6	0.5		

Input 9.1	small					0	1	small-ave	y
	average					0	0		
	large					0	0		
Input 9.2	small-av					3.75	0	average	n
	large					3.75	0.0938		
Input 9.3	small-av					0	1	small	y
	large					0	0		
Input 9.4	small-av					1.25	0.8126	average	y
	large					1.25	0.1874		
Input 9.5	small-av					0	1	small	y
	large					0	0		
Input 9.6	small					0.15	0.8125	small	y
	average					0.15	0.1875		
	large					0.15	0		
Input 9.7	small	0.1	0.3333	small	n	0	1	small	y
	average	0.1	0.0139			0	0		
	large	0.1	0			0	0		
Input 9.8	small-av					0	1	small	y
	large					0	0		

No.	test 7	f (test7)	actual term	match (y/n)
Input 1	5	0.3333	average	y
	5	1		
	5	0.25		
Input 1.1	10	1	average	n
	10	0		
Input 1.2				
Input 1.3	5	0.9167	average	n
	5	0.4947		
	5	0		
Input 2	8	0	average	n
	8	0.6		
	8	1		
Input 2.1	10	0	average	y
	10	1		
Input 2.2	10	1	small	y
	10	0		
	10	0		
Input 3	4	0.6667	poor	y
	4	0		
	4	0		
Input 3.1	10	0	average	n
	10	0		
	10	0.4286		
Input 3.2	3	1	average	n
	3	0.4		
Input 3.3	50	1	small	y
	50	1		
Input 3.4	9	0	good	y
	9	0.8889		
	9	1		
Input 3.5	15	0	large	y
	15	1		
Input 3.6	9	0	good	y
	9	0.3333		
	9	1		
Input 3.7	15	0.9372	average	n
	15	0.0628		
Input 3.8	0	1	small	y
	0	0		
Input 3.9	9	0	high	y
	9	0.8		
Input 4	8	0.5556	large	y
	8	1		
Input 4.1	1.87	1	large	n
	1.87	0.437		

Input 4.2	6	1	long	n
	6	0.3333		
Input 4.3	1.5	1	average	n
	1.5	0.3929		
Input 5	8	0	average	n
	8	0.5		
	8	1		
Input 5.1	35	0	average	n
	35	0.9175		
	35	1		
Input 5.2	55	1	average	y
	55	1		
Input 6	4	0	low	n
	4	0.75		
	4	0.3333		
Input 6.1	5	0.25	low	n
	5	0.5979		
Input 6.2	7	0	average	y
	7	0.8333		
Input 6.3	85	0.35	large	y
	85	1		
Input 6.4	7	0	average	y
	7	0.8333		
Input 7	6	0.3333	average	y
	6	1		
	6	0.0991		
Input 7.1	50	0.25	average	n
	50	0.7857		
Input 7.2	10	1	average	y
	10	1		
Input 7.3	0	1	small	y
	0	0		
Input 7.4	30	0	average	y
	30	1		
	30	0.5		
Input 7.5				
Input 8	7	0	average	y
	7	0.7778		
	7	0.75		
Input 8.1	2	0.6667	average	y
	2	0.4		
Input 8.2	2	0.5	average	y
	2	0.5		
Input 8.3	2	0.9724	average	y
	2	0		
Input 8.4	5	0.2	large	y
	5	1		
Input 8.5	75	0	long	y
	75	1		

Input 8.6	2	0.7333	average	n
	2	1		
Input 10	3	0.75	average	n
	3	0		
	3	0		
Input 10.1	1	1	short	y
	1	0.2453		
	1	0		
Input 10.2	0	1	small	y
	0	0		
	0	0.435		
Input 10.3	3	0	large	y
	3	0.75		
	3	1		
Input 10.4	0	1	small	y
	0	0		
	0	0		
Input 10.5	5	0.3333	average	y
	5	1		
	5	0.4		
Input 11	8	0.25	good	y
	8	0.6667		
Input 11.1	90	0	long	y
	90	0		
	90	1		
Input 11.2	25	1	small	y
	25	0		
Input 11.3	3	0.6	average	n
	3	0.2		
	3	0		
Input 12	5	0.25	average	y
	5	1		
	5	0.3333		
Input 12.1	1	1	small	y
	1	0		
Input 12.2	10	0.6667	small	y
	10	0.4706		
Input 13	7	0	average	n
	7	0.3333		
	7	0.6667		
Input 13.1	100	1	average	y
	100	1		
Input 13.2	1	1	small	y
	1	0		
Input 13.3	1	1	small	y
	1	0.6703		
Input 13.4	4	0	average	y
	4	1		
	4	0		
Input 13.5	28	0	average	n
	28	0.5385		

