INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

UMI®

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

University of Alberta

A Framework for Integrating Fuzzy Set Theory and Discrete Event Simulation in Construction Engineering

by

Ahmed Abdel-Hameed Ibrahim Shaheen

©

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of **Doctor of Philosophy**

in

Construction Engineering and Management

Department of Civil and Environmental Engineering

Edmonton, Alberta

Fall 2005

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Library and Archives Canada

Bibliothèque et Archives Canada

Published Heritage Branch

Direction du Patrimoine de l'édition

395 Wellington Street
Ottawa ON K1A 0N4
Canada

395, rue Wellington
Ottawa ON K1A 0N4
Canada

0-494-08730-7

Your file Votre référence
ISBN:
Our file Notre reterence
ISBN:

NOTICE:
The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:
L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

ABSTRACT

Discrete event simulation is a powerful modeling tool that has been utilized to model numerous construction engineering operations. The role of experts is very important in designing and defining many of the simulation model parameters. Current simulation modeling practices do not have a methodology for integrating experts' knowledge and opinion. This thesis presents a methodology to incorporate experts' knowledge and opinion within the simulation framework in construction engineering applications, in order to enhance discrete event simulation modeling capabilities. The concepts of fuzzy set theory are adopted to incorporate the experts' knowledge within the simulation framework, because fuzzy set theory is capable of modeling experts' way of thinking and can easily capture their decision-making processes.

The components of fuzzy modeling framework is proposed and integrated within the simulation modeling framework. Three main applications of the fuzzy modeling framework are identified. The first application is utilizing fuzzy numbers in modeling cost range estimating as compared to probabilistic range estimating. The second application is utilizing fuzzy expert system tools in predicting activity behavior within the simulation framework, using the tunnel boring machine (TBM) penetration rate prediction as a case study. The third application is utilizing fuzzy expert systems in the decision-making process within the simulation framework, using the prioritization of modules awaiting assembly in a module assembly yard as a case study.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The integrated fuzzy modeling and discrete event simulation framework has proven to be very promising in enhancing the modeling capabilities of discrete event simulation. Integrated fuzzy and discrete event simulation modeling is capable of explicitly and more confidently predicting the behavior of an activity within the simulation framework and incorporating the experts' decisions while simultaneously accounting for the uncertainty embedded within the decision making process.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

ACKNOWLEDGEMENTS

I would like to express my sincere thanks to those who guided and encouraged me in the development of this thesis. Special thanks are due to my supervisor Dr. Aminah Robinson, Dr. Simaan AbouRizk, Dr. Michael Allouche, and Dr. Yasser Mohammed.

I would also like to acknowledge with much appreciation the role of my friend and colleague Mohammed Al-Bataineh who provided me with important programming support.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE OF CONTENTS

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

LIST OF TABLES

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

LIST OF FIGURES

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

LIST OF ABBREVIATIONS

AG(A)	Ambiguity measure of fuzzy number (A).
AHP	Analytic Hierarchy Process.
$(\hat{A_m})_i$	Expected mean value of fuzzy number (i).
$(\hat{A_m})_{Faverage}$	Expected mean value of the average fuzzy number ($F_{average}$).
Avg.	Average.
Bsum	Bound sum fuzzy result aggregation.
C_n	n^{th} decision criterion.
CE	Combined effect of relative importance and impact of factors.
COA	Center of Area.
CON_n	Contribution of n^{th} decision criterion.
COM	Center of Maximum.
CR	Critical ratio.
$E(X)$	Mean value of a continuous random variable X.
$E_*(A)$	Lower expected value of a fuzzy number (A).
$E^*(A)$	Upper expected value of a fuzzy number (A).
EI(A)	Expected interval of fuzzy number (A).
EV(A)	Expected value fuzzy number (A).
EDD	Earliest due date.
F(A)	Fuzziness measure of fuzzy number (A).
$F_{average}$	Average fuzzy number.
$f_X(x)$	Probability density function.
FCFS	First come, first served.
FNQI	Fuzzy Number Quality Index.
I_y	Moment of inertia about the vertical centroidal axis.
Max.	Maximum.
MAX_{range}	Maximum range.
Min.	Minimum.
MIN_{range}	Minimum range.
MMC	Maximum module cost.
MOPNR	Most operations remaining.
MWKR	Most work remaining.
Prod.	Product fuzzy aggregation operator.
R_n	n^{th} fuzzy rule combination.
SOM	Simphony Object Model
SPT	Shortest processing time.
$Var(X)$	Variance of a continuous random variable X.
x_0	Centroidal distance of a unit area.
y^*	Defuzzified value of fuzzy number.
$\mu_i(x)$	Aggregated membership function of output variable(x).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

CHAPTER 1-INTRODUCTION

1.1 Uncertainty in construction engineering

Due to the increasing complexity of construction projects, no project can succeed by accident. Different factors affect the success or failure of any project in construction engineering. Following a good planning and control system is one of the factors that can lead to a successful project. Providing accurate quantity, cost, and time estimates can also contribute to the project success. In addition, being able to successfully predict or measure some factors such as productivity and risk can certainly lead to better project performance. However, the aforementioned factors are not easy to achieve or accomplish due to unforeseen parameters such as weather changes, breakdown of equipment, labour inefficiency, and delayed delivery of resources (Zhang et al. 2003). Consequently, these parameters may result in a great deal of uncertainty that can affect the overall construction process and may lead to its failure.

Yao and Furuta (1986) associate "uncertainty" with "ambiguity, fuzziness, randomness, vagueness, and imprecision of events under consideration". Uncertainty associated with randomness is caused by complex phenomena that are random in nature. In addition, all of the other uncertainty terms can be associated with subjectivity and imprecise knowledge describing the events which result from the use of natural language that is not clearly defined. Therefore, the uncertainty terms can be grouped into two main categories: randomness and fuzziness.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Several models and systems have been developed in order to deal with the high level of uncertainty involved within construction engineering projects. Many artificial intelligence models, such as simulation, neural networks and fuzzy logic are designed to deal with uncertainty. Simulation techniques are designed to model real-world systems. Stochastic simulation based on statistical data is more suitable for modeling construction related projects which are very susceptible to variation and interruptions (AbouRizk and Halpin 1990). Simulation has been successfully used to model many construction related problems. Some of these models are simulation analysis of construction site dewatering by Hajjar et al. (1998), simulation modeling of tunneling operations by Mohamed and AbouRizk (2001), simulation of concrete batch plant production by Zayed and Halpin (2001), and simulation modeling of earth moving operations by Marzouk and Moselhi (2003).

Artificial neural networks are capable of predicting a construction related behavior based on historical data. The power of artificial neural networks lies in their ability to learn from the different data sets used, which helps to yield more effective and reliable models. Neural networks have been used in many construction related problems such as the neural network system developed for the selection of horizontal formwork by Hanna and Senouci (1993), estimating construction formwork labour productivity by AbouRizk and Portas (1997), neural network modeling for pavement maintenance decision making system by Alsugair et al. (1998), and predicting highway construction costs by Wilmot and Mei (2005).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Fuzzy sets and fuzzy logic techniques are capable of dealing with the linguistic terms used in modeling construction related problems, and they can effectively account for the uncertainty associated with numerous construction problems. Many researchers utilized fuzzy modeling concepts to model different construction related problems such as tender evaluation using fuzzy set by Nguyen (1985), utilizing fuzzy set concepts for project scheduling and network analysis by Ayyub and Haldar (1984) and Lorterapong and Moselhi (1996), fuzzy decision making system for contractor selection by Singh and Tiong (2005) and predicting industrial construction labour productivity using fuzzy expert systems by Fayek and Oduba (2005)

The aforementioned artificial intelligence modeling tools have been increasingly utilized to model construction engineering problems. Integrating these models can yield more powerful systems. This thesis proposes possible methods of integrated fuzzy set theory modeling techniques with discrete event simulation.

1.2 Problem Statement

Construction engineering operations are characterized as being cyclic, repetitive, and resource intensive in nature (Maio et al. 2000). Hence, simulation models have proven to be effective and successful modeling tools for construction operations. AbouRizk and Halpin (1992) stated that modeling the random simulation inputs is considered the key factor behind the success of the simulation construction process. They identified two possible scenarios when modeling construction engineering inputs. In the first scenario, the modeling data are either

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

observed or provided from historical data. The second scenario represents the data-deficient environment. In the latter scenario, AbouRizk and Halpin (1992) suggest that the modeler must rely on his or her personal subjective judgment to select the appropriate probabilistic distributions that best describe the simulated construction process.

In simulation, historical data are used to fit a probabilistic distribution, from which the input variable (duration) is randomly sampled. In addition it is known that every project is unique. Therefore, conditions such as weather, labour experience, interruptions, and equipment conditions should be incorporated in the calculation of the activity duration for each project, since they may highly impact and change the historical statistical distributions used in modeling the project. AbouRizk and Sawhney (1992) showed that the uncertainties encountered in construction are usually handled by two approaches in simulation. In the first approach, aggregated input-process, the activity duration is represented by one statistical distribution that incorporates all the elements of uncertainty in it. The conditions affecting the activity duration are modeled implicitly in this approach. In the second approach, separate input-process, the influence on duration of the uncertainty elements is modeled separately. In addition, the different conditions affecting the activity duration are explicitly modeled within the separate input-process approach. In both approaches, the effect of the uncertainty elements on simulation inputs has to be addressed and modeled.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Probability based modeling techniques, such as simulation, are very effective in modeling uncertainty only when enough data sets that describe the uncertainty are available. As a result, subjectively selecting a probability distribution that approximates and best represents the missing data is not optimal, since it is not usually a straightforward and accurate process and can eventually affect the simulation output. The difficulty in approximating a probability distribution is that experts do not think in probability values, but rather they think in linguistic terms such as much, very, high, etc. (Kim and Fishwick 1997).

In conclusion, the following points record some of the limitations identified after studying simulation modeling in construction engineering:

1) Simulation modeling is based on the availability of historical data that describe the problem being modeled. Problems arise when data are not available or do not sufficiently represent the problem being modeled. Some current practices use judgment and subjectivity based on experience to assume alternative statistical distributions.
2) Construction projects are unique. Many factors have been proven to adversely affect project performance and hence, can affect the modeling process. However, these factors are usually overlooked when dealing with stochastic models and they are usually modeled implicitly.
3) In stochastic models, uncertainty is modeled based only on the frequency of occurrence of events. This means that modeling new projects is mainly based on the experience with past ones. However, the specific conditions, that every project has, require that they be accounted for in order to be

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

more representative for the changing parameters. Current statistical methods do not account for these changing conditions in every project since they only deal with past experience.

4) There is not any modeling methodology that comprehensively and effectively models uncertainty associated with both randomness and fuzziness.
5) The incorporation of experts' knowledge and information in the discrete event simulation modeling is subjectively handled.

Therefore, in light of all the limitations and problems listed, there is a need to incorporate the experts' knowledge and opinion within the simulation framework in order to enhance the modeling capabilities of discrete event simulation.

1.3 Research Objectives

Based on the problems discussed in Section 1.2, the research has the following objectives:

1) To identify the components of a framework for integrating the expert thinking using fuzzy set theory and discrete event simulation.
2) To develop a hybrid input modeling technique that is capable of modeling uncertainty caused by randomness and fuzziness in the input variables in discrete event simulation.
3) To explicitly incorporate and account for the factors affecting the simulation inputs into the simulation modeling process.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

4) To provide the modeler with different input modeling techniques to choose from based on the amount of information available and his or her preference.
5) To model the expert thinking and experience using fuzzy set theory for better modeling capability of discrete event simulation.
6) To extend the applicability and state of the art in fuzzy set research in construction by integrating it with simulation modeling techniques.

1.4 Research Summary

In order to achieve the research objectives discussed in Section 1.3, the thesis will try to combine different modeling techniques so as to create more comprehensive integrated fuzzy-simulation models in order to best model uncertainty in simulation. The integration is illustrated using different applications supported by case studies. Three main applications are illustrated:

1) Using fuzzy numbers in range estimating, supported by a cost range estimating case study.
2) Using a fuzzy expert system as a predictive tool to model the activity input in discrete event simulation models. A tunneling operation is used as a case study.
3) Using the fuzzy if-then rules for decision making in discrete event simulation models. The prioritization of modules in module assembly yards is used as a case study.

The thesis objectives will be achieved through the following steps:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

1) Showing the capabilities of using fuzzy numbers in the cost range estimating process by comparing its capabilities with Monte Carlo simulation range estimating.
2) Using fuzzy expert system modeling techniques to model the activity inputs using the Tunnel Boring Machine (TBM) advance rates as a case study and proposing a methodology for integrating the fuzzy expert system with discrete event simulation using Simphony.
3) Using fuzzy if-then rules as a decision making tool in module assembly yard scheduling. The fuzzy decision making tool will be utilized in prioritizing the modules awaiting assembly in assembly yard using Simphony©.
4) Developing a fuzzy modeling framework for integrating the fuzzy predictive modeling tools and the fuzzy decision making tools within discrete event simulation framework.

In conclusion, the proposed integrated framework tries to provide more effective models by integrating and incorporating the concepts of fuzzy set theory into simulation. More powerful and enhanced simulation models can be generated when integrated with fuzzy set theory concepts, which is what this thesis will attempt to achieve.

1.5 Thesis Organization

Chapter 1 provides an introduction and background for the research. It explains the problem statement and the overall thesis objectives.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Chapter 2 provides a methodology for utilizing fuzzy numbers in the cost range estimating process.

Chapter 3 introduces fuzzy expert systems as a predictive tool and discusses the proposed methodology for generating the rules' consequences. It also shows how fuzzy expert systems can be used to model the tunnel boring machine penetration rates in soft ground soils.

Chapter 4 discusses a methodology proposed for integrating fuzzy expert systems with discrete event simulation and illustrates the tunneling case study in which the fuzzy expert system predictive model is used to predict TBM penetration rate in a tunneling simulation model.

Chapter 5 introduces a methodology to use the fuzzy if-then rules as a decision making tool. The prioritization of modules awaiting assembly in a module assembly yard is used as a case study.

Chapter 6 illustrates the components and steps for the design and development of the fuzzy modeling framework which is integrated within the simulation framework.

Chapter 7 discusses the conclusions and comments drawn from the thesis, thesis contributions, and the future developments.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

CHAPTER 2-FUZZY NUMBERS IN COST RANGE ESTIMATING

2.1 Introduction

Range estimating is one of the most used forms of Monte Carlo Simulation in construction practice. This technique is a simple form of simulating a project estimate by breaking the project into work packages and approximating the variables in each package using statistical distributions. During simulation, these distributions are sampled and a total project cost is aggregated, and statistically analyzed to derive proper cost indicators with various probabilities of achieving them. Such an approach is useful in quantifying uncertainties with high risk work packages thus leading to better decisions regarding the project budget.

This chapter explores an alternate approach to range estimating that is grounded in fuzzy set theory. The approach addresses two shortcomings of Monte Carlo Simulation. The first is related to the analytical difficulty associated with fitting statistical distributions to subjective data, and the second relates to the required number of simulation runs to establish a meaningful estimate of a given parameter at the end of simulation.

Fuzzy set theory enables us to subjectively elicit information about parameters of interest in each work package from an estimator without having to sacrifice accuracy, as such elicitation is a cornerstone of fuzzy set theory (as opposed to approximating a statistical distribution from subjective information).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Fuzzy set theory requires only one path of calculations to establish an estimate of the parameter of concern as opposed to the multiple runs required in Monte Carlo simulation. The chapter shows that we can achieve comparable results using this approach in a more efficient manner.

2.2 Modeling Range Estimating Using Monte Carlo Simulation

Before the emergence of fuzzy set theory as a good tool for modeling uncertainty, probability theory was the only well developed mathematical tool for dealing with uncertainty (Klir and Folger 1988). One of the probabilistic methods that models uncertainty in an estimate (cost or duration) is Monte Carlo simulation. This risk analysis process is called range estimating.

Ahuja et al. (1994) define range estimating as a simulation modeling process performed after an estimate is made (i.e. estimate of duration or cost) to reflect the degree of uncertainty associated with an estimate. Taking cost range estimating as the main example, the process can be summarized:

1) Identify the major project or work components in the form of major cost packages and their related subcategories, which can be restricted to the major items that affect the total cost bottom line by a certain percentage (i.e. > 1 %).
2) Identify the uncertain items.
3) Use statistical distributions (i.e. triangular and/or uniform distributions) to model the variability of each uncertain item.
4) Use Monte Carlo simulation to provide the final outputs.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

5) Collect statistics on the mean, standard deviation and minimum/maximum values of the output.

The Monte Carlo simulation technique has been used to model the risk analysis of construction operations. Using this technique, the modeled system takes inputs in the form of random variables. The process continues by performing experiments with many variations of the input and then collects sets of outputs in the form of statistical distributions, which are analyzed to provide the measure of uncertainty and risk. The main steps followed in Monte Carlo simulation are (Ahuja et al. 1994):

1) Generate random numbers.
2) Generate random variates reflecting the true nature of the modeled item (i.e. duration and cost estimate). This step is referred to as "input modeling" which requires the modeled item be modeled by an appropriate probabilistic distribution that best represents the item.
3) Run the model and calculate the desired output parameters.
4) Steps 1 to 3 are repeated for a large number of iterations.
5) Terminate after a specified number of iterations and analyze the collected output statistics.

Range estimating can be an effective tool for modeling the uncertainty embedded within cost or duration estimates. However, this technique is based entirely on probabilistic and statistical modeling techniques that model the randomness of the problem. In addition, as explained before, the process requires a large number of

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

iterations in order to reach a reliable output. Selecting the appropriate statistical distribution that best models the inputs is an important issue in random modeling.

2.3 Modeling Uncertainty Using Fuzzy Numbers

2.3.1 Introduction to fuzzy numbers

Fuzzy set theory deals with the set of objects with a continuum of grades of membership characterized by a membership (characteristic) function that assigns to each object a grade of membership ranging between zero (no membership) and one (full membership) (Zadeh 1965). The concept of "unsharp boundaries" that fuzzy set theory tries to represent mimics the human way of thinking, which works with shades of gray rather than black and white. A fuzzy number is a fuzzy membership function that is both convex and normal. Fuzzy numbers are written in the form of a domain value and its corresponding confidence level. Figure 2.1 shows an example of a convex and a normal fuzzy number representing the concept of "about 4". A fuzzy number can be considered a generalization of the concept of interval of confidence. Therefore, the mathematical operations of fuzzy numbers (i.e. addition, subtraction, multiplication, and division) can be processed using the concepts of the interval of confidence.

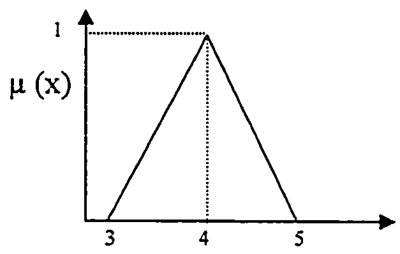


Figure 2.1: "About 4" Represented by a Triangular Fuzzy Number

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

2.3.2 Fuzzy numbers versus probability distributions

Fuzziness and probability represent different aspects of uncertainty. According to Kaufmann and Gupta (1985), a fuzzy number is not a random variable. The random variable is defined in terms of the theory of probability, whereas a fuzzy number is a subjective datum that is defined by the theory of fuzzy sets. In order to show the difference between the two types of uncertainty, the summation operation is taken as an example. Figure 2.2 shows that the summation of probability distributions will eventually achieve a Gaussian shape (normal) that obeys the Laplace-Gauss central limit theorem. On the other hand, the addition of two constant fuzzy numbers that are similar in shape will result in a larger constant fuzzy number that maintains the same shape. This example shows how the two approaches differ in the way they process uncertainty.

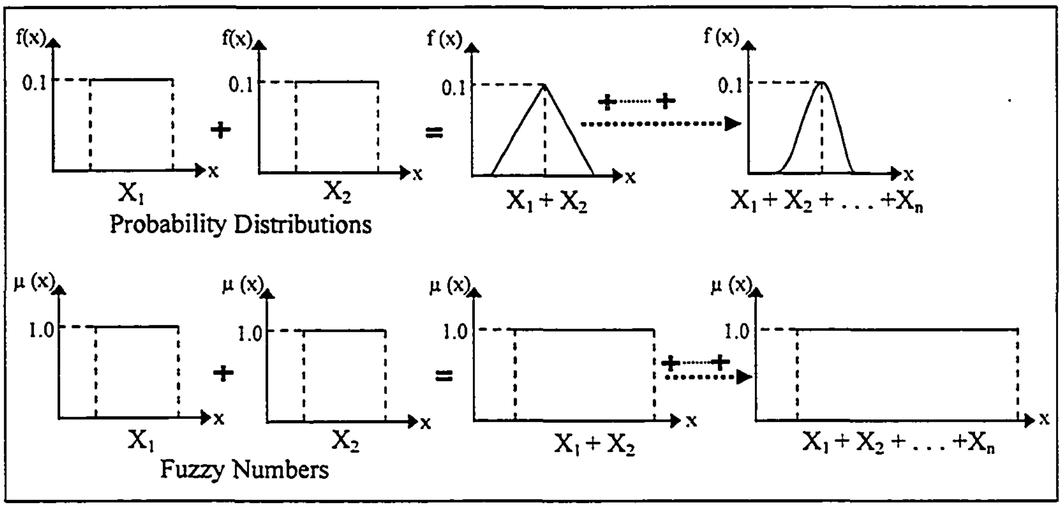


Figure 2.2: Comparison between Probability and Fuzzy Convolution

According to Ferson (2002), some of the disadvantages of Monte Carlo methods are computational burden, sensitivity to uncertainty about input distribution

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

shapes, and the need to assume correlations among all inputs. On the other hand, modeling uncertainty using fuzzy arithmetic is computationally simple, not very sensitive to moderate changes in the shapes of input distributions, and does not require the analyst to assume particular correlations among inputs. However, the results generated by fuzzy arithmetic are conservative and may overestimate uncertainty. As illustrated in Figure 2.2, the summation operation of two fuzzy numbers generated an output that is wider in range.

Fuzzy set theory can be used as an effective alternative to the random modeling of uncertainty. It is a very attractive alternative because it is more capable of extracting and representing the required information from experts by effectively capturing their linguistic and subjective evaluations. In addition, the calculations involved are much easier and faster compared to the probabilistic approach.

In order to minimize the effect of overestimation in fuzzy modeling, this chapter will introduce concepts that must be considered when modeling uncertainty using fuzzy arithmetic. The following sections present some of the concepts that need to be incorporated and how they will be used to model uncertainty in cost range estimating.

2.3.3 Crisp representation of a fuzzy number

It is also important to associate an ordinary or a crisp quantity with the fuzzy number. The crisp quantity will represent the "defuzzified" or "expected" or value of the fuzzy number. Calculating the "expected value" of the fuzzy number will render the fuzzy number ranking and comparison much easier.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Different methodologies have been developed to capture an "expected value" of a fuzzy number. Heilpern (1992) introduced the notions of expected interval EI(A) and expected value EV(A) of a fuzzy number. According to Heilpern (1992) the expected interval is defined as the expected value of an interval random set generated by the fuzzy number, and the expected value of this number is defined as the centre of the expected interval. In his approach Heilpern proved that the expected interval of a fuzzy number is equal to the mean value of this number.

Heilpern defined a lower and upper expected value of a fuzzy number as follows:

The expected interval EI(A) equals:

[2.1] $EI(A) = [E_*(A), E^*(A)]$

Where $E_*(A)$ and $E^*(A)$ are the lower and upper expected value of a fuzzy number respectively.

And the expected value EV(A) equals:

[2.2] $EV(A) = \frac{1}{2}(E_*(A) + E^*(A))$.

For a trapezoidal fuzzy number (a_1, a_2, a_3, a_4),

[2.3] $EI(A) = [(a_1 + a_2)/2, (a_3 + a_4)/2]$

[2.4] $EV(A) = (a_1 + a_2 + a_3 + a_4)/4$

For more details on the mathematical background of equations 2.1 to 2.4, the reader can refer to Heilpern (1992).

One of the most common defuzzification methods is the Center of Area "COA", which is calculated as:

[2.5] $$y^* = \frac{\int x\mu_i(x_i)}{\int \mu_i(x_i)}$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

where, y^* is the defuzzified value, $\mu_i(x)$ is the aggregated membership function and x is the output variable. Equation 2.5 represents the centroid of the fuzzy number.

In probability theory, the mean value and variance correspond, respectively, to the centroidal distance and central moment of inertia of an area (Ang and Tang 1975). The centroidal distance (x_0) of a unit area is calculated by:

$$[2.6] \quad x_0 = \frac{\int_{-\infty}^{\infty} x f(x)dx}{area} = \int_{-\infty}^{\infty} x f(x)dx$$

Refer to Figure 2.3 for the notations of the centroidal distance equation.

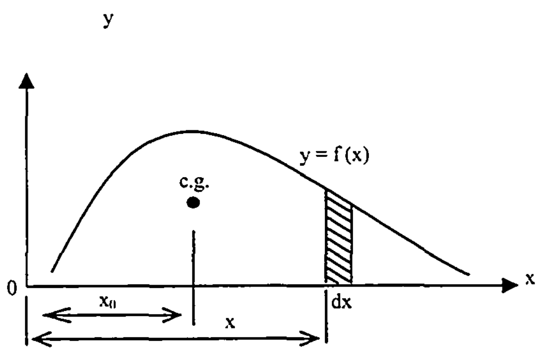


Figure 2.3: Centroid Location of an Irregular Area

Equation [2.6] is also the first moment (about 0) of the irregular-shaped area. The moment of inertia about the vertical centroidal axis (I_y) is:

$$[2.7] \quad I_y = \int_{-\infty}^{\infty} (x - x_0)^2 f(x)dx$$

In probability, the mean value of a continuous random variable X with a probability density function $f_X(x)$ is:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

[2.8] $E(X) = \int_{-\infty}^{\infty} x f_X(x) dx$

The variance of a continuous random variable X with a probability density function $f_X(x)$ is:

[2.9] $Var(X) = \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x) dx$

When comparing equations 2.6 and 2.7 with 2.8 and 2.9, respectively, it can be noticed that the mean value is equivalent to the centroidal distance, whereas the variance is equivalent to the centroidal moment of inertia of an area (Ang and Tang 1975).

Therefore, since the probabilistic mean value is equivalent to the centroidal distance of an area, the "COA" defuzzification method of any fuzzy number will generate a defuzzified value which is identical to the probabilistic mean of the normalized fuzzy number. It can be concluded that when a fuzzy number maintains the same range and the same shape of a bounded or a truncated probabilistic distribution, the defuzzified value of the fuzzy number using "COA" method will be equivalent to the mean value of the probabilistic distribution. The equations for calculating the expected values using the "COA" method of the common fuzzy numbers used in the study are as follows (the fuzzy numbers are represented by a four element notation (a, b, c, d) :

[2.10] Uniform fuzzy numbers (a, a, b, b), $EV_{uniform} = \dfrac{a+b}{2}$

[2.11] Triangle fuzzy numbers (a, b, b, c), $EV_{Triangle} = \dfrac{a+b+c}{3}$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

[2.12] Trapezoidal fuzzy numbers (a, b, c, d), $EV_{Trapezoidal}$ = $a+\dfrac{2ac+a^2+cb+ab+b^2}{3(a+b)}$

Equations 2.10, 2.11, and 2.12 calculate the expected values of the uniform, triangular, and trapezoidal fuzzy numbers respectively, which are equivalent to the mean values of uniform, triangular, and trapezoidal probability distributions respectively.

The variances of the common fuzzy numbers used in the study can be calculated using the probabilistic definition of variance as follows:

[2.13] Uniform fuzzy numbers (a, a, b, b), $Variance_{uniform} = \dfrac{(b-a)^2}{12}$

[2.14] Triangular fuzzy numbers (a, b, b, c), $Variance_{Triangular}$ = $\dfrac{a^2+b^2+c^2-ab-ac-bc}{18}$

[2.15] Trapezoidal fuzzy numbers (a, b, b, c), according to Dorp and Kortz (2003), $Variance_{Trapezoidal}$ =

$$\frac{(b-a)}{(d+c-b-a)}\left(\frac{1}{6}(a+b)^2+\frac{1}{3}b^2\right)+\frac{1}{(d+c-b-a)}\left(\frac{2}{3}(c^3-b^3)\right)+$$
$$\frac{(d-c)}{(d+c-b-a)}\left(\frac{1}{3}c^2+\frac{1}{6}(c+d)^2\right)$$

Equations 2.13, 2.14, and 2.15 calculate the variances of the uniform, triangular, and trapezoidal fuzzy numbers respectively, which are equivalent to the variance values of uniform, triangular, and trapezoidal probability distributions respectively.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Therefore, when a fuzzy number maintains the same range and the same shape of a bounded or a truncated probabilistic distribution, the variance of the fuzzy value will be equivalent to the variance of the related probability distribution.

2.3.4 Summation of Fuzzy Numbers

As indicated in Section 2.3.1, the fuzzy numbers are considered a generalization of the interval and the mathematical operations on fuzzy numbers (i.e. summation) can be processed using the concepts of the interval of confidence. Therefore, the methodology adopted for the summation of fuzzy numbers is the a-cut method and interval analysis. As shown in Figure 2.4, a-cut (a ∈ [0,1]) is a discretization technique applied on the continuous membership functions to generate a discrete set of variables in the form of intervals (a, b).

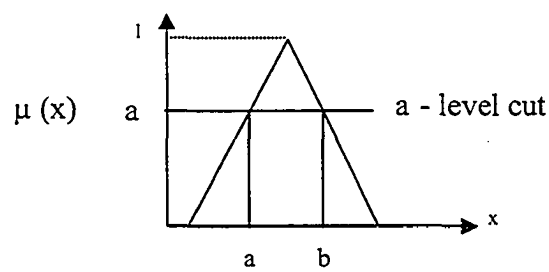


Figure 2.4: α-cut Operation on a Triangular Number

The a–cut technique is based on the "extension principle" which implies that algebraic operations on real numbers can be extended to fuzzy numbers (Zadeh 1975). Interval arithmetic is used to analyze the generated intervals.

The following example illustrates the a-cut calculation process of fuzzy number summation using interval arithmetic. Suppose we have two fuzzy numbers A =

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

(1, 4, 7) and B = (2, 5, 8, 10). We need to calculate Z = A + B using a-cut technique. The calculation process is carried out through the following steps:

1) Select a particular a-cut value ($0 \leq a \leq 1$)

2) Find the corresponding intervals of the selected a-cuts

3) Use the interval operations to calculate the intervals in the output Z

4) Repeat the steps for as many a-cuts as needed.

To calculate Z = A + B, four a-cuts are selected (0, 1/3, 2/3, and 1)

a = 0: [1,7] + [2,10] = [3,17]

a =1/3: [2,6] + [3, 9.34] = [5 ,15.34]

a =2/3: [3,5] + [4,8.67] = [7,13.67]

a = 1: [4,4] + [5,8] = [9,12]

These a-cut solutions represent the fuzzy output (Z) which is shown in Figure 2.5.

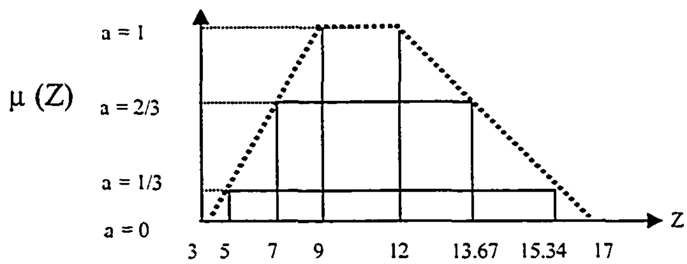


Figure 2.5: Solution Of (z = a + b) Using α-cuts

The a-cut method can be used in the different types and shapes of fuzzy numbers. The types of fuzzy numbers used in this study (uniform, triangle, trapezoidal) can be considered special cases of trapezoidal. Therefore, they can be represented by 4 variables; trapezoidal = (a,b,c,d), uniform = (a,a,b,b), and triangular = (a,b,b,d). The summation of these numbers can be performed by a-cut method or by direct summation since they are all represented by 4 variables.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

2.3.5 Fuzzy Summation and Probabilistic Central Limit Theorem

According to the central limit theorem, the sum of independent random variables tends to the normal distribution as the number of random variables, regardless of their distributions, increases without limit (Ang and Tang 1975).

Therefore, to prove that using fuzzy numbers in range estimating can yield comparable results to the probabilistic approach, we need to prove that the defuzzified expected value of the fuzzy output using "COA" is comparable to the mean of the probabilistic output, which is represented by Gaussian distribution based on the central limit theorem, and that the summation of variances of fuzzy inputs are comparable to the summation of variances of probabilistic distributions. Boswell and Taylor (1987) investigated the concept of fuzzy random variable which is a fuzzy set consisting of a membership function and a basic set whose components are ordinary mapping (real random variables) from a probability space. In addition, the random variables exhibit an infinite number of distributional types whose summation or average is a fuzzy random variable with a membership function of its own and a basic set of random variables. The conclusion of their study is that the summation of independent fuzzy random variables converges, in the limit, to a fuzzy Gaussian random variable providing a fuzzy equivalence of the central limit theorem of classical probability theory.

Boswell and Tylor's study (1987) provided a mathematical proof of the similarity between the summation fuzzy random variables and central limit theorem of classical probabilistic summation. In this study, we are only dealing with fuzzy

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

numbers which are not random. Therefore, the analysis is much easier and it is considered a special case of the summation fuzzy random variables.

In order to show how the summation of fuzzy numbers behaves, an experiment is conducted using Excel and @RISK add-in. The experiment is designed to compare the outputs of the summation of fuzzy inputs and probabilistic distributions. The experiment is designed using the following assumptions:

1) The two approaches will share the same boundaries. The boundaries are randomly created between the following ranges:
   a) Minimum range (MIN_{range}) = Random (0 to 1000).
   b) Maximum range (MAX_{range}) = Random ([Min+ 100] to [Min+ 1000]).
2) Fuzzy numbers used are:
   a) Uniform (a, b) : a = (MIN_{range}), b= (MAX_{range}).
   b) Triangular: (a, b, c): a = (MIN_{range}), c= (MAX_{range}), and (b) is randomly created between (a) and (c).
   c) Trapezoidal (a, b, c, d): a = (MIN_{range}), d= (MAX_{range}), (b) and (c) are randomly created between (MIN_{range}) and (MAX_{range}), provided that (b) ? (c).
3) Probability distributions used are:
   a) Uniform (min., max.): min (continuous boundary parameter) = (MIN_{range}), max (continuous boundary parameter) = (MAX_{range}).
   b) Triangle (min., most likely, max.): min (continuous boundary parameter) = (MIN_{range}), (max continuous boundary parameter) =

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

(MAX_{range}), most likely continuous mode parameter is randomly created between (MIN_{range}) and (MAX_{range}).

c) PERT (min., most likely, max.), which is an approximation of Beta distribution: min (continuous boundary parameter) = (MIN_{range}), (max continuous boundary parameter) = (MAX_{range}), most likely continuous parameter is randomly created between (MIN_{range}) and (MAX_{range}).

d) Generalized Beta (α_1, α_2, min., max.): min (continuous boundary parameter) = (MIN_{range}), max (continuous boundary parameter) = (MAX_{range}), and α_1 (continuous shape parameter) and α_2 (continuous shape parameter) are both randomly created between (2 and 25). The range of the shape parameters is selected between 2 and 25 to restrict the shape of the Beta distributions to unimodal distributions. In order to deal with bimodal Beta distribution, the distribution must be divided into two distributions (each with one local maximum). Two fuzzy numbers are created in order to match the parameters of the divided distributions. The scope of the experiment only covers the unimodal distributions.

e) Normal truncated (μ, σ): the distribution is truncated between (MIN_{range}) and (MAX_{range}), μ (mean) is randomly created between (MIN_{range}) and (MAX_{range}), and σ (standard deviation) is randomly created between (MIN_{range}) and (MAX_{range}).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

f) Exponential truncated (β): the distribution is truncated between (MIN_{range}) and (MAX_{range}), β (continuous scale parameter) is randomly created between (MIN_{range}) and (MAX_{range}).

g) Lognormal truncated (μ,σ): the distribution is truncated between (MIN_{range}) and (MAX_{range}) , μ (mean) is randomly created between (MIN_{range}) and (MAX_{range}), and σ (standard deviation) is randomly created between (MIN_{range}) and (MAX_{range}).

4) 1000 randomly selected combinations of fuzzy numbers and related probabilistic distributions are generated.

5) 6 inputs are used in each of the 1000 combinations.

6) The probabilistic summations of each combination are performed using a Monte Carlo simulation that has the following assumptions:

   a) Total number of iterations =1000

   b) Random generator seed = 1

7) The target of the experiment is to calculate the absolute error between the expected values and variances of the outputs:

[2.16] Absolute Error (of expected values E.V.) = $100*(|E.V._{fuzzy} - E.V._{probability}|) / E.V._{probability}$

[2.17] Absolute Error (of standard deviations)

$$=100*(|\sqrt{\sum FuzzyVariances} - \sqrt{\sum ProbabilityVariances}|)/\sqrt{\sum ProbabilityVariances}$$

8) Statistics collected for the absolute errors are:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

a) Mean absolute error for all 1000 combinations.

b) Standard deviation of absolute error for all 1000 combinations.

Refer to Figure 2.6 for a flowchart of the experiment conducted to compare the probabilistic and fuzzy approaches in range estimating.

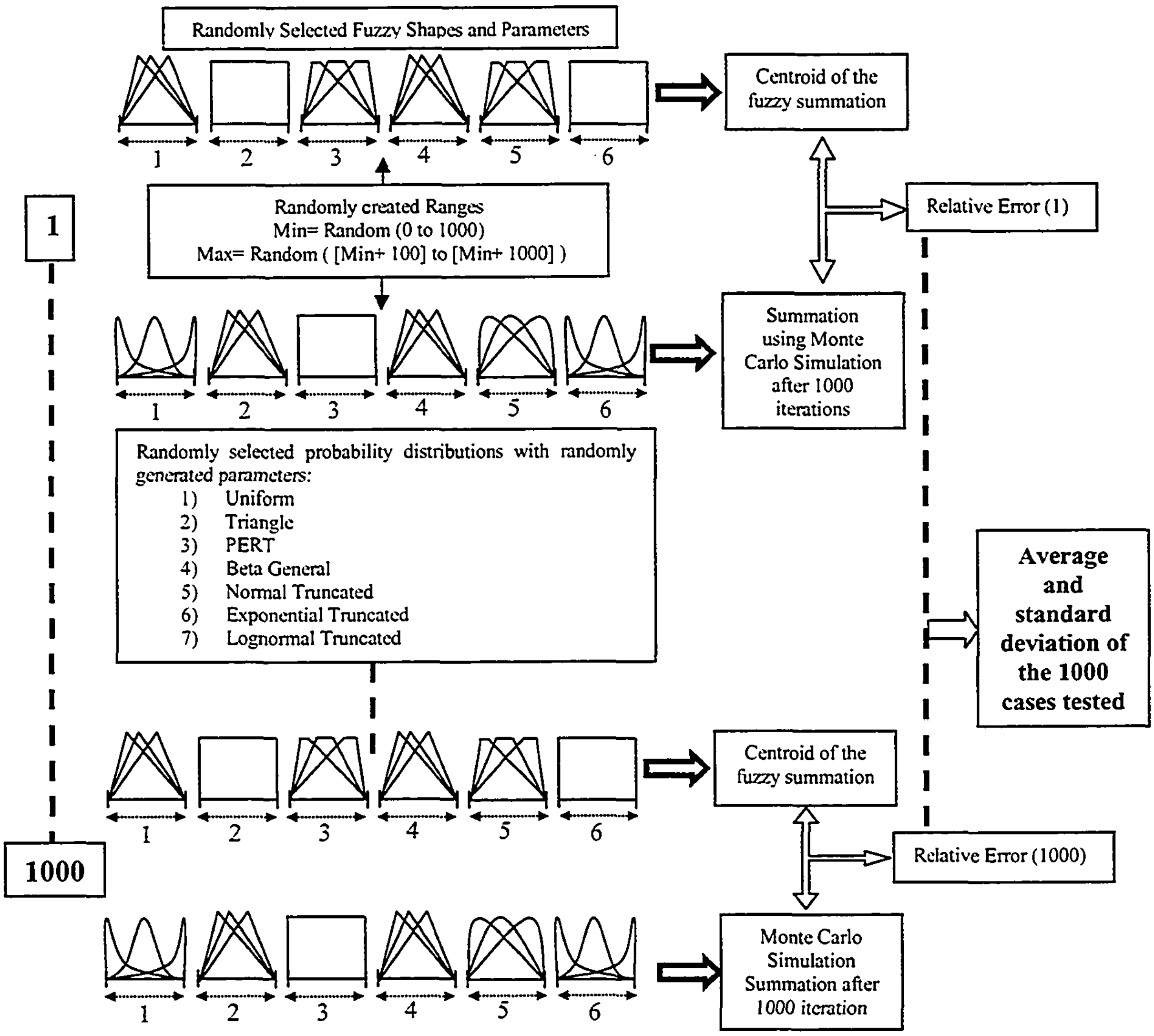


Figure 2.6: Flowchart of the Experimental Proof Conducted

The following points summarize the results of the experiment:

1) After running the experiment for 1000 iterations:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

a) The mean absolute error of the "expected values" comparison is 5.2 % with standard deviation = 4 %.

b) The mean absolute error of the "standard deviation" comparison is 13.5 % with a standard deviation = 11.6 %.

2) When the beta distribution is not included in the analysis, the results after 1000 iterations have improved:

a) The mean absolute error of the "expected values" comparison is 5.0 % with standard deviation = 3.8 %.

b) The mean absolute error of the "standard deviation" comparison is 9.6 % with a standard deviation = 6.9 %.

From the results generated we can conclude the following:

1) The fuzzy approach generated very comparable results to the probabilistic approach. This was proven experimentally when bounded and truncated probability distributions were used in the analysis.

2) When the beta distribution was removed from the experiment, better results were obtained which indicates that the beta distribution in general requires some parameter control to obtain better results. The shape parameters (α_1 and α_2) should be carefully selected to obtain the required results.

After proving the efficiency of the fuzzy approach in generating comparable results to the probabilistic approach, the following sections will show how the fuzzy numbers can effectively be used in range estimating.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

2.3.6 Fuzziness and Ambiguity Measures

There are different definitions and terms that describe "uncertainty". Uncertainty definitions can be categorized into two terms, vagueness and ambiguity. According to Klir and Folger (1988), vagueness is related to the difficulty of providing a sharp and precise distinction of a specific incident or phenomenon. As for the term "ambiguity", it describes the situation in which there is difficulty in making a specific selection or decision between alternatives. In fuzzy set theory, the concept of "fuzziness measure" deals with the first type of uncertainty (vagueness). The fuzzy measure is related to the degrees to which an arbitrary element of the universal set (X) belongs to the individual crisp subsets of X (Klir and Folger 1988). The other measure is the "ambiguity measure", which describes the lack of precision in determining the exact value of a magnitude (Delgado et al. 1998b). The following paragraphs will attempt to provide some details on the two measures and how they are obtained.

The fuzziness measure adopted in this study is the one developed by Klir and Folger (1988). Their approach defines the fuzziness of a set in terms of the lack of distinction between the set and its complement. That is to say that the more a set is different from its complement, the fuzzier it is. In order to calculate the fuzziness measure using this approach, a fuzzy complement approach and a distance function are utilized. For a given fuzzy set A(x), the measure of fuzziness, F(A) is obtained using the following equation:

[2.18] $$F(A) = \sum_{x \in X} (1 - |2A(x) - 1|)$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Equation 2.18 is only applicable to finite fuzzy sets, but it can be modified to fuzzy sets defined on infinite but bounded subsets. For example when X = [a, b], the measure of fuzziness is then obtained by:

$$[2.19]\quad F(A)=\int_a^b (1-|2A(x)-1|)\,dx$$

$$= b-a-\int_a^b |2A(x)-1|\,dx$$

The fuzziness of a crisp number or a fuzzy uniform number is zero because the lack of distinction between a fuzzy uniform number or a crisp number and their complements is zero.

As for the ambiguity measure, the approach developed by Delgado et al. (1998a) is selected as the ambiguity measure used in this study. According to (Delgado et al. 1998a), ambiguity "AG" is obtained by the following formula:

$$[2.20]\quad AG(\mu)=\int_0^1 r[R(r)-L(r)]dr$$

Where (μ) is a fuzzy number with r-cut representation (L(r), R(r)). The term [R(r)-L(r)] is the length of the r-cut interval (L(r), R(r)). Therefore, AG (μ) can be considered as a "global spread" of the fuzzy number. The ambiguity AG (μ) can be calculated for some of the most common fuzzy numbers as follows:

1) For a trapezoidal fuzzy number (a_1, a_2, a_3, a_4),

$$[2.21]\quad AG(\mu)_{Trapezoidal} = (a_3 - a_2)/2 + [(a_4 - a_3) + (a_2 - a_1)]\,/\,6$$

2) For a Triangular fuzzy number ((a_1, a_2, a_3,),

$$[2.22]\quad AG(\mu)_{Triangular} = [(a_3 - a_2) + (a_2 - a_1)]\,/\,6$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

3) For a uniform fuzzy number (a_1, a_2),

$$[2.23] \quad AG(\mu)_{Uniform} = (a_2 - a_1)/2$$

4) For a crisp number (a_1),

$$[2.24] \quad AG(\mu)_{Crisp} = 0$$

To clarify the difference between the fuzziness and ambiguity measures, Figure 2.7 provides a comparative illustration of the two measures for three types of fuzzy numbers (uniform, triangular, and trapezoidal). For comparative purposes, the three fuzzy numbers are defined on the same range. Figure 2.7 shows that the uniform fuzzy number has the least fuzziness measure, and the triangular fuzzy number has the highest fuzziness measure because the degree of belonging of the first is well defined (in terms of intervals) while the latter has a "fuzzily" defined degree of belonging depicted from the sloped lines that form the triangle. As for the trapezoidal fuzzy number, although it has sloped lines similar to those of the triangular fuzzy number, it is less fuzzy than the triangular fuzzy number because it contains a defined "flat or uniform range". This unique shape of the trapezoidal fuzzy number renders its fuzziness somewhere between that of the uniform fuzzy number and the triangular fuzzy number.

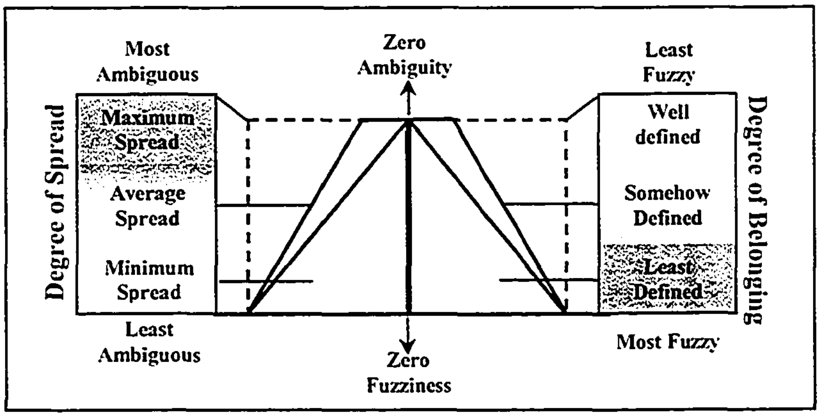


Figure 2.7: Fuzziness and Ambiguity Measure Definition

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

As for the Ambiguity measure, Figure 2.7 shows that the triangular fuzzy number has the least ambiguity measure, and that the uniform fuzzy number is most ambiguous. We also notice that the crisp number has zero fuzziness measure and zero ambiguity. If we take the crisp number as a reference, we notice that when the overall spread of any fuzzy number around this crisp number is minimized, the ambiguity is decreased. This explains why the uniform fuzzy number would certainly have the highest measure of ambiguity and the triangular fuzzy number would have the least. Again, the trapezoidal fuzzy number falls in between the two extremes, simply because it possesses, to some extent, the characteristics of both of the other fuzzy numbers.