	28	1		
Input 13.6	0	1	small	y
	0	0.4118		
Input 13.7	0	1	small	y
	0	0		
	0	0		
Input 14	9	0	favorable	y
	9	0		
	9	1		
Output 1	8	0	good	y
	8	0		
	8	1		
Output 1.1	20	1	large	y
	20	1		
Output 1.2	25	0	large	y
	25	1		
Output 1.3	15	0.5	average	y
	15	1		
Output 1.4	20	0.6	large	y
	20	1		
Output 1.5	0	1	small	y
	0	0		
Output 1.6	12	1	average	y
	12	0		
Output 2	7	0.5	average	y
	7	0.5		
Output 2.1	20	0.7333	average	y
	20	0.2667		
Output 2.1 (-)				
Output 2.2	20	0	average	y
	20	1		
	20	0		
Output 3	9	0	high	y
	9	0.25		
	9	1		
Output 3.1	5	1	small	y
	5	0		
Output 3.2	2	1	small	y
	2	0		
Output 3.3	1	1	small	y
	1	0.425		
Output 3.4	0	1	small	y
	0	0.2857		
	0	0		
Output 3.5	1	1	small	y
	1	0		
Input 9	7	0	average	n
	7	0.6667		
	7	0.75		

Input 9.1	0	1	small	y
	0	0		
	0	0		
Input 9.2	8	0	large	n
	8	0.2		
Input 9.3	0	1	average	y
	0	0		
Input 9.4	0	1	average	y
	0	0		
Input 9.5	0	1	average	y
	0	0		
Input 9.6	8	0	large	y
	8	0		
	8	1		
Input 9.7	0	1	small	y
	0	0		
	0	0		
Input 9.8	0	1	small	y
	0	0		

Appendix 7: Membership Function Testing Comparisons

	Case 1(%)		Case 2 (%)	
Input 1	100	1	67	1
Input 1.1	100	1	33	0
Input 1.2	75	1	33	0
Input 1.3	80	1	28.57	0
Input 2	80	1	85.7	1
Input 2.1	60	1	85.7	1
Input 2.2	80	1	75	1
Input 3	100	1	100	1
Input 3.1	40	0	0	0
Input 3.2	75	1	40	0
Input 3.3	80	1	50	0
Input 3.4	100	1	71.42	1
Input 3.5	60	1	85.7	1
Input 3.6	60	1	85.70	1
Input 3.7	100	1	71.42	1
Input 3.8	80	1	85.7	1
Input 3.9	100	1	71.42	1
Input 4	80	1	100	1
Input 4.1	80	1	50	0
Input 4.2	100	1	57.14	0
Input 4.3	60	1	67	1
Input 5	40	0	71.42	1
Input 5.1	80	1	85.7	1
Input 5.2	75	1	100	1
Input 6	40	0	42.86	0
Input 6.1	60	1	85.71	1
Input 6.2	60	1	85.71	1
Input 6.3	80	1	100	1
Input 6.4	100	1	100	1
Input 7	60	1	85.71	1
Input 7.1	50	0	28.57	0
Input 7.2	80	1	71.42	1
Input 7.3	80	1	85.71	1
Input 7.4	80	1	33	0
Input 7.5	100	1	80	1
Input 8	50	0	100	1
Input 8.1	60	1	100	1
Input 8.2	50	0	100	1
Input 8.3	75	1	85.71	1
Input 8.4	80	1	100	1
Input 8.5	80	1	100	1

Input 8.6	50	0	33	0
Input 10	60	1	71.42	1
Input 10.1	100	1	42.85	0
Input 10.2	50	0	67	1
Input 10.3	80	1	85.71	1
Input 10.4	100	1	83.33	1
Input 10.5	100	1	85.71	1
Input 11	100	1	85.71	1
Input 11.1	60	1	33	0
Input 11.2	100	1	83.33	1
Input 11.3	100	1	25	0
Input 12	100	1	85.71	1
Input 12.1	100	1	85.71	1
Input 12.2	67	1	85.71	1
Input 13	33	0	71.42	1
Input 13.1	100	1	100	1
Input 13.2	100	1	100	1
Input 13.3	75	1	57.14	0
Input 13.4	100	1	100	1
Input 13.5	25	0	71.42	1
Input 13.6	75	1	71.43	1
Input 13.7	100	1	100	1
Input 14	60	1	71.43	1
Output 1	80	1	100	1
Output 1.1	80	1	71.43	1
Output 1.2	20	0	83.33	1
Output 1.3	80	1	100	1
Output 1.4	50	0	100	1
Output 1.5	25	0	67	1
Output 1.6	100	1	85.71	1
Output 2	80	1	100	1
Output 2.1	100	1	100	1
Output 2.2	100	1	100	1
Output 3	60	1	85.71	1
Output 3.1	33	0	71.43	1
Output 3.2	33	0	67	1
Output 3.3	33	0	83.33	1
Output 3.4	100	1	83.33	1
Output 3.5	100	1	100	1
Input 9	75	1	80	1
Input 9.1	100	1	100	1
Input 9.2	40	0	50	0
Input 9.3	75	1	100	1
Input 9.4	75	1	80	1
Input 9.5	100	1	80	1

Input 9.6	100	1	100	1
Input 9.7	100	1	67	1
Input 9.8	100	1	100	1
total =	89	72		72
		80.89888		80.89888

Note:	1	means	successful
	0	means	failure

Appendix 8: Correlation Analysis Results

1. sub-input factors to input factors

1) input 1

Sub-factor	Sig (2 tailed)	Keep (Y/N)
11	**0.005**	Y
12	**0.035**	Y
13	0.128	N

2) input 2

Sub-factor	Sig (2 tailed)	Keep (Y/N)
21	0.236	N
22	**0.023**	Y

3) input 3

Sub-factor	Sig (2 tailed)	Keep (Y/N)
31	.720	N
32	.481	N
33	.256	N
34	.273	N
35	.264	N
36	.111	N
37	.137	N
38	**.079**	Y
39	.638	N

4) input 4

Sub-factor	Sig (2 tailed)	Keep (Y/N)
41	0.637	N
42	**0.048**	Y
43	0.992	N

5) input 5

Sub-factor	Sig (2 tailed)	Keep (Y/N)
51	**0.092**	**Y**
52	0.311	N

6) input 6

Sub-factor	Sig (2 tailed)	Keep (Y/N)
61	**0.011**	**Y**
62	**0.082**	**Y**
63	0.263	N
64	0.238	N

7) input 7

Sub-factor	Sig (2 tailed)	Keep (Y/N)
71	0.978	N
72	0.259	N
73	0.462	N
74	0.813	N
75	0.452	N

Ignore this model.

8) input 8

Sub-factor	Sig (2 tailed)	Keep (Y/N)
81	0.485	N
82	**0.110**	**Y**
83	0.688	N
84	**0.005**	**Y**
85	0.172	N
86	0.787	N

9) input 9

Sub-factor	Sig (2 tailed)	Keep (Y/N)
91	0.646	N
92	0.953	N
93	**0.060**	**Y**
94	0.539	N
95	0.420	N
96	0.767	N
97	0.330	N
98	0.544	N

10) input 10

Sub-factor	Sig (2 tailed)	Keep (Y/N)
101	**0.006**	**Y**
102	**0.019**	**Y**
103	0.967	N
104	0.496	N
105	0.877	N

11) input 11

Sub-factor	Sig (2 tailed)	Keep (Y/N)
111	0.977	N
112	0.676	N
113	0.132	N

Ignore this model.

12) input 12

Sub-factor	Sig (2 tailed)	Keep (Y/N)
121	**0.039**	**Y**
122	**0.011**	**Y**

13) input 13

Sub-factor	Sig (2 tailed)	Keep (Y/N)
131	0.732	N
132	0.973	N
133	0.539	N
134	0.358	N
135	0.274	N
136	0.241	N
137	0.891	N

Ignore this model.

2. sub-output factors to output factors

1) output 1

Sub-factor	Sig (2 tailed)	Keep (Y/N)
11	0.480	N
12	0.780	N
13	0.112	N
14	0.826	N
15	0.790	N
16	0.648	N

2) output 2

Sub-factor	Sig (2 tailed)	Keep (Y/N)
21	0.791	N
22	0.840	N

Ignore this model.

3) output 3

Sub-factor	Sig (2 tailed)	Keep (Y/N)
31	0.536	N
32	0.637	N
33	**0.086**	**Y**
34	0.909	N
35	0.988	N

3. input factors to each sub-output factor

1) sub-output 11

Factor	Sig (2 tailed)	Keep (Y/N)
1	0.165	N
2	0.539	N
3	0.157	N
4	0.175	N
5	0.989	N
6	0.219	N
7	0.660	N
8	0.477	N
9	0.764	N
10	0.258	N
11	**0.076**	**Y**
12	1.000	N
13	0.298	N
14	0.497	N

2) sub-output 12

Factor	Sig (2 tailed)	Keep (Y/N)
1	0.600	N
2	0.831	N
3	0.364	N
4	0.774	N
5	0.782	N
6	0.347	N
7	**0.074**	**Y**
8	0.140	N
9	**0.050**	**Y**
10	0.638	N
11	0.852	N
12	0.112	N
13	0.118	N
14	0.557	N

3) sub-output 13

Factor	Sig (2 tailed)	Keep (Y/N)
1	0.736	N
2	0.204	N
3	0.959	N
4	0.310	N
5	0.350	N
6	0.127	N
7	0.199	N
8	0.630	N
9	0.829	N
10	0.999	N
11	**0.024**	**Y**
12	0.282	N
13	0.446	N
14	0.355	N

4) sub-output 14

Factor	Sig (2 tailed)	Keep (Y/N)
1	0.144	N
2	0.759	N
3	0.386	N
4	0.534	N
5	0.596	N
6	**0.053**	**Y**
7	0.585	N
8	**0.085**	**Y**
9	0.787	N
10	0.590	N
11	0.381	N
12	0.287	N
13	0.589	N
14	**0.066**	**Y**