2.4 Proposed Fuzzy Range Estimating Model

In this section, a detailed description of the model developed for fuzzy range estimating is provided. The model utilizes the different concepts discussed in the previous sections. Figure 2.8 shows a flowchart of the proposed model for modeling uncertainty in cost range estimating using fuzzy arithmetic. The following points explain the steps of the proposed model:

1) By consulting the experts involved in the process, the problem under investigation is identified. In the case of uncertainty in cost estimates, the major cost packages and their subgroups have to be clearly identified by the experts.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

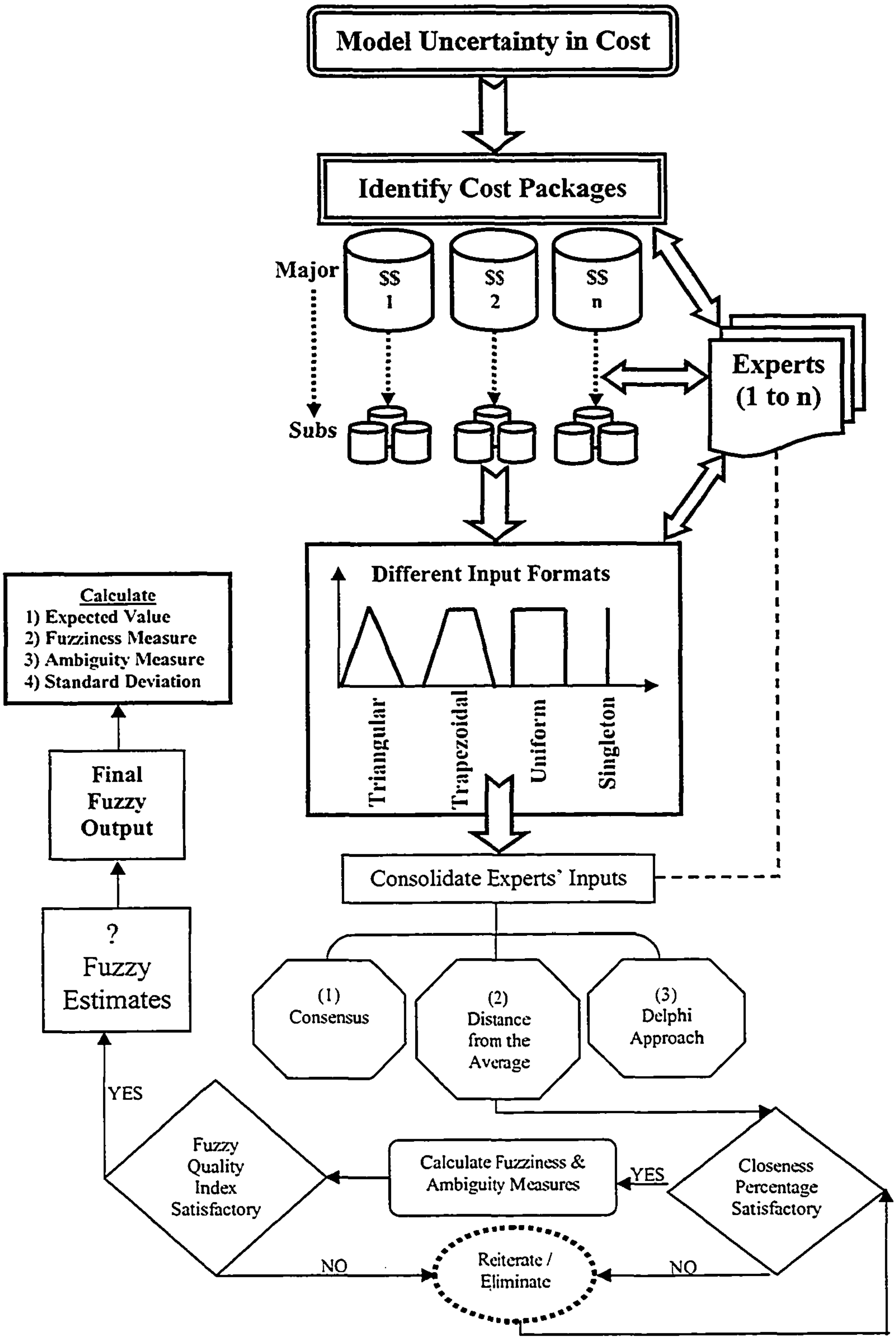


Figure 2.8: Flowchart of the Fuzzy Range Estimating Process

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

2) Each expert is then required to give his/her best estimate based on his/her experience choosing from the different input formats (triangular, trapezoidal, uniform, singleton). The logical meaning of these input formats is explained as:

a) Triangular fuzzy numbers: chosen when the expert believes that the estimate of the item has a "most likely" or "most plausible" point that is between a maximum and a minimum boundary.

b) Trapezoidal fuzzy number: the expert resorts to this input format when he/she believes that the estimate has a "most plausible" range or interval that lies between a maximum and a minimum boundary.

c) Uniform fuzzy numbers (sometimes referred to as a "crisp interval"): can be selected when the expert believes that the estimate should take the form of an interval that has a maximum and a minimum point, and when he/she finds it difficult to assign a "most plausible" estimate to a point or a range.

d) Crisp number (singleton): selected when the expert is 100 % certain that the estimate is a deterministic one that has zero uncertainty.

3) The next step is to consolidate the experts' inputs. The consolidation process can be done using three different approaches:

a) Consensus: The final estimates are made by consensus between the experts. This approach usually involves roundtable discussions between the experts involved in the estimation process. Generally, the greater the number of experts involved in the process, the more the time required to reach a consensus. This approach can be used in the estimation process within the same company or entity.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

b) Fuzzy Delphi Approach: The Delphi method is a systemic approach used for long range forecasting in which the estimates of experts are made to converge using statistical analysis (Kaufmann and Gupta 1988). It is conducted by interviewing highly qualified experts to give their opinions regarding specific issues such as providing cost and time estimates. Since it is difficult in most cases to provide precise and crisp estimates of cost or time, a fuzzy representation of the process can model this uncertainty. Fuzzy Delphi method can be an effective method for extracting subjective information from experts. However, the process can be time consuming. In addition, the way the technique is developed restricts the input option to one input type (triangular fuzzy number), which limits the expert's freedom in providing his or her estimate in different formats (i.e. trapezoidal fuzzy number).

c) Distance from the Average Approach: This approach is used in the model proposed in this paper to consolidate the fuzzy estimates. The consolidated fuzzy estimate is generated by averaging the experts' inputs and assessing them using the previously explained fuzzy arithmetic concepts. The process is as follows:

1) Inputs can be represented by a four-element format as follows:

   a) Trapezoidal Fuzzy number is (a, b, c, d)

   b) Triangular fuzzy number = (a, c, b) can be represented as (a, c, c, b)

   c) Uniform fuzzy number = (a, b) can be represented as (a, a, b, b)

   d) Crisp number = (a) can be represented as (a, a, a, a)

2) The average fuzzy number ($F_{average}$) is calculated as follows:

[2.25] $F_{average} = (\Sigma(a_1+a_2+..a_n)/n, \Sigma(b_1+b_2+..b_n)/n, \Sigma(c_1+c_2+..c_n)/n, \Sigma(d_1+d_2+..d_n)/n)$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Where, a, b, c, and d are the first, second, third, and fourth elements of the fuzzy numbers respectively, and (n) is the number of fuzzy numbers being averaged. Using equation 2.5, the expected mean value of each fuzzy number is calculated. The expected value of the average fuzzy number ($F_{average}$) is generated from equation 2.6. Another method of comparing the deviation of the fuzzy numbers from their calculated fuzzy average is by measuring the distance between their expected mean values:

$$\text{Distance } (i) = | (\hat{A_m})_i - (\hat{A_m})_{Faverage} | \quad \text{(Using the hamming distance)}$$

[2.26] $$\text{Closeness Percentage} = 100 * [1 - | (\hat{A_m})_i - (\hat{A_m})_{Faverage} | / (\hat{A_m})_{Faverage}]$$

Where, $(\hat{A_m})_i$ = the expected mean value of fuzzy number (i), $(\hat{A_m})_{Faverage}$ is the expected mean value of the average fuzzy number ($F_{average}$), and "closeness percentage" is a measure that calculates how close, by percentage, the value is to the reference. Table 2.1 shows the detailed distance calculations of expected mean values of different triangular fuzzy numbers and the expected mean value of their average ($F_{average}$). A closeness percentage can be set as an acceptance criterion and all closeness percentages that are equal or above this level are accepted.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 2.1: **Closeness Percentage of Different Triangular Fuzzy Numbers from their Average**

Parameters of Triangle Number (a, c, b)			Expected Values	Distance from Average	Closeness Percentage
a	c	b			
12.00	16.00	25.00	17.67	8.36	10.17
5.00	6.00	7.00	6.00	3.31	64.47
4.00	7.00	7.00	6.00	3.31	64.47
7.00	12.00	12.00	10.33	1.03	88.97
9.00	9.00	15.00	11.00	1.69	81.81
4.00	9.00	10.00	7.67	1.64	82.38
6.00	6.00	14.00	8.67	0.64	93.12
7.00	8.00	10.00	8.33	0.97	89.54
6.00	9.00	10.00	8.33	0.97	89.54
5.00	8.00	13.00	8.67	0.64	93.12
8.00	10.00	11.00	9.67	0.36	96.13
4.00	8.00	16.00	9.33	0.03	99.71
6.42	**9.00**	**12.50**	**9.31**	**0.00**	**100.0**

3) In the previous step, the expected mean values of the fuzzy numbers are compared to the expected mean value of the average fuzzy number in order to check the consistency of the experts' estimates and determine the estimates that are not in agreement with the overall average. This will help improve the quality of the average estimate by identifying the estimates that are not in harmony with the other estimates. However, the next step is to know what to do with these "rejected" or "eliminated" estimates. Two possible options can be taken as follows:

Option (1): Reiterate the estimating process by contacting the expert(s) who provided the rejected estimated. In this process, the experts are provided with all the information (i.e. overall fuzzy average and closeness percentage) that explains how deviant their estimates were· from the overall average estimate. The contacted experts are then asked to provide

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

different estimates that are closer to the average. Then, the averaging process is reiterated and the results are rechecked.

Option (2): Delete the "rejected" estimates and then recalculate the average. This option is chosen when contacting the experts, who provided rejected estimates, is not feasible due to time or other constraints.

The proposed modeling technique is different from the fuzzy Delphi process in that only the rejected estimates are returned to the experts for more feedback or justification, if possible, for the reasons behind providing estimates that deviated from the global mean. If the experts decide to change their rejected estimates, another averaging iteration is carried out using the adjusted estimates. If the results are satisfactory, the process moves on to the next step. If no change is made to the rejected estimates, they will be eliminated from the process. Another averaging will be carried out, and the closeness percentage recalculated and rechecked. This process is repeated until all the closeness percentages are above the acceptance level.

4) The next step is to evaluate the fuzziness and ambiguity measures of the estimates. The comparison in step (2) was necessary to check how precise the estimate is by measuring how close its expected value is to the expected value of the average. The fuzziness and ambiguity measure will try to check the "quality" of the estimates. As explained before, the fuzziness measure evaluates how vague the estimate is, and the ambiguity measure calculates the lack of precision in determining the exact value of a magnitude. Table 2.2 shows the evaluation process of the different measures.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 2.2: Comparison of Fuzziness Measure, Ambiguity Measure, and Fuzzy Quality Index

Type of Fuzzy Number	Fuzzy Parameters				Fuzziness Measure	Ambiguity Measure	Fuzzy Quality Index
	a	b	c	d			
Trapezoidal	2.0	7.0	10.0	20.0	7.5	4.0	5.8
Trapezoidal	3.0	7.0	9.0	15.0	5.0	2.7	3.9
Trapezoidal	10.0	10.0	15.0	18.0	1.5	3.0	2.3
Triangular	12.0	16.0	16.0	25.0	6.5	2.2	4.4
Triangular	7.0	12.0	12.0	12.0	2.5	0.8	1.7
Triangular	7.0	8.0	8.0	10.0	1.5	0.5	1.0
Uniform	6.0	6.0	30.0	30.0	0.0	12.0	6.0
Uniform	2.0	2.0	7.0	7.0	0.0	2.5	1.3
Uniform	6.0	6.0	7.0	7.0	0.0	0.5	0.3

As shown in Table 2.2, depending only on the fuzziness measure to assess the "quality" of the estimates is not usually enough. Although some fuzzy numbers have low measures of fuzziness, they still lack the necessary precision in determining the exact value, which renders them "ambiguous". A good example of this is uniform fuzzy numbers (sometimes referred to as crisp intervals). Although they will always have a zero measure of fuzziness, they have high measures of ambiguity. Therefore, it is important to assess the quality of the estimates using two measures (fuzziness and ambiguity). In order to provide a measure that combines the effect of both the fuzziness and ambiguity measures, a "fuzzy number quality index" is calculated using the weighted average of both measures. For a fuzzy set A, the fuzzy quality index can be measured as:

[2.27] $FNQI = [W_F * F(A) + W_{AG} * AG(A)] / [W_F + W_{AG}]$Where W_F and W_{AG} are the weights of the fuzziness measure and ambiguity measure respectively and F(A) and AG(A) are the fuzziness and ambiguity measures of fuzzy number (A) respectively.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

In this study, equal weights are assumed for both measures. Since there is no reference index against which the fuzzy quality measure can be compared, an acceptance level can be set by the user to reject the estimates with clearly high (FNQI) compared to the others.

After identifying the rejected fuzzy numbers, the same options explained in step (3) are investigated. The process terminates when the averaged fuzzy numbers for each item pass the checking and acceptance tests. The successful averaging process is the one that has the highest closeness percentages (high precision) and the least fuzziness and ambiguity measures (high consistency).

5) The process proposed in this thesis is a structured process for extracting information from the experts in the form of fuzzy numbers. Depending on the application, the process can work for both cost and activity duration evaluation. However, only the cost range estimating application is applied in this study. Therefore, in cost range estimating, the next step in the process after extracting the necessary information from the experts is to calculate the total estimated cost of the problem under consideration, which is performed using simple summation. The operation starts by adding up the cost of the sub items for each work package. The final total cost estimate is then obtained by adding up the costs of the work packages. The final evaluation, representing the total estimated cost of the modeled packages, is in the form of a fuzzy number. The following information will be calculated for the final fuzzy output:

a) Expected mean value of the output

b) Fuzziness measure

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

c) Ambiguity measure

d) Fuzzy Number Quality Index (FNQI)

These pieces of information are important in assessing the precision and quality of the output when compared to other outputs obtained from different estimating techniques. The following section provides an illustrative example comparing the fuzzy range estimating process to the Monte Carlo simulation technique.

2.5 Illustrative Example Comparing Fuzzy and Monte Carlo Simulation Approaches

The example selected as a case study is the North of Edmonton Sanitary Trunk (NEST) project. The budgeted cost of the project was $8.8 million and the preliminary estimated cost was $6 million. The City of Edmonton had concerns regarding the budget and wanted to know the chances of exceeding the preliminary estimated cost and being within the total budgeted cost. The main cost packages and their subcategories identified in the study are shown in Table 2.3.

A Monte Carlo simulation study was conducted to estimate the chances of meeting the budgeted cost (AbouRizk et al. 2004). In the study, the authors used Simphony's range estimating template to do the analysis. Simphony® is a specialized simulation tool that supports the Monte Carlo simulation technique in discrete event simulation and range estimating (AbouRizk and Mohamed, 2000). The model inputs are listed in Table 2.3.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 2.3: Data for the Case Study Project

Item	Description	Optimistic	Most Likely	Pessimistic
1	**Main Work Shaft**			
1.1.1	Mobilization--Move in	$40,000	$70,000	$100,000
1.1.2	Power Installation		$89,000	
1.1.3	Power-156 Str.	$15,000		$50,000
1.2	Excavate work shaft	$97,600	$122,000	$146,400
1.3	Excavate undercut	$200,000	$269,000	$350,000
1.4	Excavate tail tunnel to east	$100,000	$123,000	$150,000
1.5	Form and pour undercut		$80,000	
1.6	Form and Pour tail undercut		$39,000	
1.7	Form and Pour Shaft	$100,000	$120,000	$150,000
2	**Access Manhole**			
2.1	Excavate access shaft		$16,000	
2.2	Backfill shaft and install AMH		$44,000	
3	**Tunneling (866m)**			
3.1	Tunnel and install segments-866m price per m	$2,254	$2,474	$3,360
3.2	Patch and rub tunnel crown	$80	$134	$140
3.3	Patch and rub tunnel-final cleanup	$161	$188	$215
3.4	Spoil removal	$5.4	$8.1	$9.7
4	**Access manhole shaft**		$61,000	
5	**Tunneling (756 m)**			
5.1	Tunnel and install segments-756m price per m	$2,254	$2,474	$3,360
5.2	Patch and rub tunnel crown	$80	$134	$140
5.3	Patch and rub tunnel-final cleanup	$161	$188	$215
5.4	Spoil removal	$5.4	$8.1	$9.7
6	**Removal Shaft**		$101,000	

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

After 500 iterations, the following statistics were collected representing the Monte Carlo simulation model outputs:

Low estimated cost:	$ 5,486,345
High estimated cost:	$ 6,840,657
Mean estimated cost:	$ 6,059,263
Standard Deviation:	$ 280,249.7
80^{th} percentile:	$ 6,300,000

In this example, only the comparison between the final outputs of the probabilistic and fuzzy approach is performed. Therefore, for the fuzzy range estimating, the inputs represented by the fuzzy numbers are given the same shape of probability distributions used (i.e. a triangular fuzzy number is used when a probabilistic triangle distribution is used). The fuzzy output was generated after one iteration only by summing up all the fuzzy inputs involved in the analysis. The fuzzy output is a trapezoidal fuzzy number that has the following characteristics:

Fuzzy trapezoidal number parameters: a_1= $5038248.8, a_2= $5697250.2, a_3= $5732250.2, a_4= $7417863.4

Expected mean value:	$ 6054474
Standard Deviation:	$ 501,046.33
Fuzziness measure:	1172307.1 (11.7 scaled)
Ambiguity measure:	408269.1 (4.08 scaled)
FNQI:	7.89 (scaled)

The fuzziness and ambiguity measures were scaled (divided by 100,000) in order to make the comparison more readable and easier to grasp. The quality measures (fuzziness, ambiguity and FNQI) will provide more meaningful information when used as relative comparative indices, and can not be used to assess the quality of the fuzzy output unless more alternatives exist. Therefore, when more than one run is done and, hence, more than one fuzzy output is obtained, the fuzziness and ambiguity measure can be used to assess the quality of these outputs and check which one has the least FNQI.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

When comparing both outputs the following observations are made:

1) The difference between the probabilistic mean and the fuzzy expected value is 0.07% (the fuzzy output is less by 0.07 % only)
2) A unique concept in fuzzy set theory is the law of possibility or possibility measure (Kaufmann and Gupta, 1985). Using the possibility measure, one can determine which value is more plausible or possible. This type of information that fuzzy set theory provides is unique; probability theory does not support this concept, because the probability for a specific random variable to take place is almost zero (Lorterapong and Moselhi, 1996). Therefore, it is easy to determine which specific variable is more possible and plausible using this measure. For example, the possibility measure for the $6 million project estimated cost equals 0.84 which means that the project cost is possible to be $6 million with a 0.84 possibility. In addition, the most possible and plausible variable in a normal fuzzy number is the one that has possibility measure = 1.0. Therefore, the most possible and plausible output is ($ 5697250, $ 5732250).

2.6 Recommendations and Conclusions

Modeling uncertainty using fuzzy set theory was shown to be as effective as the probabilistic approach. The fuzzy set theory has the advantage of providing easier and faster to obtain outputs.

The study presented an effective approach for extracting subjective information from experts. Using the approach, the experts are free to express their knowledge in different formats (i.e., triangular or trapezoidal fuzzy numbers). The approach tries to combine and include the experts' information that is most consistent and most precise compared to

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

their total average. Outliers are either excluded or sent back to the expert for further review if time permits.

Modeling range estimating using fuzzy arithmetic yields very comparable outputs when compared to the probabilistic approach. The fuzzy approach has the advantage of being faster and easier to process because it only takes one iteration to generate the output, while it takes Monte Carlo simulation a number of iterations to generate a reasonable and reliable output. When the outputs are compared, the fuzzy expected mean value and the mean interval of confidence are found to be very comparable to the probabilistic mean, minimum and maximum values.

The possibility measure in fuzzy set theory is considered a unique concept that evaluates the plausibility for a specific variable within the fuzzy number to take place. Probability theory does not support a similar measure. In probability theory, the probability of ranges of variables can be evaluated (i.e. probability that the output is greater than or less than a certain value or between two variables).

The focus of this study is modeling cost range estimating using fuzzy set theory. Fuzzy range estimating in scheduling can also be investigated. The same methodology developed for the fuzzy cost range estimating problem can be utilized in fuzzy scheduling range estimating. However, different fuzzy arithmetic operations will be utilized in the scheduling calculations. The model developed by Lorterapong and Moselhi (1996) for project network analysis using fuzzy set theory is one of the models utilized fuzzy numbers in project scheduling. Different fuzzy arithmetic operations were utilized to calculate the forward and backward path calculations and criticality measurements.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

However, the model does not explain how the fuzzy numbers representing the activity durations are generated. Therefore, using the methodology proposed in this Chapter for extracting the fuzzy numbers from experts and a fuzzy scheduling technique similar to the one developed by Lorterapong and Moselhi (1996), a fuzzy range estimating operation in project scheduling can be developed.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

CHAPTER 3 – FUZZY EXPERT SYSTEM AS PREDICTIVE TOOLS

3.1 Introduction

In construction engineering projects, estimators and planners are required to provide estimates for the project cost items and project activity durations. The soundness and reliability of these estimates depend, to a great extent, on the estimator's intuition, judgment and past experience. In addition, there is a lot of uncertainty and subjectivity embedded within the process of providing cost and time estimates because there are many factors involved in the process which can affect the quality of the estimate if they are unaccounted for. For example, when estimating activity durations in any construction project, factors such as weather effect, labor performance, and level of work complexity can be taken into consideration in the estimating process. As the complexity of projects increases, more factors are involved. Some of these factors are subjective in nature and they are more difficult to account for. Even experienced estimators find it difficult to account for all the possible factors that can affect their decision. The estimator's decisions are usually based on experience and his/her embedded heuristic way of thinking. Therefore, for any estimate to be reasonably precise, the estimator has to have a lot of experience and the ability to account for as many factors that affect his/her decision as possible. As discussed in Section 1.1, many models have been developed to help minimize the uncertainty embedded within the prediction process. However, in the field of construction engineering, successful and reliable models should be able to mimic the human way of thinking and make use of historical data related to the problem. In this chapter, the utilization of fuzzy expert systems as a predictive tool in construction is

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

introduced. In addition, a methodology for developing the if-then rules in a fuzzy expert system is proposed.

3.2 Fuzzy Expert Systems

A fuzzy expert system is a fuzzy rule-based system that incorporates fuzzy logic concepts and approximates reasoning to reach a decision. A fuzzy inference system works by expressing human knowledge in the form of if-then rules so as to mimic the expert way of thinking, generalizing and reaching a decision. Both the premise and the conclusion of each rule can be expressed in linguistic terms, which are represented by membership functions. The structure of a fuzzy inference system is composed of four main components as shown in Figure 3.1.