5) sub-output 15

Factor	Sig (2 tailed)	Keep (Y/N)
1	0.764	N
2	0.998	N
3	0.436	N
4	0.837	N
5	0.391	N
6	0.216	N
7	0.182	N
8	0.788	N
9	0.272	N
10	0.826	N
11	0.470	N
12	**0.029**	**Y**
13	0.955	N
14	0.689	N

6) sub-output 16

Factor	Sig (2 tailed)	Keep (Y/N)
1	0.753	N
2	0.847	N
3	0.625	N
4	0.992	N
5	**0.035**	**Y**
6	0.622	N
7	**0.057**	**Y**
8	0.697	N
9	0.735	N
10	0.160	N
11	**0.015**	**Y**
12	**0.006**	**Y**
13	0.783	N
14	0.892	N

7) sub-output 21

Factor	Sig (2 tailed)	Keep (Y/N)
1	**0.106**	**Y**
2	0.837	N
3	0.341	N
4	0.995	N
5	0.335	N
6	0.735	N
7	0.976	N
8	0.476	N
9	0.592	N
10	0.952	N
11	**0.081**	**Y**
12	0.317	N
13	0.949	N
14	0.742	N

8) sub-output 22

Factor	Sig (2 tailed)	Keep (Y/N)
1	0.546	N
2	0.396	N
3	**0.070**	**Y**
4	0.794	N
5	0.869	N
6	0.158	N
7	0.830	N
8	0.565	N
9	0.253	N
10	**0.054**	**Y**
11	0.314	N
12	0.683	N
13	0.214	N
14	0.233	N

9) sub-output 31

Factor	Sig (2 tailed)	Keep (Y/N)
1	0.140	N
2	0.144	N
3	0.858	N
4	**0.029**	**Y**
5	0.330	N
6	0.958	N
7	0.253	N
8	0.513	N
9	0.876	N
10	0.781	N
11	0.889	N
12	0.259	N
13	**0.075**	**Y**
14	0.881	N

10) sub-output 32

Factor	Sig (2 tailed)	Keep (Y/N)
1	0.901	N
2	0.483	N
3	0.931	N
4	**0.080**	**Y**
5	0.908	N
6	0.909	N
7	0.549	N
8	0.291	N
9	0.956	N
10	0.384	N
11	0.182	N
12	**0.069**	**Y**
13	0.489	N
14	0.182	N

11) sub-output 33

Factor	Sig (2 tailed)	Keep (Y/N)
1	0.544	N
2	0.405	N
3	0.171	N
4	0.596	N
5	0.249	N
6	0.231	N
7	0.856	N
8	0.253	N
9	0.706	N
10	0.825	N
11	0.278	N
12	0.166	N
13	0.880	N
14	0.656	N

12) sub-output 34

Factor	Sig (2 tailed)	Keep (Y/N)
1	0.562	N
2	0.504	N
3	0.363	N
4	0.159	N
5	0.734	N
6	0.359	N
7	0.353	N
8	0.448	N
9	0.597	N
10	0.715	N
11	0.920	N
12	0.925	N
13	0.618	N
14	0.709	N

13) sub-output 35

Factor	Sig (2 tailed)	Keep (Y/N)
1	**0.005**	**Y**
2	0.530	N
3	0.506	N
4	0.618	N
5	0.615	N
6	0.540	N
7	0.962	N
8	0.157	N
9	0.734	N
10	0.715	N
11	0.917	N
12	0.900	N
13	0.470	N
14	0.433	N

Appendix 9: Complete Rulebase

Model 1

Input 1.1	Input 1.2	Input 1
Small	Small	Small
Small	Average-large	Average
Average-large	Small	Average
Average-large	Average-large	Average

Model 2

Input 2.2	Input 2
Small	High
Average	Low
Large	Average

Model 3

Input 3.8	Input 3
Small-Average	Good
Large	Average

Model 4

Input 4.2	Input 4
Short-Average	Small-Average
Long	Large

Model 5

Input 5.1	Input 5
Small	Average
Average	Average
Large	Large

Model 6

Input 6.1	Input 6.2	Input 6
Low	Low	Low
Low	Average-High	Low
Average-high	Low	Low
Average-high	Average-High	Average

Model 7

Input 8.4	Input 8
Small-Average	Average
Large	High

Model 8

Input 9.3	Input 9
Small-Average	Average
Large	High

Model 9

Input 10.1	Input 10
Short	Good
Average	Average
Long	Poor

Model 10

Input 12.1	Input 12.2	Input 12
Small-Average	Small-Average	Average
Small-Average	Large	Average
Large	Small-Average	Average
Large	Large	High

Model 11

Input 6	Ouput 1.1
Low	Average-Large
Average	Average-Large
High	Small

Model 12

Input 7	Input 9	Output 1.2
Low	Low	Average-Large
Low	Average	Average-Large
Low	High	Average-Large
Average	Low	Average-Large
Average	Average	Average-Large
Average	High	Average-Large
High	Low	Average-Large
High	Average	Average-Large
High	High	Small

Model 13

Input 11	Output 1.3
Poor-Average	Small
Good	Average-Large

Model 14

Input 6	Input 8	Input 14	Output 1.4
Low	Low	unfavorable	Small-Average
Low	Low	average	Small-Average
Low	Low	favorable	Large
Low	Average	unfavorable	Small-Average
Low	Average	average	Small-Average
Low	Average	favorable	Small-Average
Low	High	unfavorable	Small-Average
Low	High	average	Small-Average
Low	High	favorable	Large
Average	Low	unfavorable	Small-Average
Average	Low	average	Small-Average
Average	Low	favorable	Small-Average
Average	Average	unfavorable	Small-Average
Average	Average	average	Small-Average
Average	Average	favorable	Small-Average
Average	High	unfavorable	Small-Average
Average	High	average	Small-Average
Average	High	favorable	Small-Average
High	Low	unfavorable	Large
High	Low	average	Small-Average
High	Low	favorable	Small-Average
High	Average	unfavorable	Small-Average
High	Average	average	Small-Average
High	Average	favorable	Small-Average
High	High	unfavorable	Small-Average
High	High	average	Small-Average
High	High	favorable	Large

Model 15

Input 12	Output 1.5
Low	Small-Average
Average	Small-Average
High	Large

Model 16

Input 5	Input 7	Input 11	Input 12	Output 1.6
Small	Low	Poor-Average	High	Small-Average
Small	Low	Poor-Average	Average	Small-Average
Small	Low	Poor-Average	Low	Small-Average
Small	Low	Good	High	Small-Average
Small	Low	Good	Average	Small-Average

Small	Low	Good	Low	Large
Small	Average	Poor-Average	High	Small-Average
Small	Average	Poor-Average	Average	Small-Average
Small	Average	Poor-Average	Low	Small-Average
Small	Average	Good	High	Small-Average
Small	Average	Good	Average	Small-Average
Small	Average	Good	Low	Large
Small	High	Poor-Average	High	Small-Average
Small	High	Poor-Average	Average	Small-Average
Small	High	Poor-Average	Low	Large
Small	High	Good	High	Large
Small	High	Good	Average	Large
Small	High	Good	Low	Large
Average	Low	Poor-Average	High	Small-Average
Average	Low	Poor-Average	Average	Small-Average
Average	Low	Poor-Average	Low	Small-Average
Average	Low	Good	High	Small-Average
Average	Low	Good	Average	Small-Average
Average	Low	Good	Low	Large
Average	Average	Poor-Average	High	Small-Average
Average	Average	Poor-Average	Average	Small-Average
Average	Average	Poor-Average	Low	Small-Average
Average	Average	Good	High	Small-Average
Average	Average	Good	Average	Small-Average
Average	Average	Good	Low	Small-Average
Average	High	Poor-Average	High	Small-Average
Average	High	Poor-Average	Average	Small-Average
Average	High	Poor-Average	Low	Large
Average	High	Good	High	Large
Average	High	Good	Average	Large
Average	High	Good	Low	Large
Large	Low	Poor-Average	High	Small-Average
Large	Low	Poor-Average	Average	Small-Average
Large	Low	Poor-Average	Low	Large
Large	Low	Good	High	Large
Large	Low	Good	Average	Large
Large	Low	Good	Low	Large
Large	Average	Poor-Average	High	Small-Average
Large	Average	Poor-Average	Average	Small-Average
Large	Average	Poor-Average	Low	Large
Large	Average	Good	High	Large
Large	Average	Good	Average	Large
Large	Average	Good	Low	Large
Large	High	Poor-Average	High	Large
Large	High	Poor-Average	Average	Large
Large	High	Poor-Average	Low	Large
Large	High	Good	High	Large
Large	High	Good	Average	Large
Large	High	Good	Low	Large

Model 17

Input 1	Input 11	Output 2.1
Small	Poor-Average	Large
Small	Good	Small-Average
Average	Poor-Average	Large
Average	Good	Small-Average
Large	Poor-Average	Small-Average
Large	Good	Small-Average

Model 18

Input 3	Input 10	Output 2.2
Poor	Poor	Large
Poor	Average	Large
Poor	Good	Average
Average	Poor	Large
Average	Average	Average
Average	Good	Small
Good	Poor	Average
Good	Average	Small
Good	Good	Small

Model 19

Input 4	Input 13	Output 3.1
Small-Average	Low	Small
Small-Average	Average	Small
Small-Average	High	Average-Large
Large	Low	Average-Large
Large	Average	Average-Large
Large	High	Average-Large

Model 20

Input 4	Input 12	Output 3.2
Small-Average	Low	Small
Small-Average	Average	Small
Small-Average	High	Average-Large
Large	Low	Average-Large
Large	Average	Average-Large
Large	High	Average-Large