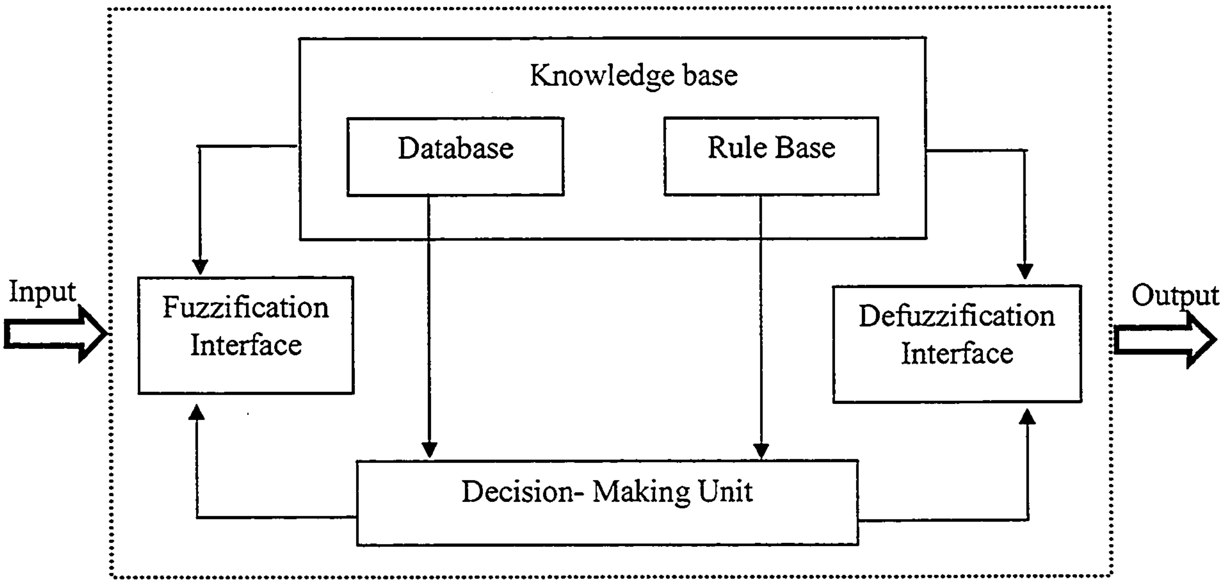


Figure 3.1: Components of a Fuzzy Expert System (Jang, 1993)

As Jang (1993) indicated, the four main components of a fuzzy inference system are:

1) Knowledge Base: which contains the fuzzy if-then rules, and a database that defines the membership functions of the fuzzy sets used in the rules.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

2) Decision-Making Unit: this component performs the inference operations on the fuzzy rules.

3) Fuzzification Interface: this part of the model transforms the crisp inputs into levels of truth of the linguistic terms, which are represented by fuzzy membership functions.

4) Defuzzification Interface: this component transforms the fuzzy results (i.e., output) of the model into a crisp value.

The way the system carries out the logic operation is mainly governed by the choice of fuzzy reasoning mechanism. The most important mechanisms are illustrated by Jang (1993) in Figure 3.2.

As shown in Figure 3.2, the steps of the fuzzy reasoning are (Jang, 1993):

1) Input variables are compared with the membership functions in the premise part to get the "level of truth" of each linguistic term (fuzzification).

2) The firing strength of each rule is calculated by combining the membership values in the premise part most commonly by multiplication or min (T-norm operators).

3) Depending on the firing strength, the qualified consequence is generated (either fuzzy or crisp).

4) The qualified consequent is aggregated to produce a crisp output(defuzzification).

As for the fuzzy reasoning, it can be classified into three types:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

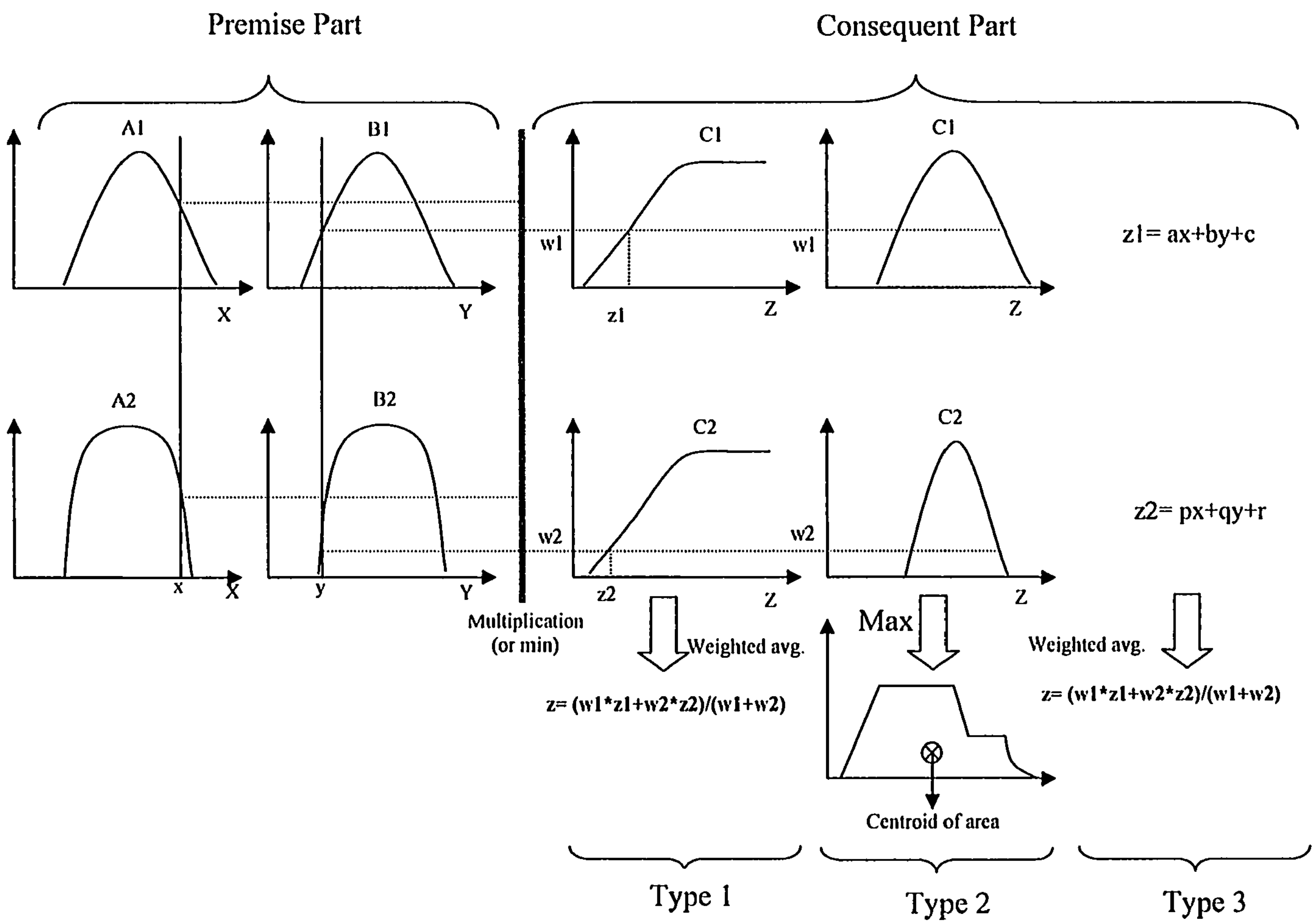


Figure 3.2: Fuzzy Reasoning Mechanisms (Jang, 1993)

Type 1: The overall output of the system is the weighted average of the crisp output of each rule which is measured by the rule's firing strength.

Type 2: The overall fuzzy output, as indicated from Figure 3.2, is generated by applying the max operation to the qualified fuzzy outputs. The final crisp output is measured by one of several different defuzzification schemes (e.g., centroid of area).

Type 3: The output of each rule in this type is a linear combination of the input variables plus a constant term. The final output is then calculated by taking the weighted average of the output of each rule.

Figure 3.3 shows a structure of a typical fuzzy expert system.

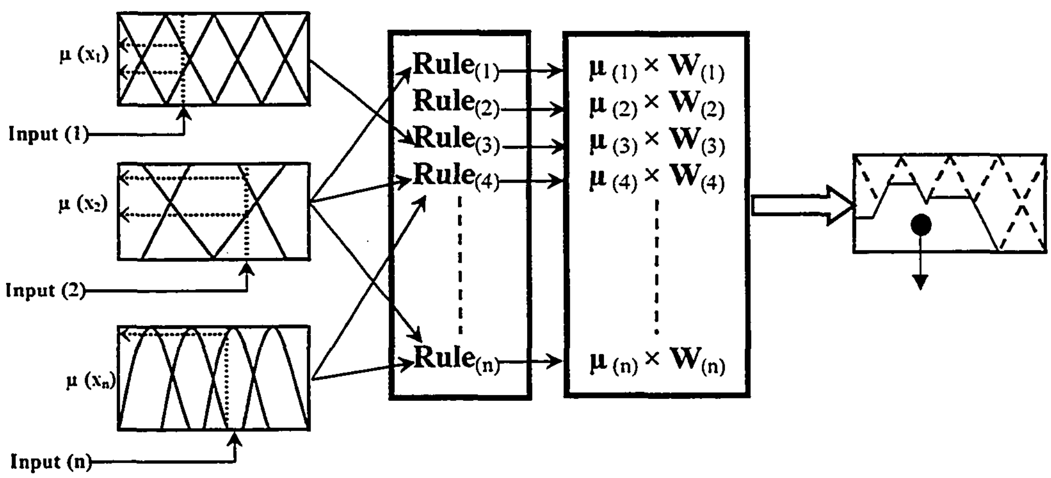


Figure 3.3: Structure of a Fuzzy Expert System

As indicated in Figure 3.3, a fuzzy expert system model has a systematic and layered structure, which can be explained in the following steps:

1) The first layer in the model is the "fuzzification" layer in which the system inputs are represented by linguistic variables in the form of membership functions. The membership functions are used to model the uncertainty

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

caused by the fuzziness and subjectivity of the input variables. The membership functions can take different shapes and numbers depending on the input variable being modeled. Triangular and trapezoidal membership functions are some of the most commonly used types of fuzzy membership functions since they mimic the way most experts think. Experts tend to provide their estimates in the form of most possible ranges "guess on the low/high and most likely".

2) The second layer is the fuzzy inference layer. It contains the if-then rules that control the fuzzy logic. Each rule has a premise part and a consequent part. The former represents the "IF" part of the rules and the latter represents the "THEN" part. The following is a sample of a fuzzy if-then rule.

   IF x_1 is M_1 (AND).....(AND) x_n is M_n THEN y_1 is N_1 (AND).....(AND) y_n is N_n

   Where ($x1 \ldots x_n$) and ($y_1 \ldots y_n$) are linguistic variables corresponding to input and output, respectively. ($M_1 \ldots M_n$) and ($N_1 \ldots N_n$) are the membership functions representing the linguistic terms of input and output respectively. (AND) is one of the operators or logical connectives. According to Rutkowska (2002), the two main classes of fuzzy operators are the intersection and union operators. The first is defined by so called triangular norms or T-norms. The latter is defined by the S-norms or T-conorms. The triangular norms are applied in fuzzy sets theory as logical connective "AND" which depicts the intersection between two terms. On

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the other hand, the S-norms are used to model logical connective "OR", which depicts the union between two terms. The most common forms of T-norms and S-norms are explained in the following examples:

For a and b $\in$ [0, 1], the T-norms are defined as follows:

a) min (a, b)

b) a b

c) max (a + b-1, 0)

And the S-norms are defined as follows:

a) max (a, b)

b) a+ b-a b

c) max (a + b, 1)

A combination between minimum and maximum operators with different degrees of compensations (parameters) can be a good logical operator. It is based on the belief that if both operators (minimum and maximum) are fulfilled, with a degree of compensation between both, it will yield better approximation of the logic (Wanous, 2000). The Min-Max operator is defined as follows (FuzzyTECH® 5.5 User's Manual, 2001):

$$\text{Min-Max}(\mu_1 \text{ to } \mu_n) = (1\text{-}?) \min (\mu i) + ? \ (\mu i)$$

Where ? is a degree of compensation from (0 to 1), for ? = 0, Min-Max = Min denoting AND operator and for ? = 1, Min-Max =Max denoting "OR" operator, and (i) is from 1 to n. Another compensatory operator is the Min-Avg which is defined as:

$$\text{Min-Avg}(\mu_1 \text{ to } \mu_n) = (1\text{-}?) \min (\mu i) + ? \sum_{i=1}^{n} (\mu i/n)$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

When ? = 0, Min-Avg. = min operator and when ? =1, Min-Avg = Avg. operator, and (i) is from 1 to n.

Another logical operator is "NOT" which is defined as (1-a), for "a" $\in$ [0, 1]. The "NOT" operator depicts the complement of a term.

3) The logic proceeds by calculating the corresponding level of truth of the different input variables. The membership level (μ) of each input variable is first calculated. The membership function(s) used will determine which rule(s) will be fired. The operators on the premise part of the fired rules determine the aggregation logic (i.e. min or max) as discussed in step (2). The degree to which each part of the premise is fulfilled can be referred to as the qualified level of truth. Then, on the consequent part of the rule, the fired rule's weight (degree of support) is multiplied by the qualified level of truth of the premise part. The multiplied value represents the total level of truth of the premise part. When more than one rule has the same conclusion, an output aggregation method is used. One of the output aggregation methods is the (max) in which the rule that has the maximum level of truth of the premise part is considered.

4) The final aggregated fuzzy output can be defuzzified into a crisp one using one of the different defuzzification methods. Some of the most common defuzzification methods are (Altrock, 1995):
    1) Center of area method: Is sometimes referred to as center of gravity. This method calculates the centroid of the area under the resulting functions of all the terms which represents the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

compromise value of all the terms. This method can be applicable to control and decision making modeling.

2) Center of maximum method: It also calculates the compromise value of all the terms by balancing out the resulting functions. It considers the resulting functions as weights at the position of the maximum values of the terms. The center of maximum value is calculated by multiplying the inference results (or weights) by the maximum values of each corresponding term. Similar to the center of area method, the center of maximum method is applicable to control and decision making modeling.

3) Mean of maximum method: This method is developed to measure the "most plausible" result rather than the compromise value that the first two methods measure. It calculates the typical value of the term that is most valid by selecting the maximum value of a membership function that has the maximum resulting weight. This method is more applicable for pattern recognition and data analysis.

It can be noticed that one of the major steps in developing a fuzzy expert system is to generate the if-then rules that encompass all the expert knowledge and experience about a specific domain. Therefore, a methodology is proposed to maximize the efficiency when generating the if-then rules. The following section explains how the methodology can help achieve a better rule developing process.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

3.3 A Proposed methodology to Develop Fuzzy Expert System Rules' Consequents

Developing the fuzzy expert system rules is one of the most important steps because the rule base represents the reasoning and logic mechanism of the system. The rule base is usually determined by the experts subjectively. Therefore, a structured methodology is proposed to develop the rule base system and the related degrees of support to guarantee better representation of the factors included within the rule blocks. The main idea of the methodology is to aggregate the effect of the factor's "relative importance" within the rule block and its "impact on the output" being at a specific state (i.e. Medium). The methodology is shown in Figure 3.4 and is explained in the following sections.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

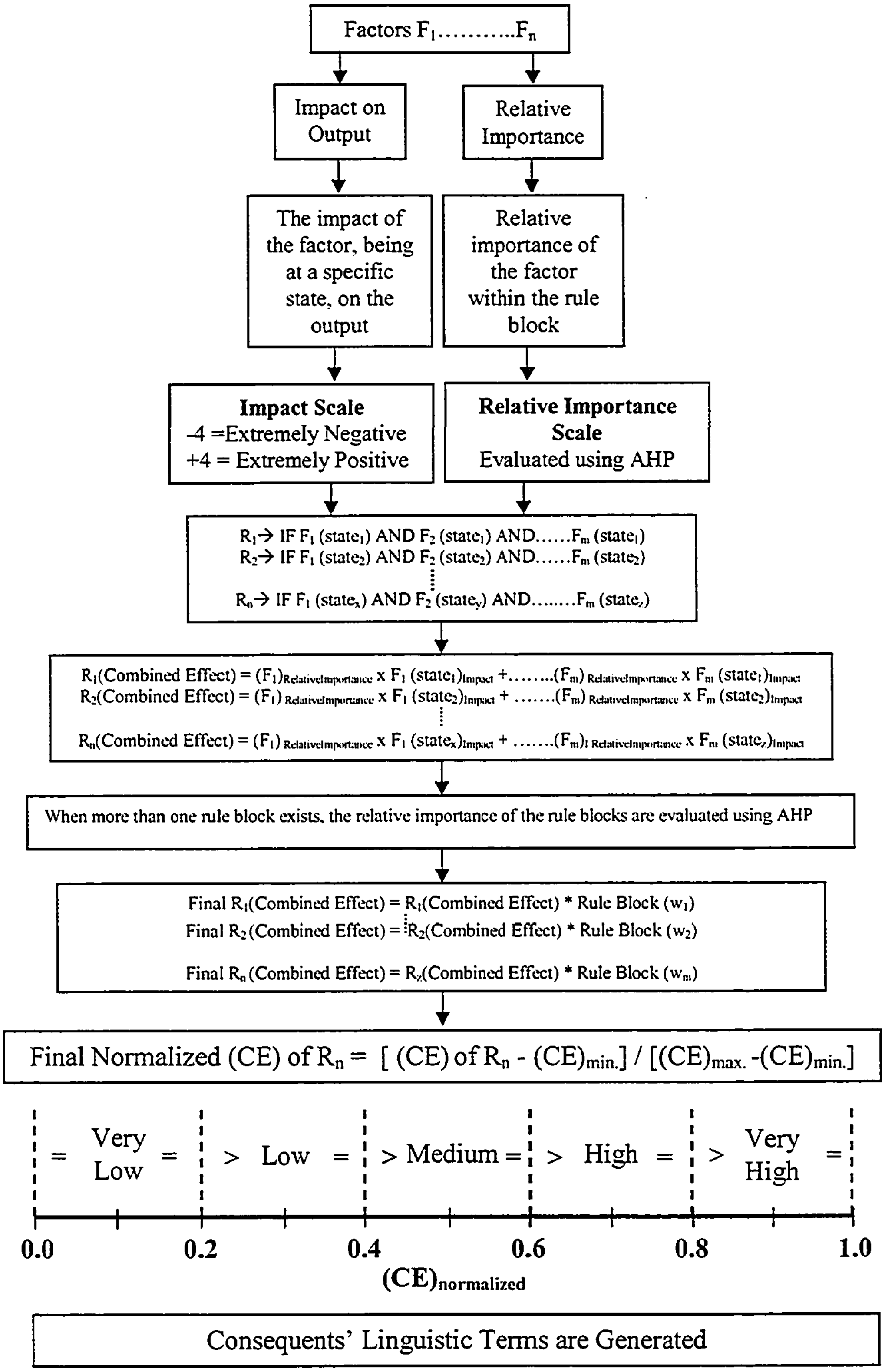


Figure 3.4: A Methodology to Evaluate Rules' Consequents

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

3.3.1 Initial Steps

The rules' antecedents are determined by calculating the total number of possible combinations of the factors' linguistic terms. The number of rules generated is determined using the following formula:

$$\text{No. of Rules} = (\text{no. of input terms})^{(\text{no. of inputs})}$$

One way to control the exponential growth of rules generated is to group the related factors in different "rule blocks". Factors within rule blocks are grouped based on their class commonalities. For example, if wind speed, precipitation and average daily temperature are some of the factors that affect an activity duration, it can be noticed that these factors belong to the class of "weather factors". Therefore, these factors can be grouped in one rule block since they share common characteristics. The rules are then generated using the factors grouped within the rule blocks. Fired rules from different rule blocks are then aggregated using a specific inference and aggregation operator as discussed in Section 3.2 in order to generate the final output of the system. Figure 3.5 shows a schematic diagram of the factors being grouped in different rule blocks in a fuzzy expert system.

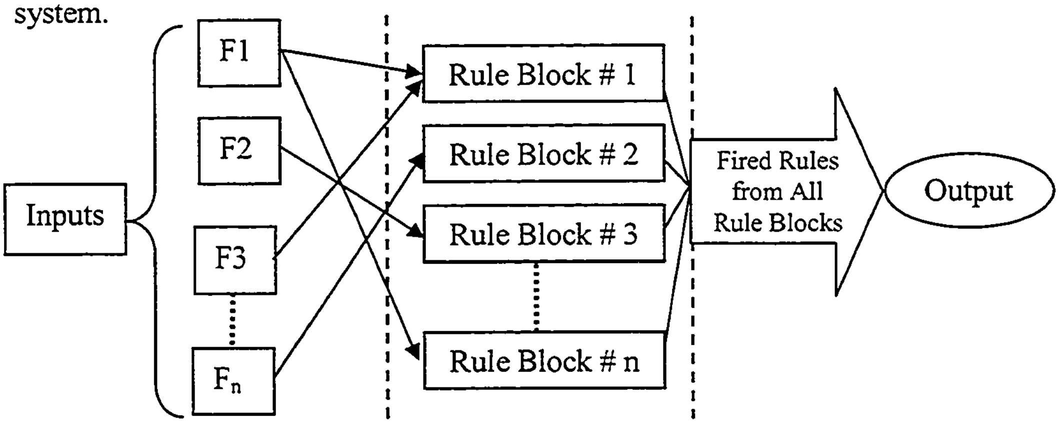


Figure 3.5: Rule Blocks in Fuzzy Expert Systems

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

After grouping the factors and determining the total number of rules within each rule block based on the linguistic term combinations of the factors, the consequents of each rule has to be evaluated. The following sections explain how this process can be done in a structured methodology.

3.3.2 Relative Importance of Factors Using AHP

The relative importance or "contribution to the output" of each factor relative to the other factor has to be evaluated first. The relative importance of each factor is evaluated relative to the factors within the same rule block. One of the methodologies used to evaluate the relative importance of a multi-objective problem is the Analytic Hierarchy Process (AHP) which was first developed by Saaty (1980). In Saaty's approach, the multi-attribute problem is structured into a hierarchy of interrelated elements, and then a pairwise comparison of elements in terms of their dominance is conducted. The weights were given by the eigenvector associated with the highest eigenvalue of the reciprocal ratio matrix of pairwise comparisons. A detailed description of the AHP steps and calculations can be can be found in Saaty (1980).

3.3.3 Impact of Factor on Final Output

After evaluating the relative importance of each factor compared to the other factors, the next step is to evaluate the impact of a factor being at a specific state (i.e. medium) on the final output. For example, if factor (x) is high, what would be the impact of this state on the final output? The impact is measured by a scale ranging from -4 (extremely negative impact) to +4 (extremely positive impact).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The full adopted scale is shown in Figure 3.6. The following paragraph will illustrate how the factors' impacts on final output are evaluated using the proposed scale.

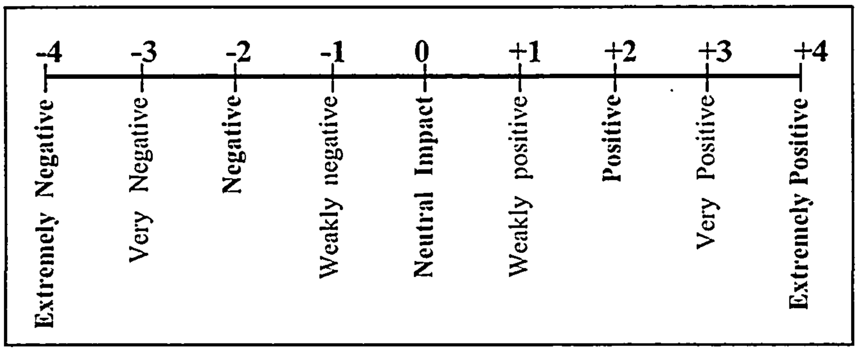


Figure 3.6: Impact Scales of Factors on Final Fuzzy Output

Suppose that we are trying to predict the crew's efficiency by percentage (%). Four factors are assumed to be affecting the crew's efficiency. The first and second factors are weather related factors which are the average temperature on site (F1) and wind speed (F2). The second and third factors are crew related, which are the crew's average experience (F3) and crew's skill level (F4). The linguistic terms (states) that describe the inputs are High-Medium-Low for both factors.