Model 21

Input 1	Output 3.5
Small	Small-Average
Average	Small-Average
Large	Large

Appendix 10: LOM Testing Results

model 1	Input 1	output value	crisp value	error %	match (y/n)	match %	actual term	defuzzified term	match (y/n)	match %
	test 1	n/a		n/a			n/a			
	test 2	5	5.8	16	y		average	average	y	
	test 3	5	5.8	16	y		small-average	average	y	
	test 4	n/a		n/a			n/a			
	test 5	3	3	0	y		small	small	y	
	test 6	n/a		n/a			n/a			
	test 7	5	3	40	n	75	average	small	n	75
model 2	input 2	output value	crisp value	error %	match (y/n)	match %	actual term	defuzzified term	match (y/n)	match %
	test 1	n/a		n/a			n/a			
	test 2	8	4	50	n		high	low	n	
	test 3	7	4	42.857	n		average	low	n	
	test 4	n/a		n/a			n/a			
	test 5	8	10	25	y		high	high	y	
	test 6	n/a		n/a			n/a			
	test 7	8	10	25	y	50	average	high	n	20
model 3	input 3	output value	crisp value	error %	match (y/n)	match %	actual term	defuzzified term	match (y/n)	match %
	test 1	8	10	25	y		average-good	good	y	
	test 2	7.5	5	33.333	y		good	average	n	
	test 3	3	10	233.33	n		poor	good	n	
	test 4	8.5	10	17.647	y		good	good	y	
	test 5	9	10	11.111	y		good	good	y	
	test 6	7	10	42.857	n		average-good	good	y	
	test 7	4	10	150	n	57.1429	poor	good	n	57.1429

model 4	input 4	output value	crisp value	error %	match (y/n)	match %	actual term	defuzzified term	match (y/n)	match %
	test 1	5	4	20	y		average	small-ave	y	
	test 2	6.5	10	53.846	n		average	large	n	
	test 3	9	10	11.111	y		large	large	y	
	test 4	7	10	42.857	n		average-large	large	y	
	test 5	6	4	33.333	y		average	small-ave	y	
	test 6	8	10	25	y		average-large	large	y	
	test 7	8	4	50	n	57.1429	large	small-ave	n	71.4286
model 5	input5	output value	crisp value	error %	match (y/n)	match %	actual term	defuzzified term	match (y/n)	match %
	test 1	6	6.9	15	y		average	average	y	
	test 2	7.5	10	33.333	y		average-large	large	y	
	test 3	8	10	25	y		average	large	n	
	test 4	8	5	37.5	n		average-large	average	y	
	test 5	9	10	11.111	y		large	large	y	
	test 6	8	10	25	y		average-large	large	y	
	test 7	8	10	25	y	85.7143	average	large	y	85.7143
model 6	input6	output value	crisp value	error %	match (y/n)	match %	actual term	defuzzified term	match (y/n)	match %
	test 1	8	3	62.5	n		average	average	y	
	test 2	7.5	3.3	56	n		average-high	average	y	
	test 3	2	3	50	n		low	average	n	
	test 4	7	3.6	48.571	n		average	average	y	
	test 5	6	3	50	n		average	average	y	
	test 6	3	5.6	86.667	n		low-average	high	n	
	test 7	4	4.6	15	y	14.2857	low	average	n	57.1429

model 7	input8	output value	crisp value	error %	match (y/n)	match %	actual term	defuzzified term	match (y/n)	match %
	test 1	8	7.4	7.5	y		average-high	high	y	
	test 2	3.5	5	42.857	n		low	average	n	
	test 3	9	10	11.111	y		average-high	high	y	
	test 4	8	10	25	y		average-high	high	y	
	test 5	5	5	0	y		average	average	y	
	test 6	8	5	37.5	n		average-high	average	y	
	test 7	7	10	42.857	n	57.1429	average	high	n	71.43
model 8	input9	output value	crisp value	error %	match (y/n)	match %	actual term	defuzzified term	match (y/n)	match %
	test 1	n/a	n/a	n/a			n/a			
	test 2	3.5	5.5	57.143	n		low-average	average	y	
	test 3	9	10	11.111	y		high	high	y	
	test 4	n/a	n/a	n/a			n/a			
	test 5	3	5.5	83.333	n		low	average	n	
	test 6	6	5.5	8.3333	y		average	average	y	
	test 7	7	5.5	21.429	y	60	average	average	y	80
model 9	input10	output value	crisp value	error %	match (y/n)	match %	actual term	defuzzified term	match (y/n)	match %
	test 1	8	10	25	y		good	good	y	
	test 2	6.5	1	84.615	n		average-good	poor	n	
	test 3	4	10	150	n		poor-average	good	n	
	test 4	6	7	16.667	y		average	average	y	
	test 5	10	7.1	29	y		good	average	n	
	test 6	8	7.9	1.25	y		good	good	y	
	test 7	3	10	233.33	n	57.1429	average	good	n	42.8571