The experts' role is to evaluate the impact or effect of these linguistic terms on the final output (crew's efficiency) in general as:

1) Average temperature (F1)

IF (F1) is LOW→ impact on crew's efficiency = -3 (very negative)

IF (F1) is MEDIUM→ impact on crew's efficiency = 3 (very positive)

IF (F1) is HIGH→ impact on crew's efficiency = -1 (weakly negative)

2) Wind Speed (F2)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

IF (F1) is LOW→ impact on crew's efficiency = 3 (very positive)

IF (F1) is MEDIUM→ impact on crew's efficiency = -2 (negative)

IF (F1) is HIGH→ impact on crew's efficiency = -3 (very negative)

3) Crew's average experience (F3)

IF (F2) is LOW→ impact on crew's efficiency = -1 (weakly negative)

IF (F2) is MEDIUM→ impact on crew's efficiency = 3 (very positive)

IF (F2) is HIGH→ impact on crew's efficiency = 4 (extremely positive)

4) Crew's skill level (F4)

IF (F2) is LOW→ impact on crew's efficiency = -1 (weakly negative)

IF (F2) is MEDIUM→ impact on crew's efficiency = 3 (very positive)

IF (F2) is HIGH→ impact on crew's efficiency = 4 (extremely positive)

3.3.4 Rule Consequents

Once the relative importance of all the factors within the rule block and their related impacts on the final output are evaluated, the consequents of the rules and the degrees of support can be evaluated. The following steps show the calculations required to generate the rule consequents and degrees of support:

1) The number of possible rule combinations is determined as explained in Section 3.7.1. Assume that the total number of rule combinations is R_n.

$$R_1 \rightarrow \text{IF } F_1\ (\text{state}_1) \text{ AND } F_2\ (\text{state}_1) \text{ AND} \ldots\ldots F_m\ (\text{state}_1)$$
$$R_2 \rightarrow \text{IF } F_1\ (\text{state}_2) \text{ AND } F_2\ (\text{state}_2) \text{ AND} \ldots\ldots F_m\ (\text{state}_2)$$
$$\vdots$$
$$R_n \rightarrow \text{IF } F_1\ (\text{state}_x) \text{ AND } F_2\ (\text{state}_y) \text{ AND} \ldots\ldots F_m\ (\text{state}_z)$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

2) The Combined Effect (CE) of the relative importance and the impacts of the factors in the different generated rules as:

Equation [2.1]:

$$R_1(CE) = (F_1)_{RelativeImportance} \times F_1\,(state_1)_{Impact} + ..(F_m)_{RelativeImportance} \times F_m(state_1)_{Impact}$$
$$R_2(CE) = (F_1)_{RelativeImportance} \times F_1\,(state_2)_{Impact} + ..(F_m)_{RelativeImportance} \times F_m(state_2)_{Impact}$$
$$\vdots$$
$$R_n(CE) = (F_1)_{RelativeImportance} \times F_1\,(state_x)_{Impact} + ..(F_m)_{RelativeImportance} \times F_m(state_z)_{Impact}$$

3) When rules are classified into more than one rule block, the importance of each rule block representing the contribution of the rules contained within it to the final output has to be calculated. Weights or importance of rule blocks are evaluated using a scale from 0 (no weight or no importance) to 10 (most weight or most important). Each weight is then normalized by dividing it by the maximum weight given to a rule block. The normalized weights of the blocks are then multiplied by the combined effects (CE) of the rules within a rule block.

4) The final combined effects of all the rules (irrespective of the rule block) are then normalized between 0 and 1 using the following formula:

Equation [2.2]:

$$\text{Final Normalized (CE) of } R_n = [\,(CE) \text{ of } R_n - (CE)_{min.}] \,/\, [(CE)_{max.} - (CE)_{min.}]$$

5) The last step is to estimate the rule consequent using a scale whose number and names of linguistic terms are equivalent to that of the system output. The scale used ranges from 0 to 1.0. Figure 3.7 illustrates a five-term scale ranging from very low to very high.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

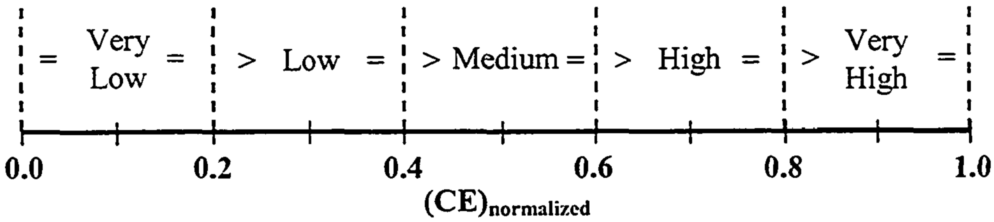


Figure 3.7: The Linguistic Term Zones to Evaluate Rules' Consequences

The main goal of this scale is to cluster the normalized combined effects of the rules' antecedents into the linguistic terms used by the system output.

Let us take the example of crew's efficiency used in Section 3.7.3 to illustrate how the scale is used to evaluate the rules' consequents. We will assume that the number of the linguistic terms of the crew's efficiency are five and range from very low to very high as shown in Figure 3.8.

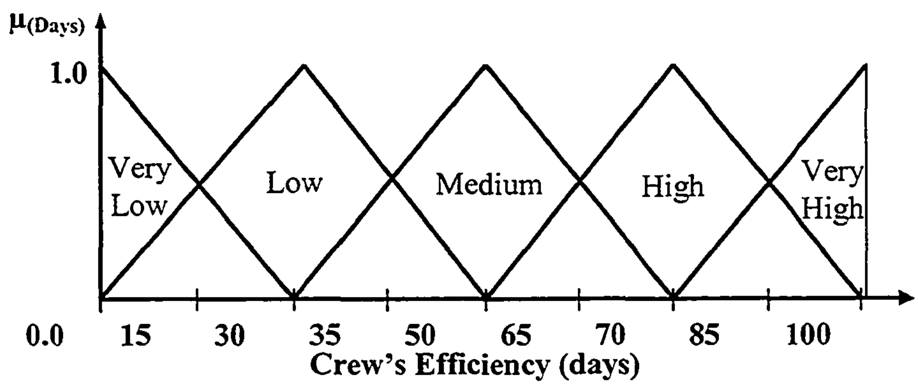


Figure 3.8: Membership Functions of the Crew's Efficiency of the Illustrative Example

The related factors are grouped in two rule blocks (weather related and crew related rule blocks). The weather related rule block contains temperature (F1) and wind speed (F2). The crew related rule block contains crew's experience (F3) and crew's skill level (F4). The values of the relative importance of the different factors are: 0.67 for temperature (F1), 0.33 for wind speed (F2), 0.67 for crew's experience (F3) and 0.33 for crew's skill level (F4). . The total number of rule

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

combinations in each rule block is 9. The rules' antecedents of the 18 rules are as follows:

R_1→ IF Temperature is LOW and Wind speed is LOW
R_2→ IF Temperature is LOW and Wind speed is MEDIUM
R_3→ IF Temperature is LOW and Wind speed is HIGH
R_4→ IF Temperature is MEDIUM and Wind speed is LOW
R_5→ IF Temperature is MEDIUM and Wind speed is MEDIUM
R_6→ IF Temperature is MEDIUM and Wind speed is HIGH
R_7→ IF Temperature is HIGH and Wind speed is LOW
R_8→ IF Temperature is HIGH and Wind speed is MEDIUM
R_9→ IF Temperature is HIGH and Wind speed is HIGH

R_{10}→ IF Crews' Experience is LOW and Crews' skill Level is LOW
R_{11}→ IF Crews' Experience is LOW and Crews' skill Level is MEDIUM
R_{12}→ IF Crews' Experience is LOW and Crews' skill Level is HIGH
R_{13}→ IF Crews' Experience is MEDIUM and Crews' skill Level is LOW
R_{14}→ IF Crews' Experience is MEDIUM and Crews' skill Level is MEDIUM
R_{15}→ IF Crews' Experience is MEDIUM and Crews' skill Level is HIGH
R_{16}→ IF Crews' Experience is HIGH and Crews' skill Level is LOW
R_{17}→ IF Crews' Experience is HIGH and Crews' skill Level is MEDIUM
R_{18}→ IF Crews' Experience is HIGH and Crews' skill Level is HIGH

The weights/importance of each of the rule blocks are assumed as follows: crew related = 10 and weather related = 7. After normalization, the weight of the crew related rule block = 1.0 and the weight of the weather related rule block = 0.7. The combined effects of each of the rules and the linguistic terms of the consequents are shown in Table 3.1.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 3.1: Consequent Linguistic Terms of the Illustrative Example

Rule	Factor Status	Factor Status	Weight of Rule Block	(CE)	Norm. (CE)	Output Linguistic terms and Degree of support				
						VL	L	M	H	VH
Weather Related Rule Block										
R_1	L	L		-0.7	0.23		X			
R_2	L	M		-1.9	0.04	X				
R_3	L	H		-2.1	0.00	X				
R_4	M	L		2.1	0.69				X	
R_5	M	M	**0.7**	0.9	0.50			X		
R_6	M	H		0.7	0.46			X		
R_7	H	L		0.2	0.38		X			
R_8	H	M		-0.9	0.19	X				
R_9	H	H		-1.2	0.15	X				
Crew Related Rule Block										
R_1	L	L		-1.0	0.18	X				
R_2	L	M		0.3	0.40		X			
R_3	L	H		0.7	0.45			X		
R_4	M	L		1.7	0.62				X	
R_5	M	M	**1.0**	3.0	0.84					X
R_6	M	H		3.3	0.89					X
R_7	H	L		2.4	0.73				X	
R_8	H	M		3.7	0.95					X
R_9	H	H		4.0	1.00					X

Section 3.7 has introduced a structured methodology to generate the rule consequents based on the relative importance of the system inputs and their impact on the system output when being at a specific linguistic term. The methodology is designed to minimize the subjectivity involved in developing the fuzzy expert system by generating the rules' consequents in a systematic and structured methodology. The following sections will show how the fuzzy expert system is utilized to model the Tunnel Boring Machine (TBM) penetration rate in soft ground soils.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

3.4 Introduction to Tunneling Operations

Tunnel Boring Machine (TBM) penetration operation is a multi-attribute process that involves many quantitative and qualitative factors affecting it. In order to model the TBM penetration operation and be able to predict the rate of penetration, all the significant factors affecting the operation must be identified and modeled. The operation must be developed by a system that is able to handle and model both the subjective and quantitative nature of the factors. As explained in Chapter 1, the input modeling in discrete event simulation models only account for these factors implicitly within the probability distributions. Accordingly, many parameters related to these factors will be overlooked. Therefore, other modeling techniques should be used to model this type of problems. A fuzzy expert system will be used to model the TBM advance rate prediction in soft ground soils. The model development steps will follow the same guidelines and information introduced Sections 3.2 and 3.3.

In general, tunneling operations in construction engineering involve a number of different processes. As shown in Figure 3.9, the tunneling operation involves the following major processes:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

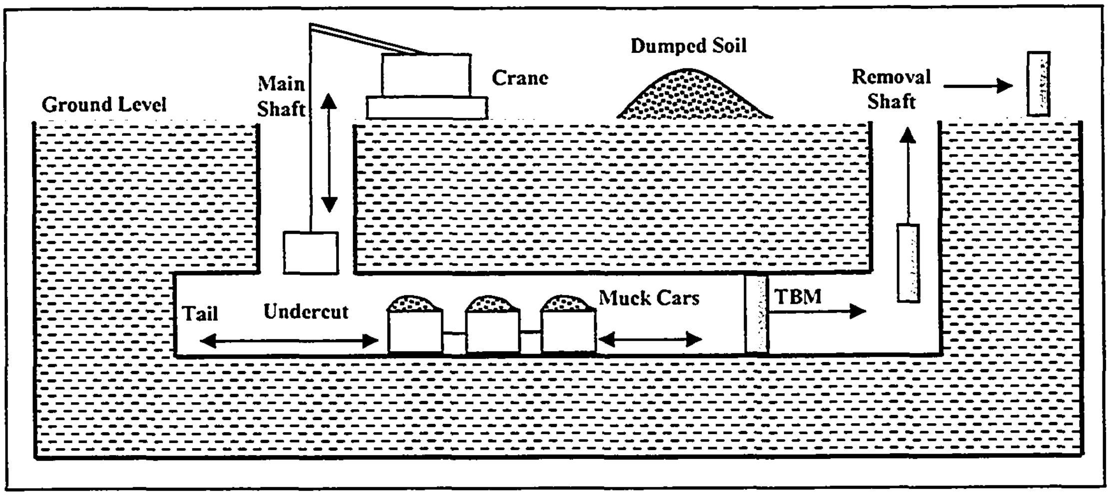


Figure 3.9: Cross-Sectional View of a Tunneling Operation

1) Shaft Construction: the excavation and support of a vertical shaft to the level of the tunnel.

2) Undercut / Tunnel Tail Construction: the excavation and support of an area under the shaft that is used in the processes of dirt removal and material handling.

3) Tunnel Construction: the main tunneling operation, which involves three main processes: excavation, dirt removal, and tunnel support (lining). These components are detailed as follows:

    a) Excavation: The main tunneling component that involves digging horizontally along the tunnel direction. This process can be done using different techniques depending on the tunnel length and complexity. Tunnel Boring Machines (TBM) is one of the techniques used for excavation. Also, the excavation process can be done by hand excavation.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

b) Dirt Removal: Excavation requires a process of dirt disposal from the tunnel face. This process includes transferring the dirt to the undercut area and hoisting the dirt up to the ground level. This can be done using one or two trains of muck cars that move back and forth on rail tracks.

c) Tunnel Support (Lining): This process deals with supporting the excavated tunnel by an appropriate material. One lining material is pre-cast concrete segments, which can be installed by a TBM as it advances. Other lining materials are PVC or steel pipes.

4) Removal Shaft: this process is done if a TBM machine is used in the tunneling process. At the end of the excavation, a removal shaft must be excavated to hoist the TBM machine up to the ground level so it can be used on other projects.

The excavation operation is one of the major controlling operations in tunneling. The TBM penetration rate will highly contribute to the final tunneling productivity. Faster TBM penetration rates will increase the number of tunnel meters excavated per shift. Therefore, it is important to accurately model the TBM penetration rate in order to study and predict the performance of the TBM within different conditions. Modeling the TBM penetration rate will help provide better planning and decision making in the tunneling operation. The following section reviews the different research work developed to model tunneling operations. Section 4 introduces the studies conducted on TBM advance rates.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

3.5 Background and Literature Review

No comprehensive study for predicting the TBM penetration rate in soft ground soils has been developed yet. Instead, several researchers have studied the TBM performance in rock. The studies conducted on TBM performance in rocks can be divided into two groups. The first includes the studies conducted based on empirical methods. The second group includes the studies that utilized artificial intelligence to model the TBM penetration rate in rock. The following paragraphs discuss some of the research work conducted on TBM performance prediction in rock.

Snowdon et al. (1982) developed an empirical relation between the TBM penetration rate and the cutter-head diameter (m), rock uniaxial compressive strength (MPa), and TBM thrust force (kN). In another study, Innaurato et al. (1988) developed an empirical equation relating the TBM penetration rate (mm/round) to the uniaxial compressive strength (MPa) and rock structure rating (a rating used to assess the rock quality).

Grima and Bruines (2000) modeled the TBM performance in rock using NeuroFuzzy modeling. Based on literature and statistical analysis, they concluded that the most influential factors affecting the TBM penetration rate in rock are the rock mass properties, machine characteristics and the geometry of the tunnel. They studied 640 TBM projects world wide and concluded using statistical analysis that eight factors can be considered for the model development. Some of the selected factors were the core fracture frequency (a parameter that measures the discontinuity in a rock mass), the unconfined compressive strength,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

tunnel diameter, torque, and revolution per minutes (RPM). By investigating and testing several NeuroFuzzy models with different factor combinations, they found that a NeuroFuzzy model, based on Takagi-Sugeno method, with five factors yielded the best results. The final factors used in the model were the core fracture frequency, unconfined compressive strength, RPM, thrust per cutter and cutter diameter.

Okubo et al. (2003) developed an expert system model to study the feasibility of using tunnel boring machines in certain projects in Japan when limited information is available for pre-feasibility studies. The model is divided into three stages. In stage (A), the authors developed a set of conditions or requirements above or below which the use of TBM should be reconsidered. For example, if the tunnel length is less than 500 m and if the excavation is less than 2 meters or greater than 10 meters, then the TBM will not be feasible to use. In stage (B), using an iterative procedure the total advance rate per day is estimated. First the value for the penetration rate (m/h) is assumed then used to calculate the penetration (m), uniaxial compressive strength and rolling and thrust forces using previously developed empirical equations. The calculated values are checked against a point system to see whether they are within a reasonable range and if not a new estimate of penetration rate is proposed until all calculated parameters are within a reasonable range which will be used later to calculate the working advance rate (m/day). The final stage (C) was developed to assess the estimates generated in step (B) using a knowledge base extracted from different experts.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Benardos and Kaliampakos (2004) developed a neural network model to predict the overall TBM advance rate (m/day). In their model, the authors considered 8 factors related to the rock properties and geological setting such as the rock mass permeability, weathering degree of rock mass, uniaxial compressive strength and water table surface. The data used in the analysis were collected from boreholes that were spatially modeled so as to identify the properties within the 12 m thick stratum along a tunnel selected as a case study. The tunnel was dividend into 11 control segments, which represent the total number of data set used in training and testing the model. The relative testing error of the systems ranged from 6 to 8 %.

3.6 Factors Affecting TBM Penetration Rate

Identifying the key factors affecting TBM penetration rate in soft ground soil is considered the most important, yet most time consuming stage for modeling the TBM penetration rate. Several methods are adopted to identify the factors that highly affect the TBM penetration rate. Literature, previous projects, and experts' opinions are some of the techniques used to identify the factors. The following sections elaborate on the different techniques used.

3.6.1 Literature

The following points represent some of the information that can be extracted from literature:

a) How the tunneling operations using TBM's are handled.

b) Types of soft ground soils.

c) Soil properties and behaviors.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

d) TBM types.

e) TBM performance and specifications.

f) TBM penetration mechanism.

These pieces of information can be found in books, machine specification documents and previous models developed in the tunneling fields as explained in Section 3.4.

3.6.2 Previous Tunneling Projects Using TBM

To develop more efficient and reliable models, a full and clear understanding of the process being modeled must be first achieved. Understanding the tunneling operation can be achieved by extracting information from literature as indicated in section 3.6.1 and reviewing the previous tunneling projects. Previous tunneling projects can provide the following types of data:

a) Actual productivity data.

b) Actual TBM performance.

c) Actual soil types encountered and their properties.

d) Problems encountered and sources of delays.

e) Labour and operators performance.

The City of Edmonton is very experienced in the field of utility tunneling. Many utility tunneling projects are being handled yearly. More details on the City of Edmonton tunneling history can be found in Ruwanpura (2001). Several site visits were made to actual tunneling sites. In addition, studies made on actual tunneling projects in Edmonton were studied and reviewed. For example, the study by AbouRizk et al. (2004) and explained in Section 4.3 is one of the studies

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

used as a review reference.

3.6.3 Experts' Opinion

The different types of information discussed in sections 3.6.1 and 3.6.2 were used to formulate an interview questionnaire that was used to extract the experts' knowledge in the field of tunneling. The experts' feedback and comments are important to validate the information presented and to account for any missing piece of information that is deemed essential to the TBM penetration process. The objective of the interview questionnaire is to capture the experts' ways of thinking with respect to the TBM penetration operation. Capturing the experts' knowledge will help develop the rule-base that will be used to model the TBM penetration operation. Therefore, to achieve the previously discussed objectives, a questionnaire was formulated to collect the participants' feedback on the factors affecting the TBM penetration operation. The complete forms used in the interview questionnaire and a description of the factors included in the interview questionnaire can be found in Appendix (A). Several qualitative and quantitative factors were included in the questionnaire. The interview questionnaire is divided into the following three parts:

a) The objective of the first part was to collect information on how significant each listed factor is to the TBM penetration operation.

b) The second part was used to help build the rule base of the fuzzy expert system. It is designed to study the effect of each factor on the TBM penetration rate. It investigates how positively or negatively each factor affects the TBM penetration rate. For example, if the increase of a certain

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

factor will cause the overall TBM penetration rate to increase, it implies that the factor is positively affecting the TBM penetration rate. If the increase of a certain factor will cause the overall TBM penetration rate to decrease, it means that the factor is negatively affecting the TBM penetration rate.

c) The last part of the survey questionnaire will help develop the fuzzy membership functions of the subjective factors. In this part, the expert is asked to attach a numerical value of the linguistic term describing each subjective factor. For example, for the "Large tunnel diameter", the expert is required to provide his/her numerical representation of the term "Large tunnel diameter" in a triangular format (i.e. least possible, most likely, and largest possible) or a trapezoidal format (i.e. least possible, a range for the most likely, and largest possible).

The factors are divided into the following six sections:

1) Tunnel properties: It lists all the major properties of the tunnel such as diameter, depth, and layout.
2) Soil Properties: It includes the factors related to soil such as soil type, moisture content, and water level.
3) TBM properties: It lists the TBM properties such as TBM thrust and age.
4) Operator's performance: This section includes the factors that assess the operator's performance such as experience and skill level.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

5) Shift related: It includes the factors that describe the way the shift is being managed such as shift duration, and day of week.

6) Weather related: This section lists some of the major weather factors that can affect the overall work progress such as temperature and wind speed.

The experts are given the freedom to add any comments that they deem important to the study. In addition, they are free to add any new factor, which they consider significant and vital to the TBM penetration operation.

3.6.4 Experts' Feedback Analysis

Four experts in the field of tunneling were interviewed to identify the factors affecting the TBM advance rate in soft ground soils. The experts interviewed were experienced foremen whose experiences ranged from 25 to 37 years in the field of tunneling in soft ground soils. Table 3.2 shows the results of the experts' responses to the different questions asked. The answers were reached consensually. Consensual answers imply high consistency among the experts' opinions. In addition, some factors which were not originally included in the forms were added by the experts.

The following points discuss the factors selected and added by the experts which are listed in Table 3.2:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 3.2: Significant Factors Based on Experts Opinions

Factor	Significance	Comments
Tunnel Properties		
Tunnel alignment	Very Significant	Curved is less productive
Soil Properties		
Soil Behavior	Very Significant	Combined properties (type, plasticity, moisture content) below and above water level
Inclusion of Boulders	Very Significant	
Contaminated Soil	Very Significant	Time lost in ventilation, testing and safety
Inclusion of cobbles	Somewhat Significant	
TBM Properties		
TBM age (in meters)	Very Significant	
TBM type	Very Significant	Use the right machine for the right soils (slurry/EPB)
TBM boring diameter	Somewhat Significant	Only when tunneling in rock
Operator's Properties		
Operator's experience (in meters)	Very Significant	How many meters excavated
SHIFT RELATED		
Shift Type (Day vs. Night)	Very Significant	Less productive in night shifts
Shift duration(8,10,12)	Somewhat Significant	12 hr shift is less productive when working 2 weeks non stop

1) Tunnel Properties: The most significant factor in this category is the tunnel alignment. It has been indicated by the experts that the TBM penetration rate is less productive at the curved sections of the tunnel.

2) Soil properties: The experts have indicated that the most important issue within the soil properties is the behavior of the soft ground soils below and above the water table level. Therefore, the classification developed by Terzaghi (1950) in his tunnelman's ground classification system which was later modified by Heuer (1974) was adopted in the model because it investigates the behavior of the soft ground soils below and above the water table. The classification is regenerated in Figure 3.10 based on Heuer's modified classification.

   The different soil behaviors are explained as (Heuer, 1974)):

   a) Firm: the tunneling can advance without initial support and the lining can be constructed before the ground starts to move.

   b) Slow raveling: chunks or flakes of material begin to drop out the arch or walls.

   c) Fast raveling: due to loosening or overstress, ground separates or breaks along distinct surfaces. In fast raveling ground, the process starts within a few minutes; otherwise the ground is slow raveling.

   d) Cohesive running: granular materials without cohesion are unstable at a slope grater than their angle of repose (+/- $30^{\circ} - 35^{\circ}$)

   e) Running: when exposed at steeper slopes, they run like granulated sugar or dune sand until the slop flattens to the angle of repose.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

f) Flowing: a mixture of soil and water flows into the tunnel like a viscous fluid. The materials can completely fill the tunnel in some cases.