model 10 input 12	output value	crisp value	error %	match (y/n)	match %	actual term	defuzzified term	match (y/n)	match %
test 1	1	5	400	n		average	average	y	
test 2	7.5	5.5	26.667	y		average-high	average	y	
test 3	6	6	0	y		average	average-high	y	
test 4	5	5.3	6	y		average	average	y	
test 5	2	6	200	n		low	average-high	n	
test 6	7	5.9	15.714	y		average-high	average	y	
test 7	5	6	20	y	71.4286	average	average-high	y	85.7143
model 11 output 1.1	output value	crisp value	error %	match (y/n)	match %	actual term	defuzzified term	match (y/n)	match %
test 1	5	25	400	n		average	small-ave-large	y	
test 2	5	27	440	n		average	small-ave-large	y	
test 3	25	100	300	n		large	ave-large	y	
test 4	30	29	3.3333	y		large	small-ave-large	y	
test 5	25	33	32	y		average	ave-large	y	
test 6	12	100	733.33	n		average	ave-large	y	
test 7	20	100	400	n	28.5714	large	ave-large	y	100
model 12 output 1.2	output value	crisp value	error %	match (y/n)	match %	actual term	defuzzified term	match (y/n)	match %
test 1	n/a	n/a	n/a			n/a			
test 2	98	100	2.0408	y		large	ave-large	y	
test 3	95	13	86.316	n		large	ave-large	y	
test 4	n/a	n/a	n/a			large			
test 5	90	100	11.111	y		large	ave-large	y	
test 6	15	100	566.67	n		small-average	ave-large	y	
test 7	25	100	300	n	40	large	ave-large	y	100

model 13 output 1.3	output value	crisp value	error %	match (y/n)	match %	actual term	defuzzified term	match (y/n)	match %
test 1	5	100	1900	n		average	ave-large	y	
test 2	2	11	450	n		small	small-ave-large	y	
test 3	5	100	1900	n		average	ave-large	y	
test 4	5	10	100	n		average	small-ave-large	y	
test 5	10	100	900	n		small	ave-large	n	
test 6	1	10	900	n		small	small	y	
test 7	15	100	566.67	n	0	average	ave-large	y	85.7143
model 14 output 1.4	output value	crisp value	error %	match (y/n)	match %	actual term	defuzzified term	match (y/n)	match %
test 1	5	10	100	n		average	small-ave	y	
test 2	n/a	11				n/a	n/a		
test 3	60	100	66.667	n		large	large	y	
test 4	30	100	233.33	n		n/a	large		
test 5	25	13	48	n		n/a	large		
test 6	25	10	60	n		large	small-ave	n	
test 7	20	100	400	n	0	large	large	y	75
model 15 output 1.5	output value	crisp value	error %	match (y/n)	match %	actual term	defuzzified term	match (y/n)	match %
test 1	0	5		y		n/a	small-ave		
test 2	98	100	2.0408	y		average	large	n	
test 3	95	100	5.2632	y		large	large	y	
test 4	25	5	80	n		large	small-ave	n	
test 5	25	5	80	n		small	small-ave	y	
test 6	20	100	400	n		average	large	n	
test 7	0	5		y	57.1429	small	small-ave	y	50

model 16 output 1.6	output value	crisp value	error %	match (y/n)	match %	actual term	defuzzified term	match (y/n)	match %
test 1	10	100	900	n		average	large	n	
test 2	5	20	300	n		small-average	small-ave	y	
test 3	11	100	809.09	n		average	large	n	
test 4	1	16	1500	n		small	small-ave	y	
test 5	25	100	300	n		average	large	n	
test 6	5.5	24	336.36	n		small	small-ave	y	
test 7	12	100	733.33	n	0	average	large	n	42.8571
model 17 output 2.1	output value	crisp value	error %	match (y/n)	match %	actual term	defuzzified term	match (y/n)	match %
test 1	n/a	n/a	n/a			average			
test 2	5	100	1900	n		small-average	large	n	
test 3	10	25	150	n		average	small-ave	y	
test 4	n/a	100	n/a			n/a			
test 5	10	25	150	n		small	small-ave	y	
test 6	0	100				average	large	n	
test 7	20	25	25	y	20	average	small-ave	y	60
model 18 output 2.2	output value	crisp value	error %	match (y/n)	match %	actual term	defuzzified term	match (y/n)	match %
test 1	0	10		y		n/a	average		
test 2	5	3	40	y		small	small	y	
test 3	30	100	233.33	n		large	large	y	
test 4	0	5		y		small	small	y	
test 5	0	0	0	y		small	small	y	
test 6	0	10		y		small	small-ave	y	
test 7	20	100	400	n	71.4286	average	large	n	83.3333