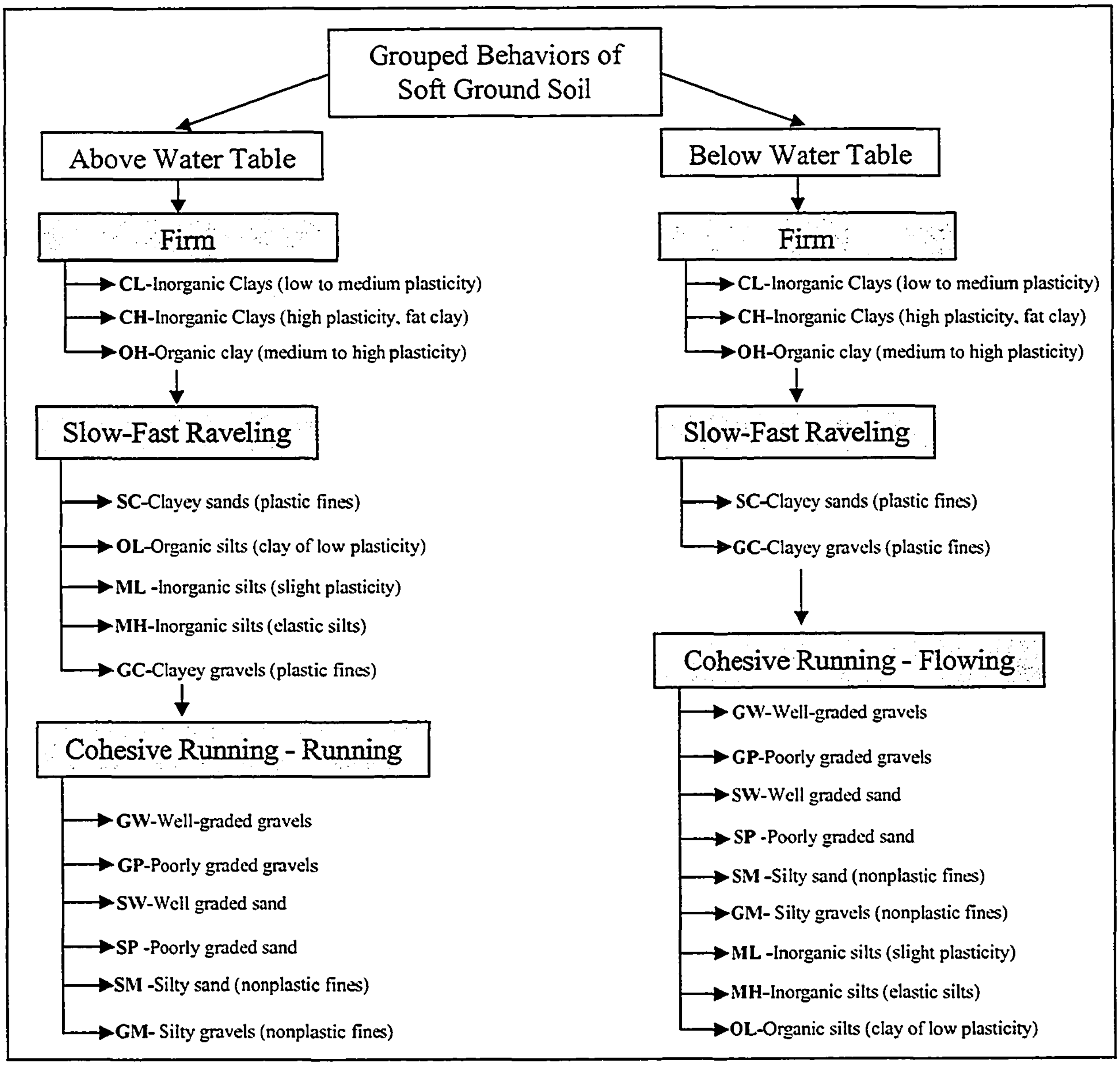


Figure 3.10 Soft Ground Soil Behaviors Above and Below Water Table Adapted from (Heuer, 1974)

3) Inclusion of boulders is considered a very significant factor that affects the TBM performance. In addition, the inclusion of cobbles, which are smaller in

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

size than boulders (Appendix A), is considered a concern but with a lower significance compared to effect that boulders make. The other very significant factor in this category is the soil contamination. When contaminated soils exist, they cause safety concerns that involve dealing with harmful gases. Therefore, dealing with contaminated soils affects the productivity of the TBM machines because a lot of time is wasted in ventilation and safety precautions.

4) TBM properties: It contains 3 subcategories which are:

    a) The TBM age which is measure by the total tunneling distance excavated by the TBM machine was considered a very significant factor. Newer machines are believed to perform better than older ones.

    b) The experts indicated that TBM type is another very significant factor which was not listed in the questionnaire. According to (Milligan, 2000), in soft ground tunneling, there are two major types of TBM machines which are the slurry machine and Earth Pressure Balancing (EPB) machine. The two types are suited for less stable ground, such as softer clays, cohesionless soil or highly fractured rocks. The main function of the two types is to support the tunnel face while excavation proceeds. In slurry machines, the tunnel face is supported by pressure from a fluid, usually either a bentonite slurry or slurry formed from water mixed with some of the excavated spoil. The excavated material is transported away from the face to the ground surface in the supporting slurry; and the spoil is then separated from the slurry so that the slurry

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

may be re-used. In EPB machines the tunnel face is supported by pressure from excavated material within the working chamber of the shield. The spoil is extracted from the pressure chamber by a screw conveyor and removed in muck wagons or by a conveyor system. According to Milligan (2000), slurry machines can be used in a wider range of soils when compared to the EPB machines. The EPB machines are limited to relatively soft and fine-grained soils (when soil additives are not used). The slurry machines, on the other hand, can be used in all soils except for those containing numerous and large rocks. Therefore, both types are assumed to be functioning with the same degree of efficiency. However, the slurry type machines will be less efficient than the EPB machines when boulders exist in the soil. This fact will be reflected in the developed model by giving less efficiency and negative impact for the slurry type machines when boulders exist.

c) The TBM diameter is considered somewhat significant only when the excavation is performed in rocks. Since we are only dealing with soft ground soils, this factor is not included in the analysis.

6) Operator's properties: The operator's experience was considered the most significant factor in this category. The operator's experience is measured by how many years of experience the operator has as a direct TBM operator.

7) Shift Related: The shift type whether it is a day or night shift was considered very significant. The TBM operator is believed to be less productive during night shifts. In addition, the duration of the shift was considered somewhat

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

significant only when the shift is a 12-hour shift that extends for two consecutive weeks or more. It was anticipated that the productivity of the operator will be less if he/she works for a 12-hour shift for extended continuous weeks.

3.7 Development of Fuzzy Expert System for Predicting TBM Penetration Rate

Based on the analysis made in Section 3.6, the factors affecting the TBM penetration rate in soft ground soils are identified. Using the same development steps introduced and explained in Sections 3.2 and 3.3, the fuzzy expert system can be now developed. The development steps are explained in the following sections.

3.7.1 Initial Model Structure

An initial model is first developed based on the information collected in Section 3.6. Figure 3.11 shows the initial model structure using a fuzzy expert system commercial package called FuzzyTECH© (FuzzyTECH, 2002). Related factors are grouped into rule blocks. A total of 5 rule blocks were generated representing the different factors listed and grouped in Table 3.2.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

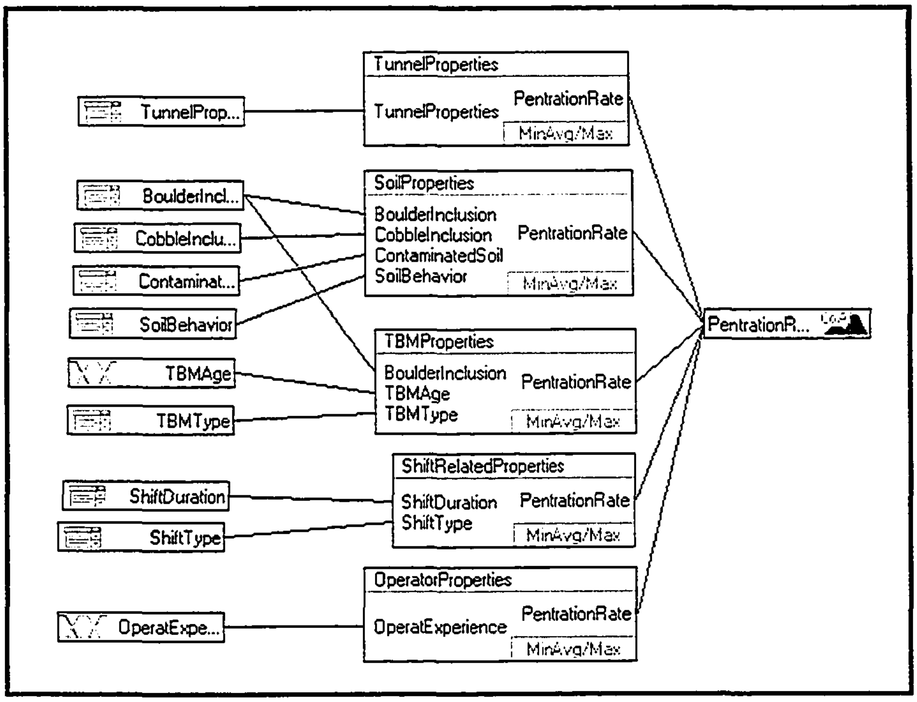


Figure 3.11 Fuzzy Expert System Initial Model

The model inputs are represented using the following input types depending on the nature of the factor:

1) Categorical inputs: Most of the model inputs are represented by this type of input. The inputs are listed as follows:

a) Tunneling alignment (1= Straight, 0 = Curved)

b) Boulder inclusion (1= No, 0 = Yes)

c) Cobble inclusion (1= No, 0 = Yes)

d) Soil contamination (1= No, 0 = Yes)

e) Soil behavior

0= Firm above water table

1= Slow-fast raveling above water table

2= cohesive running-running above water table

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

3= Firm below water table

4= slow-fast raveling below water table

5= Cohesive running-flowing below water table

f) TBM type (1= Slurry, 0 = EPB)

g) Shift duration (1= 8/10 hours, 0 = 12 hours)

h) Shift type (1= Day, 0 = Night)

2) Fuzzy inputs: Two inputs are represented by three membership functions for each. The first input is the TBM age in meters and the second is the operator's experience. Each input is represented by Low, Medium, and High linguistic terms. The two inputs are shown in Figure 3.12

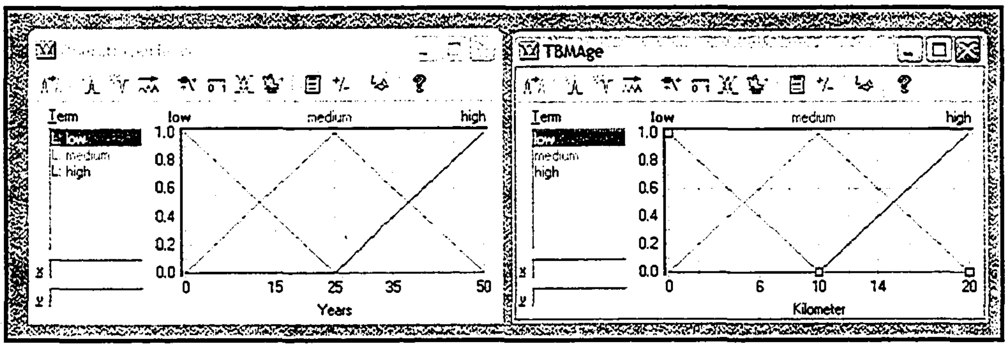


Figure 3.12 TBM Age and Operator's Experience Membership Functions

The model's main output is the TBM penetration rate in meters/hour and it is represented by five membership functions as shown in Figure 3.13.

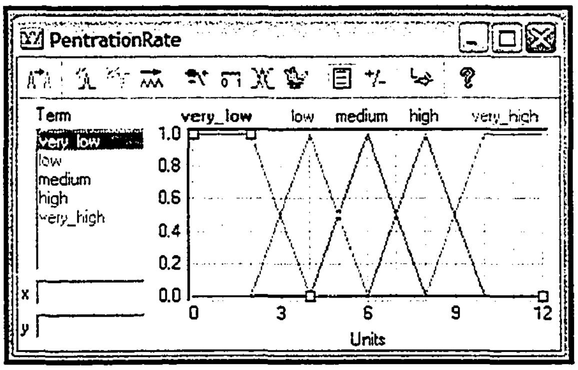


Figure 3.13 Membership Functions of the Model's Output

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

There are five rules blocks grouping the related factors together. The generation of the rules is considered one of the key components for developing the fuzzy expert system model as indicated in Section 3.2. Adopting the methodology explained in Section 3.3, the rules antecedent and consequents are generated using the knowledge gained in the experts' interview. The ranking weights of factors within the different rule blocks and their related impact on output are shown in Table 3.3.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 3.3: Factors' Ranking Weights within the Rule Blocks and Impacts

Rule Block	Block Weight	Factor	Ranking Weight	Factor State	Factor Impact
SOIL PROPERTIES	**1**	**Soil Behavior**	**0.49**	0	4
				1	2
				2	-1
				3	3
				4	-4
				5	-4
		Inclusion of Boulders	**0.12**	Yes	-3
				No	0
		Inclusion of Cobbles	**0.04**	Yes	-2
				No	0
		Contaminated Soil	**0.35**	Yes	-4
				No	0
TUNNEL PROPERTIES	**0.56**	**Tunnel Alignment**	**1.0**	Straight	0
				Curved	-2
OPERATOR'S PROPERTIES	**0.78**	**Operator's Experience**	**1.0**	High	4
				Medium	2
				Low	-1
TBM PROPERTIES	**0.56**	**TBM Type**	**0.5**	Slurry with Boulders	-3
				Slurry without Boulders	3
				EPB with Boulders	-1
				EPB without Boulders	3
		TBM Age	**0.5**	High	4
				Medium	2
				Low	-2
SHIFT RELATED PROPERTIES	0.33	**Shift Type**	**0.67**	Day	3
				Night	-2
		Shift Duration	**0.33**	8 to 10 hrs	3
				12 hrs	-1

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The complete rule base generated is shown in Table 3.4.

Table 3.4: Model's Generated IF-THEN Rules

Rule Number	Rule block: Soil Properties										
	Rule # in rule block	Factor 1 Soil Type	Factor2 Boulder	Factor 3 Cobble	Factor 4 Contaminate	Normalized CE	Rule Consequent VL	L	M	H	VH
1	1	1	y	y	y	0.58			X		
2	2	1	y	y	n	0.79				X	
3	3	1	y	n	y	0.59			X		
4	4	1	n	y	y	0.62				X	
5	5	1	y	n	n	0.79				X	
6	6	1	n	n	y	0.62				X	
7	7	1	n	y	n	0.82					X
8	8	1	n	n	n	0.83					X
9	9	2	y	y	y	0.44			X		
10	10	2	y	y	n	0.64				X	
11	11	2	y	n	y	0.44			X		
12	12	2	n	y	y	0.47			X		
13	13	2	y	n	n	0.65				X	
14	14	2	n	n	y	0.48			X		
15	15	2	n	y	n	0.68				X	
16	16	2	n	n	n	0.68				X	
17	17	3	y	y	y	0.15	X				
18	18	3	y	y	n	0.35		X			
19	19	3	y	n	y	0.15	X				
20	20	3	n	y	y	0.18	X				
21	21	3	y	n	n	0.36		X			
22	22	3	n	n	y	0.19	X				
23	23	3	n	y	n	0.39		X			
24	24	3	n	n	n	0.39		X			
25	25	4	y	y	y	0.44			X		
26	26	4	y	y	n	0.64				X	
27	27	4	y	n	y	0.44			X		
28	28	4	n	y	y	0.47			X		
29	29	4	y	n	n	0.65				X	
30	30	4	n	n	y	0.48			X		
31	31	4	n	y	n	0.68				X	
32	32	4	n	n	n	0.68				X	
33	33	5	y	y	y	0.00	X				
34	34	5	y	y	n	0.21		X			
35	35	5	y	n	y	0.01	X				
36	36	5	n	y	y	0.04	X				
37	37	5	y	n	n	0.21		X			
38	38	5	n	n	y	0.04	X				
39	39	5	n	y	n	0.24		X			
40	40	5	n	n	n	0.25		X			
41	41	6	y	y	y	0.00	X				
42	42	6	y	y	n	0.21		X			
43	43	6	y	n	y	0.01	X				
44	44	6	n	y	y	0.04	X				
45	45	6	y	n	n	0.21		X			
46	46	6	n	n	y	0.04	X				
47	47	6	n	y	n	0.24		X			
48	48	6	n	n	n	0.25		X			
\	Rule block: TBM Properties										
\	Rule	Factor 1 Age	Factor2 Type	Factor 3 \	Factor 4 \	Normalized CE	Rule Consequent VL	L	M	H	VH
49	1	L	EPB	n	\	0.83					X
50	2	M	EPB	n	\	0.70				X	
51	3	H	EPB	n	\	0.58			X		
52	4	L	Slurry	n	\	0.83					X
53	5	M	Slurry	n	\	0.70				X	
54	6	H	Slurry	n	\	0.58			X		
55	7	L	EPB	n	\	0.66				X	
56	8	M	EPB	n	\	0.54			X		
57	9	H	EPB	n	\	0.42			X		
58	10	L	Slurry	n	\	0.58			X		
59	11	M	Slurry	n	\	0.46			X		
60	12	H	Slurry	n	\	0.33		X			
\	Rule block: Shift										
\	Rule	Factor 1 Age	Factor2 Type	Factor 3 \	Factor 4 \	Normalized CE	Rule Consequent VL	L	M	H	VH
62	1	Day	8 to 10	\	\	0.70				X	
63	2	Day	12	\	\	0.66				X	
64	3	Night	8 to 10	\	\	0.51			X		
65	4	Night	12	\	\	0.46			X		
\	Rule block: Tunnel Properties										
\	Rule	Factor 1 Alignment	Factor2 \	Factor 3 \	Factor 4 \	Normalized CE	Rule Consequent VL	L	M	H	VH
66	1	Straigt	\	\	\	0.87					X
67	2	Curved	\	\	\	0.29		X			
\	Rule block: Operator										
\	Rule	Factor 1 Experience	Factor2 \	Factor 3 \	Factor 4 \	Normalized CE	Rule Consequent VL	L	M	H	VH
68	1	High	\	\	\	1.00					X
69	2	Medium	\	\	\	0.77				X	
70	3	Low	\	\	\	0.42			X		

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The initial model assumptions are chosen as follows:

1) Input aggregation is min.
2) Output aggregation is max.
3) Fuzzification method is center of area method (COA).
4) It is assumed that the experts are confident about the rules they provided and they are 100% certain about the information provided. The degrees of support are interpreted as the weight of each rule and its contribution to the output. Two degrees of support are tested. The first degrees of support are set to 1 for all rule blocks which means that all rules contribute by the same weight to the output. The second degrees of support are set to the weights generated in Table 3.4 which means that some rule blocks will have more weight or contribution to the final output. Both cases are tested in section 4.5.2 to study which one is able to capture most of the variation within the rule base.

3.7.2 System Stability Testing

After the model is built, the system needs to be tested by checking its stability. One way of verifying the system's stability is to study the behavior of the system under different model parameters (i.e. input aggregation parameter) and different input variations. This verifying methodology will show whether each of the factors is behaving as expected or not and will help select the model parameters under which the system is behaving reasonably. The methodology is explained as follows:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

3) A base case (case # 1) is selected by setting all the factors at their best possible status (i.e. shift type = Day and TBM age = 0)

4) 27 scenarios are created by changing the status of each factor at a time at different increments

5) Factors in scenarios 22-27 are all set to the lowest possible status (i.e. shift type = Night). Soil Behaviors in these scenarios are changed from Soil 1 to 6.

6) Different model parameter combinations are tested. Input aggregations, output aggregation, and fuzzification method are all changed.

7) Two different degrees of support are tested. The first degrees of support are set to the weights of the rules blocks. The second degrees of support are all set to one.

Table 3.5 shows the different scenarios tested and the system outputs for each scenario. Figures 3.14 and 3.15 show a graphical representations of all the scenarios tested for the two cases of degrees of support.

The following points are observed when studying the results of the scenarios:

1) Some scenarios were not able to capture the input variation and generated relatively constant output such as:

   a. Max-Max-COA

   b. Min/Max (0.5)-Max-COA

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 3.5: Different Scenarios Tested for the Fuzzy Expert System Model Developed

Variation number #			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
Input Variations	Soil Behavior		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	6	5	4	3	2	1
	Boulders Inclusion		NO	Yes	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	YES	YES	YES	YES	YES	YES
	Cobbles Inclusion		NO	NO	YES	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	YES	YES	YES	YES	YES	YES
	Contaminated Soil		NO	NO	NO	YES	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	YES	YES	YES	YES	YES	YES
	TBM Age		0	0	0	0	0	0	0	0	5	10	15	20	0	0	0	0	0	0	0	0	0	20	20	20	20	20	20
	TBM Type		EPB	EPB	EPB	EPB	EPB	EPB	EPB	Slurry	EPB	EPB	EPB	EPB	EPB	EPB	EPB	EPB	EPB	EPB	EPB	EPB	EPB	Slurry	Slurry	Slurry	Slurry	Slurry	Slurry
	Shift Type		DAY	DAY	DAY	DAY	DAY	NIGHT	NIGHT	DAY	DAY	DAY	DAY	DAY	DAY	DAY	DAY	DAY	DAY	DAY	DAY	DAY	DAY	NIGHT	NIGHT	NIGHT	NIGHT	NIGHT	NIGHT
	Shift Duration		8 or 10 Hrs	8 or 10 Hrs	8 or 10 Hrs	8 or 10 Hrs	12 hrs	8 or 10 Hrs	8 or 10 Hrs	8 or 10 Hrs	8 or 10 Hrs	8 or 10 Hrs	8 or 10 Hrs	8 or 10 Hrs	8 or 10 Hrs	8 or 10 Hrs	8 or 10 Hrs	8 or 10 Hrs	8 or 10 Hrs	8 or 10 Hrs	8 or 10 Hrs	8 or 10 Hrs	8 or 10 Hrs	12 Hours	12 Hours	12 Hours	12 Hours	12 Hours	12 Hours
	Tunnel Alignment		Straight	Straight	Straight	Straight	Straight	Straight	Curved	Straight	Straight	Straight	Straight	Straight	Straight	Straight	Straight	Straight	Straight	Straight	Straight	Straight	Straight	Curved	Curved	Curved	Curved	Curved	Curved
	Experience		50	50	50	50	50	50	50	50	50	50	50	50	40	35	30	25	20	15	10	5	0	0	0	0	0	0	0
Different Weights for rule Nodes	Min-Max-COA	Advance Rate Predicted (m/hr)	10.0058	9.3192	10.0058	9.3192	10.0058	9.6466	8.5926	10.0058	10.0058	9.7778	9.4716	9.063	10.0058	9.8624	9.7252	9.6074	9.4704	9.3332	9.188	8.9552	8.836	3.2602	3.2602	5.28	3.2602	5.28	5.28
	Min/Max(.5)-Max-COA		6.8188	6.7328	6.818	6.7328	6.818	6.818	6.8414	6.818	6.888	6.907	6.818	6.8742	6.888	6.888	6.9246	6.9664	6.9246	6.888	6.818	6.85	6.8266	5.1552	5.1552	5.9648	5.1552	5.9648	5.9648
	Max-Max-COA		6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
	Avg-Max-COA		6.7142	6.537	6.7142	6.2656	6.7142	6.7142	6.7142	6.7142	6.7142	6.7142	6.7142	6.7728	6.7142	6.7142	6.7142	6.7932	6.7142	6.7142	6.7142	6.7618	6.7356	4.5918	4.5918	5.2238	4.5918	5.2238	5.5914
	Min/Avg(0.5)-Max-COA		7.7388	7.4532	7.7388	7.2526	7.7388	7.6932	7.51	7.7388	7.7388	7.7552	7.6666	7.5845	7.7388	7.7474	7.76	7.7712	7.76	7.7112	7.616	7.5384	7.4578	4.0316	4.0316	5.2402	4.0316	5.2402	5.5794
	Min-Bsum-COA		10.0058	9.4664	10.0058	9.4664	10.0058	9.6466	8.5926	10.0058	9.7382	9.5336	9.2584	9.063	9.713	9.5962	9.4948	9.4664	9.2852	9.1656	9.0458	8.926	8.836	3.524	3.524	5	3.524	5	5
	Min-Bsum-COM		10.2588	9.5	10.2588	9.5	10.2588	9.7648	8.3966	10.2588	9.854	9.5866	9.2892	8.9918	9.8236	9.6696	9.536	9.5	9.2742	9.126	8.9778	8.8296	8.6816	3.666	3.666	5	3.666	5	5
	Min-Max-COM		10.2588	9.3156	10.2588	9.3156	10.2588	9.7648	8.3966	10.2588	10.2588	9.9196	9.5146	8.9918	10.2588	10.0426	9.846	9.6842	9.509	9.3334	9.148	8.896	8.6816	3.3456	3.3456	5.28	3.3456	5.28	5.28
	Prod-Max-COA		10.0058	9.3192	10.0058	9.3192	10.0058	9.6466	8.5926	10.0058	10.0058	9.7778	9.4716	9.063	10.0058	9.8624	9.7252	9.6074	9.4704	9.3332	9.188	8.9552	8.836	3.2602	3.2602	5.28	3.2602	5.28	5.28
	Gamma(0.5)-Max-COA		10.0058	9.3192	10.0058	9.3192	10.0058	9.6466	8.5926	10.0058	10.0058	9.7778	9.4716	9.063	10.0058	9.8624	9.7252	9.6074	9.4704	9.3332	9.188	8.9552	8.836	3.2602	3.2602	5.28	3.2602	5.28	5.28
	Gamma(1)-Max-COA		10.0058	9.3192	10.0058	9.3192	10.0058	9.6466	8.5926	10.0058	10.0058	9.7778	9.4716	9.063	10.0058	9.8624	9.7252	9.6074	9.4704	9.3332	9.188	8.9552	8.836	3.2602	3.2602	5.28	3.2602	5.28	5.28
All weights are set to one	Min-Max-COA	Advance Rate Predicted (m/hr)	9.4664	9.4664	9.4664	9.4664	9.4664	8.6664	7.9044	9.4664	9.4664	9.4664	8.888	8.4758	9.4664	9.4664	9.4664	9.4664	9.2058	8.5884	8.7554	8.626	8.4758	3.524	3.524	5	3.524	5	5
	Min/Max(.5)-Max-COA		7.0194	7.0194	7.0194	7.0194	7.0194	6.7142	6.7016	7.0194	7.0194	7.0194	6.9628	6.9122	7.0194	7.0194	7.0194	7.0194	7.0194	7.0194	6.996	6.9522	6.9122	5.0876	5.0876	5.75	5.0876	5.75	5.75
	Max-Max-COA		6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
	Avg-Max-COA		6.8214	6.8068	6.8214	6.5458	6.8214	6.7016	6.6778	6.8214	6.8214	6.8214	6.793	6.7664	6.8214	6.8214	6.8214	6.8214	6.8214	6.8214	6.8214	6.7984	6.7984	4.7902	4.7902	5.2506	4.7902	5.2506	5.5314
	Min/Avg(0.5)-Max-COA		7.744	7.724	7.744	7.5336	7.744	7.379	7.188	7.744	7.744	7.744	7.5956	7.4704	7.744	7.744	7.744	7.744	7.744	7.7122	7.6232	7.543	7.4704	4.236	4.236	5.172	4.236	5.172	5.3712
	Min-Bsum-COA		9.4664	9.4664	9.4664	9.4664	9.4664	8.6664	7.9044	9.4664	9.4664	9.4664	8.888	8.4758	9.4664	9.4664	9.4664	9.4664	9.2058	8.5884	8.7554	8.626	8.4758	3.524	3.524	5	3.524	5	5
	Min-Bsum-COM		9.5	9.5	9.5	9.5	9.5	8.4958	7.6666	9.5	9.5	9.5	8.8	8.3332	9.5	9.5	9.5	9.5	9.1818	8.9166	8.6922	8.4958	8.3332	4.75	4.75	6	4.75	6	6
	Min-Max-COM		9.5	9.5	9.5	9.5	9.5	8.4958	7.6666	9.5	9.5	9.5	8.8	8.3332	9.5	9.5	9.5	9.5	9.1818	8.9166	8.6922	8.4958	8.3332	4.75	4.75	6	4.75	6	6
	Prod-Max-COA		9.4664	9.4664	9.4664	9.4664	9.4664	8.6664	7.9044	9.4664	9.4664	9.4664	8.888	8.4758	9.4664	9.4664	9.4664	9.4664	9.2058	8.5884	8.7554	8.626	8.4758	3.524	3.524	5	3.524	5	5
	Gamma(0.5)-Max-COA		9.4664	9.4664	9.4664	9.4664	9.4664	8.6664	7.9044	9.4664	9.4664	9.4664	8.888	8.4758	9.4664	9.4664	9.4664	9.4664	9.2058	8.5884	8.7554	8.626	8.4758	3.524	3.524	5	3.524	5	5
	Gamma(1)-Max-COA		9.4664	9.4664	9.4664	9.4664	9.4664	8.6664	7.9044	9.4664	9.4664	9.4664	8.888	8.4758	9.4664	9.4664	9.4664	9.4664	9.2058	8.5884	8.7554	8.626	8.4758	3.524	3.524	5	3.524	5	5

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

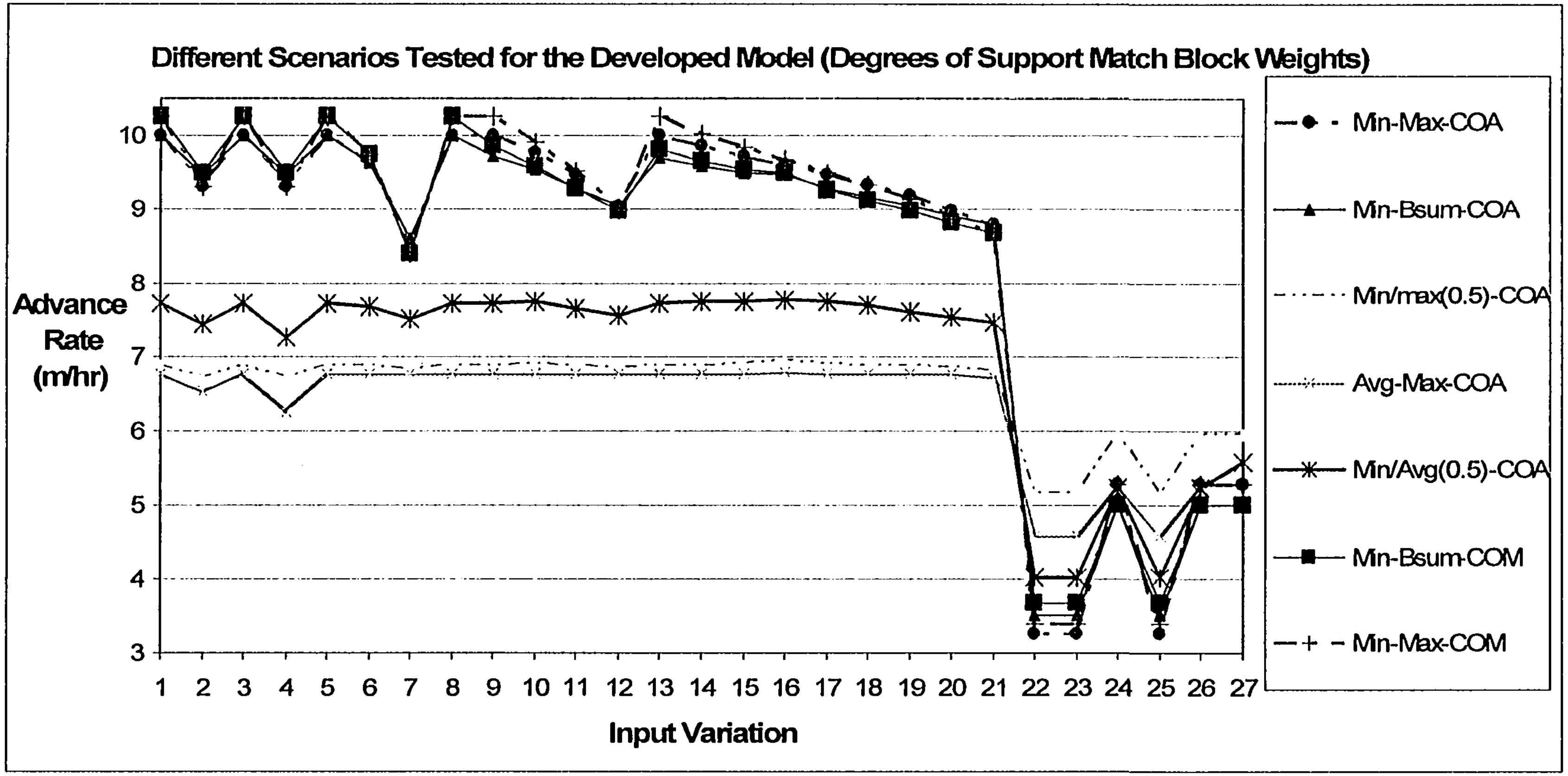


Figure 3.14 Graphical Representation of Some of the Scenarios Tested (Degrees of Support Match Block Weights)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

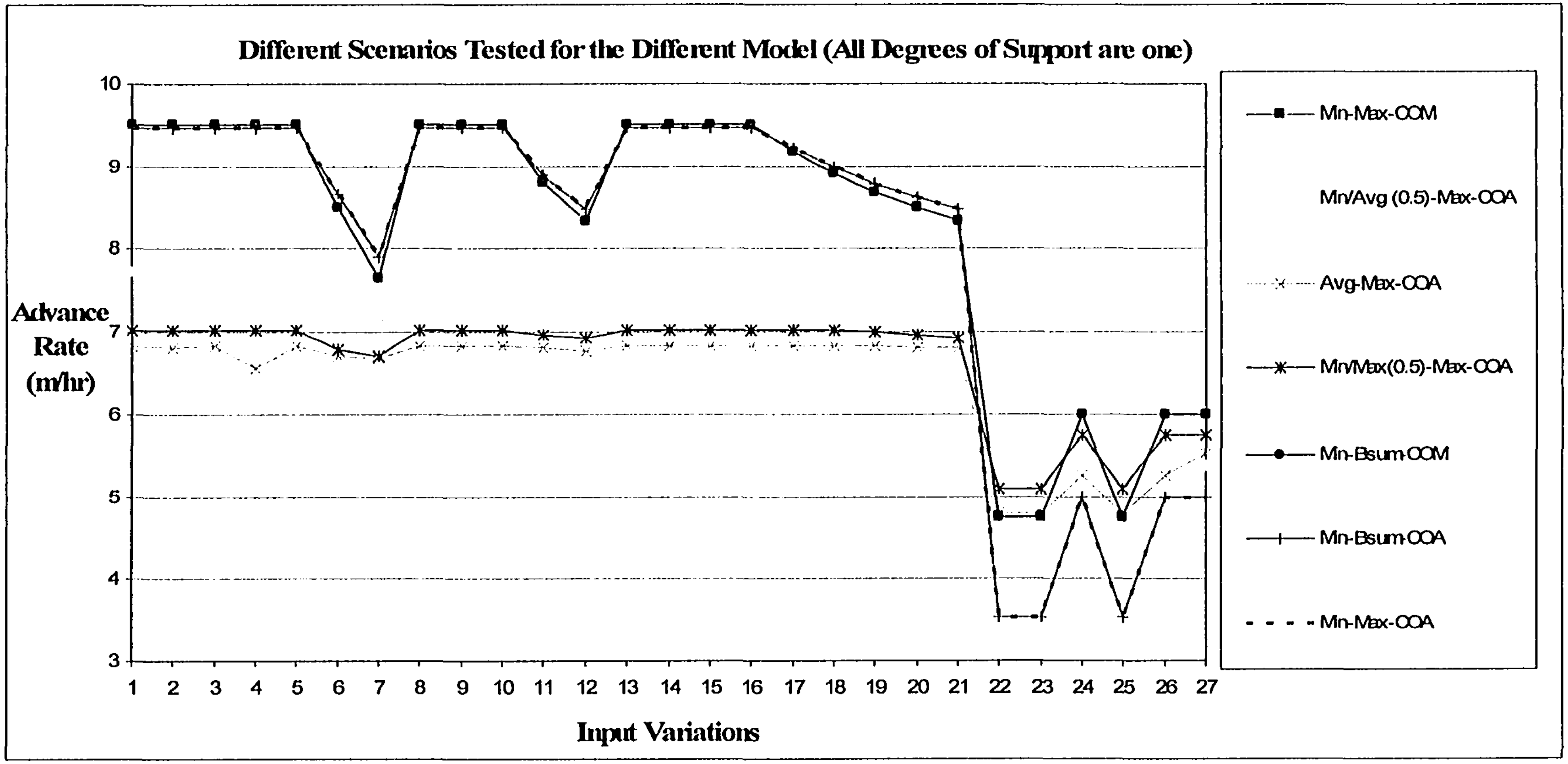


Figure 3.15 Graphical Representation of Some of the Scenarios Tested (All Degrees of Support Equal 1)

c. Average-Max-COA

d. Min/Average(0.5)-Max-COA

Therefore, these scenarios were excluded from the analysis.

2) Some scenarios were able to capture the model input variations which are:

   a. Min-Max-COA

   b. Min-Bsum-COA

   c. Min-Max-COM

   d. Prod.-Max-COA

   e. Gamma (0.5)-Max-COA

   f. Gamma (1)-Max-COA

3) Relatively similar output trends were generated for both cases of degrees of support. However, the case where the degrees of support were set to the weights of the blocks captured more input variations when compared to the case where all weights are set to 1. However, both cases were very comparable.

4) The output trends of the scenarios listed in point # 2 show that the factors generated consistent behaviors in response to the changes made.

5) For the scenarios listed in point # 2, the system is considered reasonably stable and is not highly sensitive to the degrees of support selected or the fuzzification method used.

Any of the scenarios listed in point # 2 can be selected since they all behave in a similar manner. It is important to mention that any selection made is totally based on trial and error and it can only indicate the soundness of the model parameters

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

based on the assumptions and feedback of the experts. The selection of the model is made based on the stability indicators observed when different scenarios were tested by trial and error. In order to test and check the accuracy of the model output, different methodologies are required which are discussed in the following section.

3.8 System Accuracy and Fine-Tuning

Section 3.7.2 showed how the system model can be checked for stability and consistency. As for the prediction accuracy, it can be verified by the experts' testing and feedback and by checking the soundness of the system against actual data. When system accuracy is not accepted, an optimization method can be utilized. One of the methods that has recently gained a lot of research attention is NeuroFuzzy optimization tools.

3.8.1 Testing System Accuracy

As mentioned before, testing can be done using the experts' opinion and feedback or by comparing the system outputs with the outputs of the data related to the modeled problem. Some of the data of a City of Edmonton tunneling project called NEST tunnel was used to compare the outputs. NEST tunnel data are shown in Appendix 2 representing 63 data points of the TBM penetration rate (m/hr) at different conditions. When outputs are compared, it is found out that the total average absolute error between the outputs equals 83 %. Therefore, the selected model is predicting higher rates than the actual in this specific project. By reducing the maximum limit of the "Very High" advance rate linguistic term

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

of the output shown in Figure 3.15 from 12 m/h to 10 m/h, it was found that the total average absolute error dropped to 53 %. The next step is to utilize the available NEST data to optimize the system parameters using the NeuroFuzzy optimization technique which will be introduced in the next section.

3.8.2 Optimize the Developed Model Using NeuroFuzzy Technique

NeuroFuzzy technique combines the explicit knowledge representation of fuzzy expert systems with the learning powers of adaptive neural networks that are used to fine-tune the system. The learning capabilities of the adaptive neural networks are used to fine tune and optimize the parameters of the fuzzy rule-based models in order to improve the predictive performance and reliability of the model. The NeuroFuzzy optimization capabilities have proven to be very effective in optimizing the fuzzy expert system's parameters. Detailed description of the NeuroFuzzy tools with case studies can be found in the work of Altrock (1995), Boussabaine, (2001a and 2001b) and Rutkowska (2002). In the next paragraphs, the NeuroFuzzy tools are introduced and utilized to optimize the fuzzy expert system developed for the TBM advance rate prediction.

Neural networks are adaptive networks that have the ability to learn and generalize from examples. Neural networks have many advantages that make them very good modeling tools. As discussed by Lisboa and Vellido (2000), neural networks can approximate complex non linear mapping with reasonably high accuracy. In addition, neural networks perform quite efficiently with incomplete and nosy data sets, which make them very noise tolerant. Updating

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

neural networks with fresh and new data is relatively easy which makes them very flexible in a dynamic environment.

However, neural networks still suffer some disadvantages that limit their powerful modeling capabilities. According to Lisboa and Vellido (2000), neural networks are considered "black boxes" due to their poor transparency. In addition, the neural networks require relatively large amounts of data to generate reasonable and reliable results. Another problem that neural networks may suffer is losing the ability to generalize due to over-fitting in the training stage. The previously discussed disadvantages encouraged the users to seek other modeling techniques that do not suffer from similar problems.

Therefore, when incorporating fuzzy expert systems with the neural networks in the form of NeuroFuzzy models, the generated system can offer a good modeling alternative to neural networks. NeuroFuzzy technique offers a modeling tool that is explicit and easy to interpret. In addition, the learning power of neural networks renders the fuzzy expert systems adaptive to real world data. In other words, NeuroFuzzy models are developed to combine the best of both fuzzy expert systems and neural networks.

Real world data sets related to the problem being modeled are used in the training and optimizing of the fuzzy expert system parameters. The parameters that are subject to tuning and optimizing are: the rule degree of support (rule weight) and the input and output membership functions. Figure 3.16 shows a structure of a typical NeuroFuzzy model and the parameter that can be optimized.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

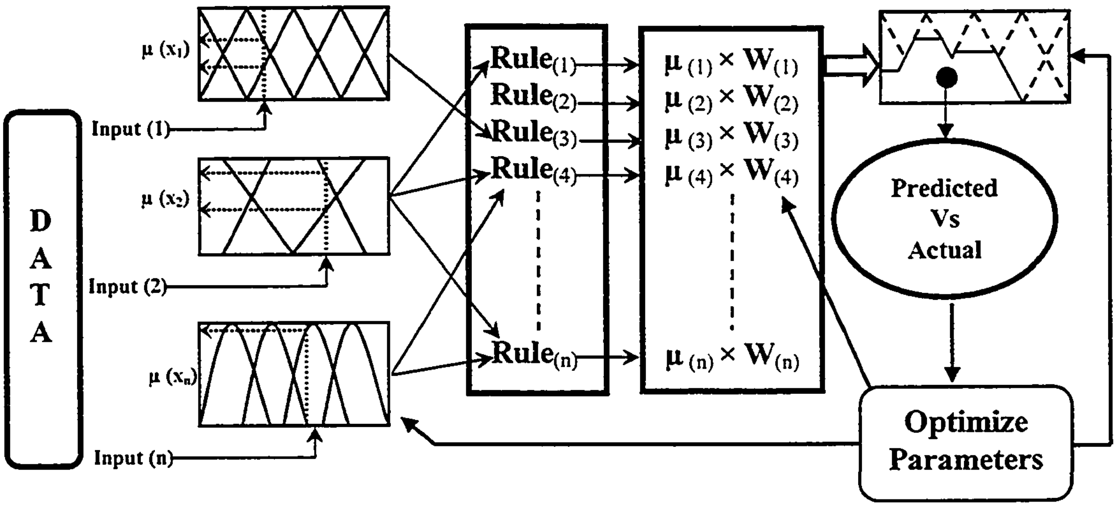


Figure 3.16 Structure of a NeuroFuzzy Model

The following paragraphs give some more details about the optimization of the model parameters.

The rule's weight or degree of support (DoS), as introduced by Kosko (1992), represents the influence and contribution of a rule to the output, for example:

IF x_1 is M_1 (AND) x_2 is M_2 THEN y is N (0<=DoS <=1)

Where DoS is the degree of support, which takes a value between 0 (no contribution) to 1 (full contribution). The (DoS) is one of the parameters that can be optimized within the fuzzy expert system The process starts by calculating the error generated by comparing the model output (predicted) with the actual output (expected). Then the system error is back-propagated using the gradient descent algorithm to tune the rule weights. The process is iterative and is terminated when the overall system error reaches a specific threshold specified by the modeler. More details about how the fine tuning process works can be found in Rutkowska (2002).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The other component in the NeuroFuzzy model that can be optimized and tuned is the parameters of the membership functions that are used to model the input and output variables. Any membership function is characterized by a set of parameters that control its shape. For example, a triangle membership function can be characterized by three parameters (center, left width, and right width) and a bell-shaped membership function is characterized by two parameters (c, a). Figure 3.17 shows the key parameters of a triangle and a Gaussian membership functions.

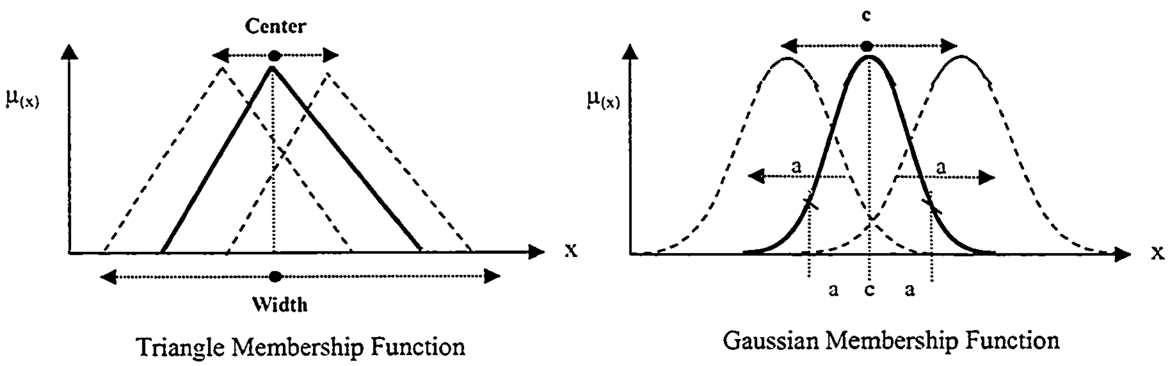


Figure 3.17 Parameter Tuning of Two Different Membership Functions

The membership parameters can be optimized using the same learning algorithm utilized in training the rules' weights. The error generated by comparing the model output (predicted) with the actual output (expected) is back-propagated to fine tune the membership parameters using the gradient descent method explained. The training process terminates when the overall output error equals a predefined threshold. Again, more details about how the optimization process is handled can be found in Rutkowska (2002).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The optimization process is an iterative operation that utilizes portion of the data as a training set and the other portion as a testing set in order to check the system behavior.

The following paragraphs show how the NeuroFuzzy techniques are utilized to fine tune the model parameters of the fuzzy expert system developed for predicting the TBM advance rate.

The first step is to prepare the data used in optimization process. Appendix 2 shows the data table used in the process. Unfortunately, the data are collected from one type of project that utilized a TBM machine of the Earth Pressure Balancing (EPB) type and shows two soil behaviors only. In order for the optimization process to be very effective and comprehensive, data that covers most of the input variables must be available. However, in this particular project, the data listed in Appendix 2 will be used to optimize the system for illustrative purposes and to show how the system is being optimized using real data.

The optimization process is summarized in the following points:

1) The first step of the optimization process is to prepare the data used in training the model and the data used in testing. As a general rule of thumb, 80% of the data is used in training and 20% will be used in testing. The total number of data used is equal to 63 data points. Therefore, 50 data points are randomly selected form the 63 data points and used to train the system and the remaining 13 points will be used to test it.
2) Different model parameters, including the scenarios generated in Section 3.7.2, are experimented with and the parameters that will generate the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

lowest average training and testing absolute error will be used as the final optimized system. A total of 27 scenarios were generated and experimented with.

3) 300 iterations were used in training the different scenarios.

4) Only the degrees of support of the rules are optimized because the fuzzy membership functions of the inputs and output require more data that cover the different linguistic terms in order to optimize their parameters. However, the data available only cover some linguistic terms which makes the fine tuning process less effective.

Table 3.6 shows the different scenarios studied and the generated results.