model 19 output3.1	output value	crisp value	error %	match (y/n)	match %	actual term	defuzzified term	match (y/n)	match %
test 1	20	100	400	n		Average	large	n	
test 2	80	12	85	n		large	small	n	
test 3	40	100	150	n		large	large	y	
test 4	10	100	900	n		Average	large	n	
test 5	7	13	85.714	n		Average	small	n	
test 6	22.5	100	344.44	n		Average	large	n	
test 7	5	100	1900	n	0	small	large	n	14.2857
model 20 output 3.2	output value	crisp value	error %	match (y/n)	match %	actual term	defuzzified term	match (y/n)	match %
test 1	10	15	50	n		average	small	n	
test 2	n/a	100				average	ave-large	y	
test 3	40	100	150	n		large	ave-large	y	
test 4	50	100	100	n		large	ave-large	y	
test 5	n/a	20				n/a	small		
test 6	40	100	150	n		average	ave-large	y	
test 7	2	100	4900	n	0	small	ave-large	n	66.67
model 21 output 3.5	output value	crisp value	error %	match (y/n)	match %	actual term	defuzzified term	match (y/n)	match %
test 1	n/a	n/a	n/a			small			
test 2	5	5	0	y		small	small-ave	y	
test 3	20	5	75	n		large	small-ave	n	
test 4	0	5		y		small	small-ave	y	
test 5	n/a	5	n/a			n/a	small-ave		
test 6	7.5	5	33.333	y		small	small-ave	y	
test 7	1	5	400	n	60	small	small-ave	y	80
				count >= 50 ratio	12 0.57143			count>=50 ratio	17 0.80952

Appendix 11: Linguistic Term Analysis Results for Base Case

predicted \ actual

model 1

	small	average	large	(actual)	(predicted)
	small	average	large	1 case	
samll	2	1		small-ave	small
average		1			
large					

model 2

	low	average	high
low		1	
average			1
high		1	1

model 3

	poor	average	good		
				two cases	
poor				ave-good	good
average			1		
good	2		4		

model 4

	small	average	large		
				two cases	
samll-ave		4	2	ave-large	small-ave
large		1			

model 5

	small	average	large		
				3 cases	
samll				ave-large	average
average		6	1		
large					

model 6

	low	average	high		
				2 cases	
low				ave-high	average
average	2	5		low-ave	average
high					

model 7

	low	average	high		
				3 cases	
low				ave-high	high
average	1	2		1 case	
high		1	3	ave-high	average

model 8

	low	average	high		
				I case	
low				low-ave	average
average	1	3			
high			1		

model 9

	poor	average	good		
				I case	
poor			1	ave-good	poor
average		3	3	I case	
good				poor-ave	average

model 10

	low	average	high
low			
average	1	6	
high			

2 cases
ave-high average

model 11

	small	average	large
samll			
ave-large		4	3

2 cases
average small-ave-large
1 case
large small-ave-large

model 12

	small	average	large
samll			
ave-large		1	4

1 case
small-ave ave-large

model 13

	small	average	large
samll	1		
ave-large	2	4	

1 case
small small-ave-large

model 14

	small	average	large
small-ave			
large		1	3

model 15

	small	average	large
small-ave	2		2
large		2	

model 16

	small	average	large
small-ave	1		
large	2	4	

1 case
small-ave large

model 17

	small	average	large
small-ave	1	2	
large	1	1	

1case
small-ave large

model 18

	small	average	large
samll	1		
average	2		
large	1	1	1

model 19

	small	average	large
samll			
ave-large	1	4	2

model 20

	small	average	large
samll			
ave-large	1	3	2

model 21

	small	average	large
small-ave	2		
large	2		1

Appendix 11: Linguistic Term Analysis Results for LOM Method

(actual) (predicted)

model 1

	small	average	large
small	1	1	
average		2	
large			

1 case
small-ave average

model 2

	low	average	high
low		1	1
average			
high		1	1

model 3

	poor	average	good
poor			
average			1
good	2		4

two cases
ave-good good

model 4

	small	average	large
small-ave		3	1
large		1	2

one case
ave-large small-ave
one case
ave-large large

model 5

	small	average	large
small			
average		2	
large		2	3

1 case
ave-large average
2 cases
ave-large large

model 6

	low	average	high
low			
average	2	4	
high	1		

2 cases
ave-high average
low-ave high

model 7

	low	average	high
low			
average	1	2	
high		1	3

3 cases
ave-high high
1 case
ave-high average

model 8

	low	average	high
low			
average	1	3	
high			1

I case
low-ave average

model 9

	poor	average	good
poor			1
average		1	1
good	1	1	2

I case
ave-good poor
I case
poor-ave good

model 10

	low	average	high		
low				2 cases	
average	1	6		ave-high	average
high					

model 11

	small	average	large	2 cases	
small				average	small-ave-large
				1 case	
ave-large		4	3	large	small-ave-large

model 12

	small	average	large	1 case	
small				small-ave	ave-large
ave-large		1	4		

model 13

	small	average	large	1 case	
small	2			small	small-ave-large
				1 case	
ave-large	1	4		average	small-ave-large

model 14

	small	average	large
small-ave		1	1
large			2

model 15

	small	average	large
small-ave	2		1
large		2	1

model 16

	small	average	large	1 case	
small-ave	3			small-ave	small-ave
large		4			

model 17

	small	average	large	1case	
small-ave	1	2		small-ave	large
large	1	1			

model 18

	small	average	large
small	4		
average			
large		1	1

model 19

	small	average	large
small		1	1
ave-large	1	3	1

model 20

	small	average	large
small		1	
ave-large	1	2	2

model 21

	small	average	large
small-ave	4		1
large			