Table 3.6: Different Scenarios Tested and Results of Optimization

Scenario Number	Input Aggregation	Parameter	Output Aggregation	Defuzzification	Training Average Absolute Error	Testing Average Absolute Error	Average Training and Testing Error
1	Min	1	Max	COA	8.26	19.03	13.64
2	Min	1	Bsum	COA	9.53	22.97	16.25
3	Prod.	1	Max	COA	8.59	19.34	13.97
4	Gamma	0.5	Max	COA	9.01	21.42	15.21
5	Gamma	1	Max	COA	9.56	19.04	14.30
6	Min/Max	0.1	Max	COM	8.57	21.26	14.91
7	Min/Max	0.2	Max	COA	7.59	19.08	13.33
8	Min/Max	0.3	Max	COA	7.74	18.16	12.95
9	Min/Max	0.5	Max	COA	7.49	19.16	13.33
10	Min/Max	0.7	Max	COA	7.66	18.53	13.10
11	Min/Max	0.9	Max	COA	7.49	16.89	12.19
12	Min/Max	0.8	Max	COA	7.7	18.24	12.97
13	Min/Max	0.85	Max	COA	7.79	17.15	12.47
14	Min/Max	0.95	Max	COA	7.79	17.78	12.78
15	Min/Max	0.9	Bsum	COA	7.84	17.71	12.77
16	Min/Avg	0.1	Max	COA	9.83	18.75	14.29
17	Min/Avg	0.2	Max	COA	8.35	18.92	13.63
18	Min/Avg	0.5	Max	COA	7.48	19.16	13.32
19	Min/Avg	0.7	Max	COA	7.4	18.74	13.07
20	Min/Avg	0.9	Max	COA	7.43	16.57	12.00
21	Min/Avg	0.95	Max	COA	7.44	18.87	13.15
22	Min/Avg	0.85	Max	COA	7.45	19.28	13.36
23	Gamma	0.1	Max	COA	8.87	22.09	15.48
24	Gamma	0.3	Max	COA	9.28	22.65	15.97
25	Gamma	0.7	Max	COA	9.32	20.41	14.86
26	Gamma	0.9	Max	COA	9.5	20.11	14.81
27	Min/Avg	0.9	Bsum	COA	7.84	17.71	12.77

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The following points conclude the optimization process using NeuroFuzzy techniques:

1) Most of the scenarios generated very close results because the data do not cover all the input variations.

2) The scenario that had the lowest average training and testing error has the following parameters:

    a. Input aggregation → Min/Average with parameter 0.9

    b. Output aggregation → Max.

    c. Defuzzification → COA

3) The optimization process of the fuzzy expert system model for predicting the TBM advance rate is incomplete due to lack of sufficient data. When more data that cover the different input variations are available, the optimization process can be expected to yield better results.

3.9 Summary and Conclusions

The chapter has illustrated the development cycle of the fuzzy expert system for predicting the TBM advance rate. Figure 3.20 shows a flowchart that summarizes the developed cycle of a fuzzy expert system as proposed in this Chapter. Based on the analysis and discussions in Chapter 3, the following conclusions can be drawn:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

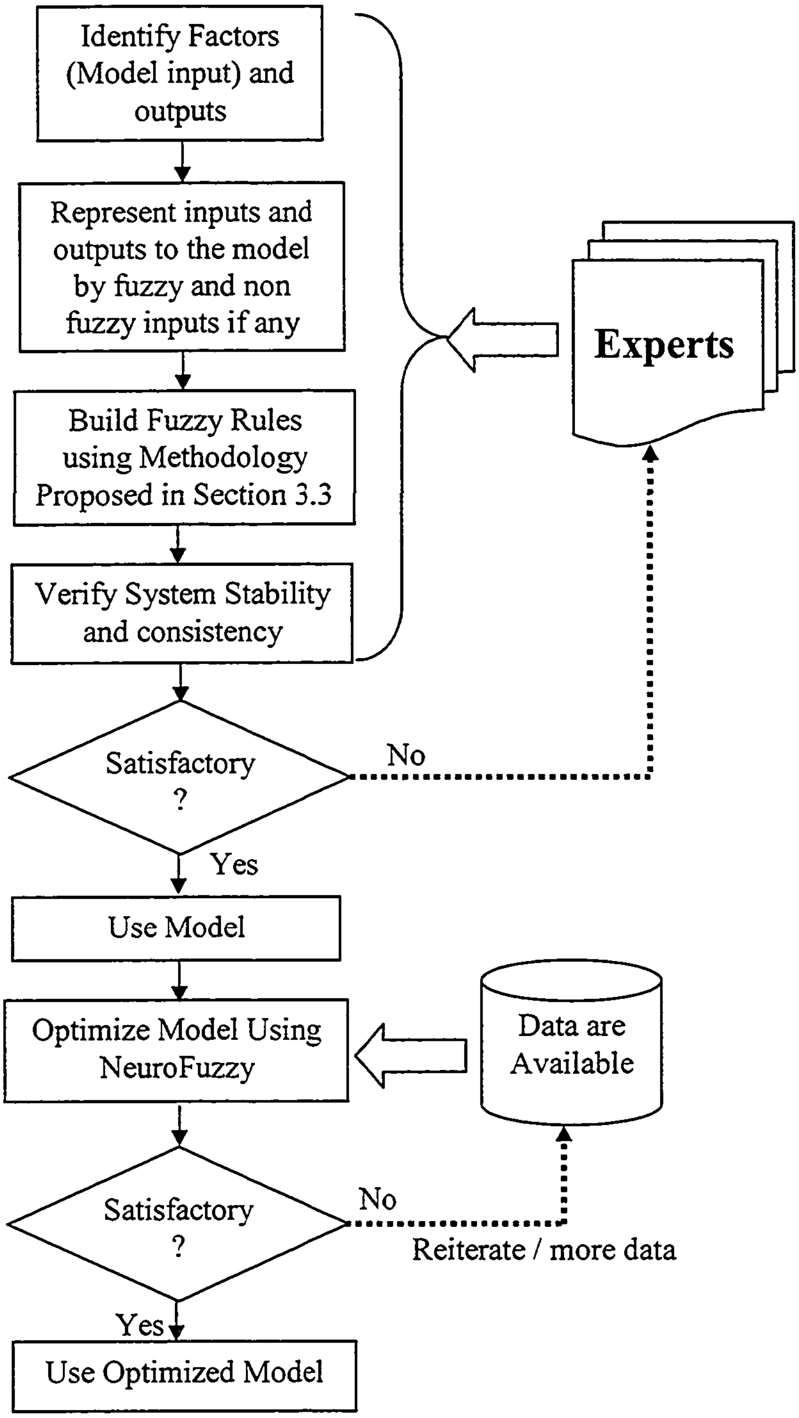


Figure 3.18: Development Flowchart of a Fuzzy Expert System

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

1) The design and development of a predictive fuzzy expert system is a cyclic and iterative process that is mainly based on the feedback and knowledge of the experts. The experience of the experts and their certainty in the knowledge they provide are the key factors for developing an effective fuzzy expert system.
2) A systematic and structured methodology was proposed to help extract the related knowledge and information from the experts to develop the fuzzy expert system. Due to the complexity and multi-attribute nature of the problems modeled, the methodology tries to minimize the current subjectivity and ad hoc manner of developing the fuzzy if-then rules by the experts by allowing the experts to evaluate and assess the importance and impact of the factors on the final output systematically.
3) The experts will also be consulted to validate the system. Optimizing the system and fine tuning its parameters for better predictive power is a very important step which can be accomplished using the optimizing powers of NeuroFuzzy technique. However, the data used in optimizing the system should cover the different input domains for more comprehensive and effective fine tuning process.

Chapter 4 will show a methodology for integrating fuzzy expert systems with discrete event simulation to improve the input modeling process in simulation. In addition, the integration is illustrated using the fuzzy expert system of the TBM advance rate developed in this Chapter.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

CHAPTER 4 –A METHODOLOGY FOR INTEGRATING FUZZY EXPERT SYSTEM AND DISCRETE EVENT SIMULATION

4.1 Introduction

This chapter will show how the fuzzy expert system modeling concept will be incorporated with discrete event simulation to model some of the simulation inputs as compared to the current practices in discrete event simulation.

4.2 Current Discrete Event Simulation Methodology

According to Halpin and Riggs (1992), simulation is an abstraction of a real-world system. It includes examining the interaction between flow units (i.e. activities), determining the idleness of productive resources, and estimating production of the system as constituted. When movements of units (i.e. activities) in simulation take place at discrete points in time which means that any activity in the system is defined in terms of its starting and end events, the system is referred to as "discrete event simulation". The duration of the activities in the simulation model are usually randomly sampled from probability distributions that represent the randomness of the system activity durations. Construction engineering projects are usually modeled using the discrete event simulation method because activities in construction engineering projects are usually defined in terms of their start and end events.

There are different simulation strategies in discrete-event simulation modeling that guide the modeling process and control it (Martinez et al. 1999). One of

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

these strategies is called "activity scanning" (AS), which is controlled by the start and end events of an activity. Halpin and Riggs (1992) explain that the way discrete event simulation works is by arranging the entire scheduled event chronologically into a "scheduled events list". Then a simulation clock (Sim_Clock) keeps track of the simulation time (Sim_Time). The Sim_Clock is advanced based on the scanning of end-event times of the activities. The process contains two main lists that organize the advancement of events: event list and chronological list. The generated event times are listed in the event list and then recorded in order on the chronological list when they occur. The last entry of the chronological list represents time now (T_Now) in the simulation and refers to an event the already took place. The event time is transferred to the chronological list and is then crossed off the event list. Finally, Sim_Time is updated with every transfer to the cumulative simulation time at this moment. Refer to Figure 4.1 for an example of the event and chronological lists.

Transfer	**Event List**			
	Activity	T_NOW	Duration	E.E.T.*
v	2	0.0	3.0	3.0
v	2	3.0	4.9	7.9
v	5	3.0	1.0	4.0
v	6	4.0	4.5	8.5

*E.E.T. = End Event Time

Chronological List	
Activity	Sim_Time
2	3.0
5	4.0
2	7.9
6	8.5

Figure 4.1: An excerpt of Event and Chronological Lists of a Project (Halpin and Riggs 1992)

As indicated by Halpin and Riggs (1992), the two major phases that control and manage the simulation of discrete event systems are the "event generation" phase and "advance" phase. Figure 4.2 shows a flowchart of the discrete event

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

simulation phases. The parts that explain the “event generation” phase and “advance” phase in the flowchart are taken from Halpin and Riggs (1992). The following paragraphs will explain the two major discrete event simulation phases.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

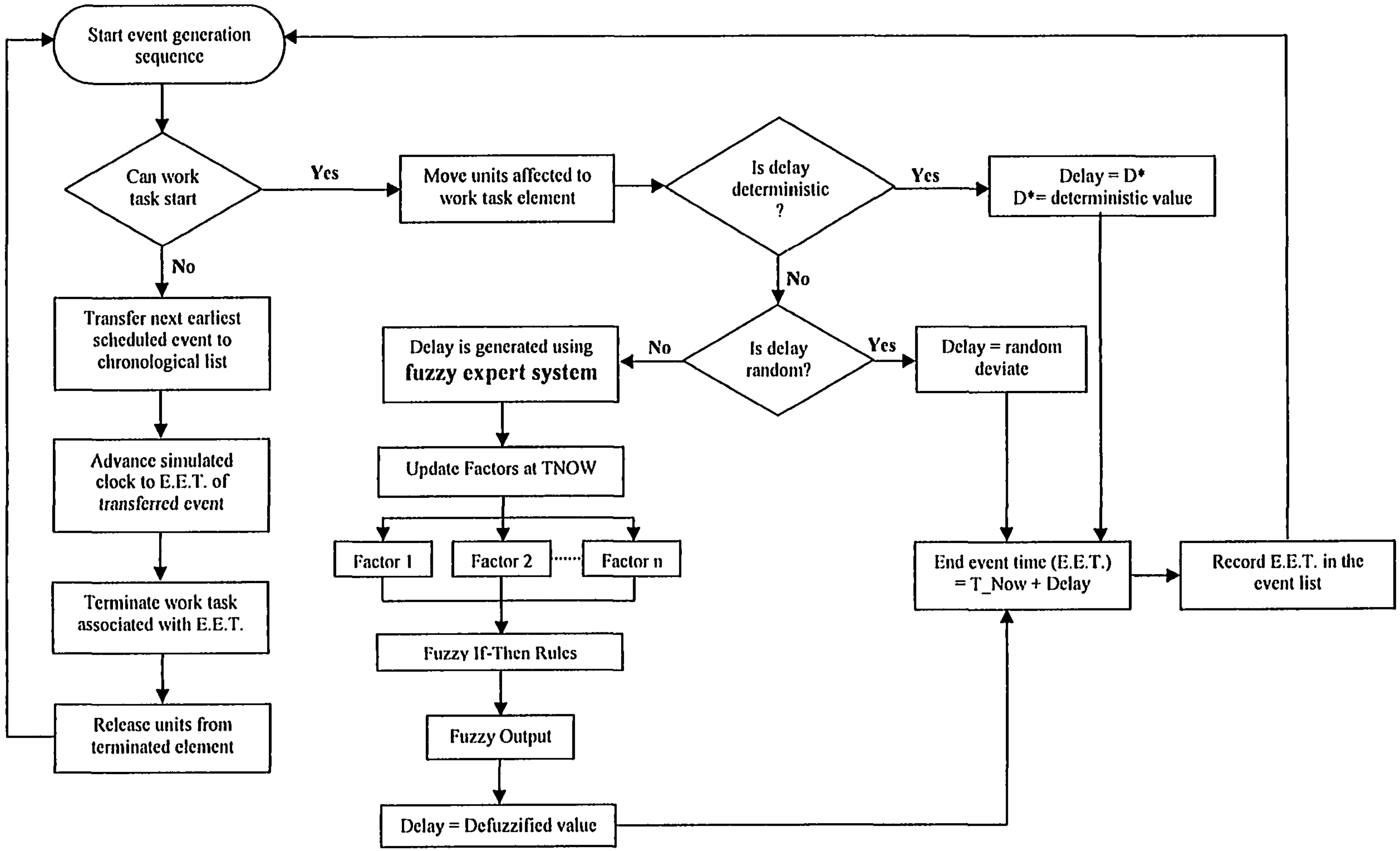


Figure 4.2: Integrated Fuzzy-Simulation Model Flow Diagram

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The "event generation" phase starts by identifying the work tasks that can start. Once a task is determined to start, the flow units that transit the work task to the graphical element representing it are moved. The next step is to generate the activity delay time.

The different possible delay time generation methods are:

1) Deterministic: Selected when the work task has a constant duration value that never changes through all the simulation runs.
2) Random: Selected when the task delay is random. A random variate is generated from the corresponding cumulative probability distribution. Figure 4.3 shows a flowchart of the procedure for generating random variates according to Halpin and Riggs (1992).
3) Fuzzy expert system: The time delay is generated using a fuzzy expert system. Section 4.2 elaborates on the fuzzy expert system method.

After the time delay is generated using one of the three methods previously introduced, the next step in the even generation phase is to calculate the event times corresponding to the termination of these work task. As shown in Figure 4.2, the end event time (E.E.T.) is calculated by adding the simulated time now (T_Now) to the event time delay. The last step in the event generation phase is to record the end event times of the work tasks in the event list.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

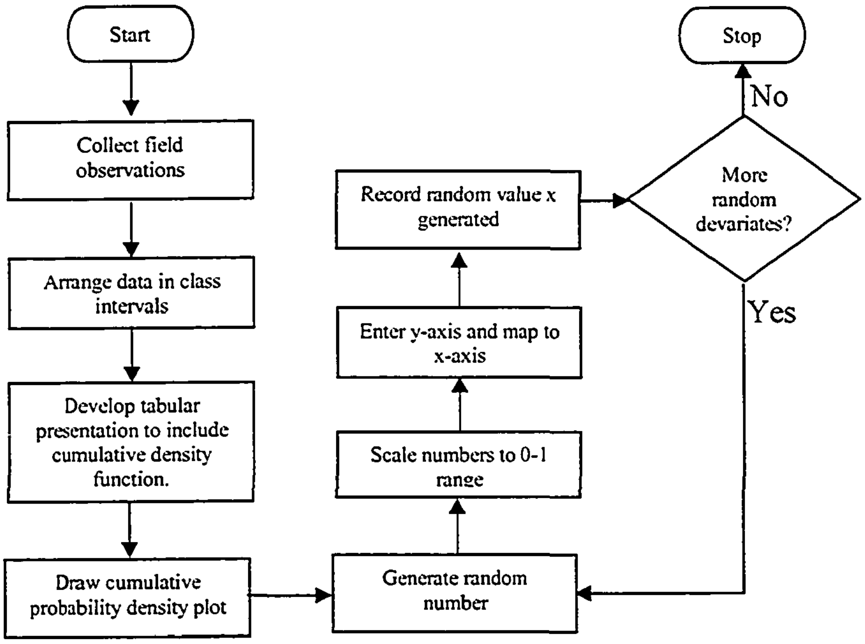


Figure 4.3 Generating Random Variates Flowchart (Halpin and Riggs, 1992)

By listing all the events that can start at T_NOW in the event list, the "advance" phase is now ready to start. By advancing the simulation clock (Sim_Clock), the next earliest scheduled event is moved from the event list to the chronological list. The simulation clock (Sim_Clock) is then advanced from its previous setting to the simulation time of the transferred event, which is (T_Now). All the activities that can be terminated when Sim_Clock is advanced are ended and corresponding units held in transits are released (i.e. resources). After the release of all units, the event generation phases start again. The simulation process continues between the two phases until a stopping criterion is applied. For example, the simulation will be terminated if it reaches a predetermined maximum simulation time or when no more units are transferred from a scheduled event (i.e. all resources used up).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

4.3 Input Modeling Using Fuzzy Expert System

As discussed in Section 4.2, input modeling using probabilistic methods will only implicitly account for the changing conditions, factors and circumstances affecting the input model (i.e. activity). Probabilistic methods are capable of handling and modeling the randomness and stochastic nature of some of these surrounding conditions and factors. Therefore, modeling methods that account for all the different surrounding conditions and changes of an input are required for better modeling process in discrete event simulation. As introduced in full detail in Chapter 3, fuzzy expert systems are capable of efficiently modeling the input accounting for the different factors and conditions affecting the process. In addition, it was shown that the fuzzy expert system is capable of handling the subjectivity and linguistic terms that describe the modeled process. The fuzzy expert system works best when enough expert knowledge describing the modeled process is available. The following paragraphs will explain how fuzzy expert systems are integrated with the discrete event simulation.

As explained in Chapter 3, when fuzzy expert system is used to model an activity input, the following steps are considered:

1) The modeled input is first studied and analyzed (identify the factors affecting the input.
2) The experts are consulted to refine and determine the best factors.
3) A fuzzy expert system is developed which can be then incorporated with the discrete event simulation model to predict the delay time of the activity.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

4) If model optimization is necessary, related data are collected to optimize the system using NeuroFuzzy tools.

5) The final optimized model is then integrated with the discrete event simulation model.

As shown in Figure 4.2, when the fuzzy expert system is selected to model the input, the first step is to update the system inputs. Figure 4.4 shows the event generation phase using fuzzy expert system.

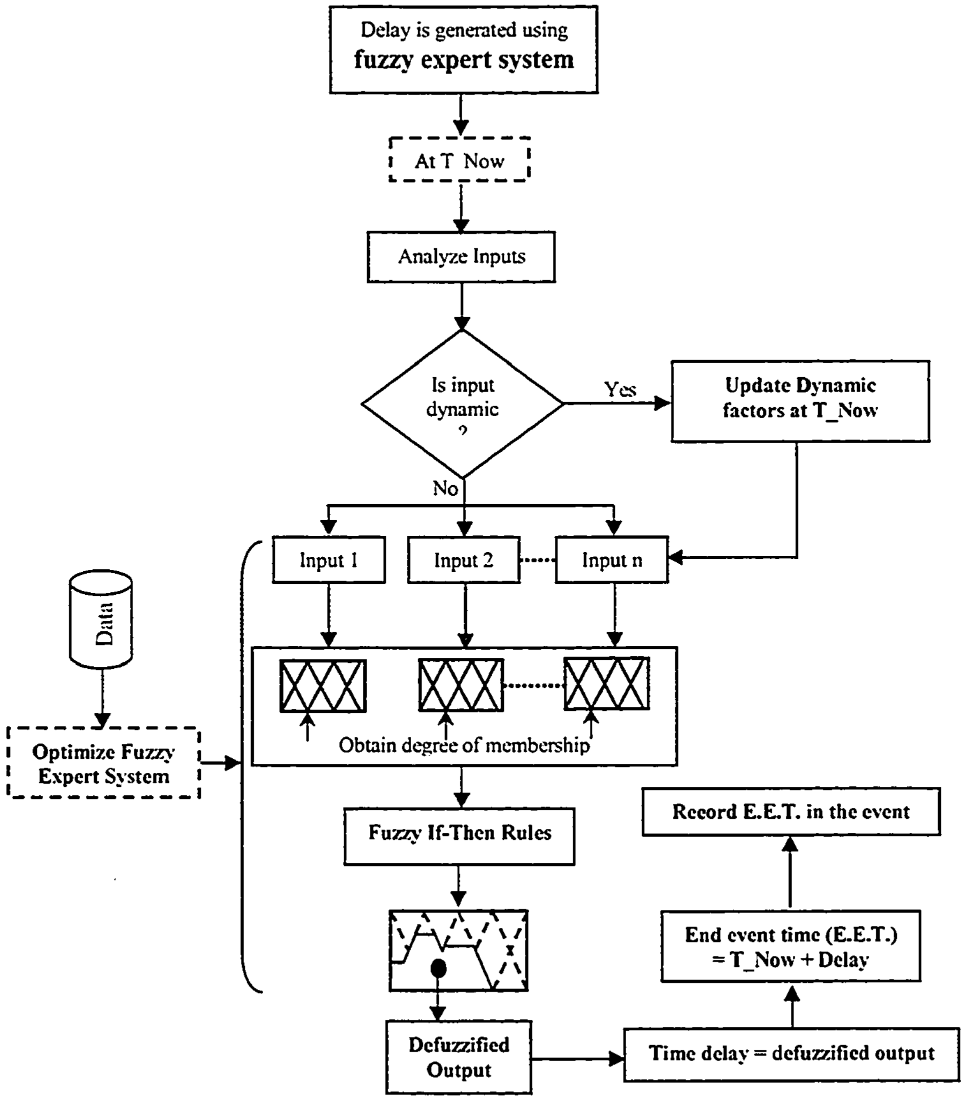


Figure 4.4: Event Generation Using Fuzzy Expert System

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The first step of the fuzzy event generation is to update the input status of the model used. Two types of inputs are identified; "static inputs" and "dynamic inputs". The first refers to the inputs that do not change in time. Regardless of being fuzzy or non-fuzzy, the "static inputs" do not change their value with simulation time. For example, it is assumed that the laborers' experience is a factor that affects the duration of an activity and that this factor is represented by fuzzy membership functions. If it is determined that the modeler will only run the simulation model using laborers that have an average of 20 years of experience, the labors' experience input will be a "static input" in this case. Therefore, the degree of membership value for that input will remain the same during the simulation time and there is no point for updating its status each time the model is invoked during the simulation runs. The modeler is free to rerun the simulation model with different average experience (i.e. 10 years). Again, because of the static nature of the factor, the new degree of membership value will remain constant during the simulation time.

On the other hand, "dynamic inputs" are the input whose values and corresponding degrees of membership are expected to change during the simulation time. A good example of a "dynamic input" is the temperature. If the average temperature, in degrees Celsius, is assumed to be a fuzzy factor that affects the duration of an activity, it is expected that the temperature will definitely change during the simulation run. Therefore, each time the fuzzy model is invoked to predict the activity duration, the temperature input is updated to find out the temperature at T_Now. The main criteria that determine the type

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

of an input, being "dynamic" or "static", are the nature of the modeled process (i.e. activity duration), the factors affecting it and the design of the simulation model. The modeler's preference and assumptions are also major criteria.

After updating the status of the "dynamic inputs", the next step is to run the fuzzy expert system model to predict the activity duration. The first stage, as shown in Figure 4.4 and explained earlier in Chapter 3, is to measure the membership degree for each input. Then the if-then rules are run and the activated rules are fired. The last stage is the output calculation. The output is first generated in the form of a fuzzy number, which will be defuzzified to a crisp output using the defuzzification method utilized. The next step in the event generation phase using fuzzy expert system is to calculate the end event time (E.E.T.) by adding the defuzzified output to the value of the (T_Now). The last step is to list the calculated (E.E.T.) of the task in the event list ready for the next phase as explained in Section 5.6. In case some data are or will be available with time, they can be used to optimize the parameters of the developed fuzzy expert system for more accurate and efficient predictive capabilities. This optimization can be accomplished using the NeuroFuzzy modeling technique, which was discussed in Chapter 3.

An example of how the inputs are handled in a typical fuzzy expert model is shown in Figure 4.5. In this hypothetical example, it is assumed that the delay time of an activity is controlled by four factors, two of which are fuzzy and the other two are non fuzzy. The two fuzzy factors are the temperature and laborers' average experience, and the non-fuzzy factors are the project location and the day

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

of the week. As indicated in Figure 4.5, the temperature and day of the week are designed as dynamic inputs, and the average experience and project location are static inputs. When the fuzzy expert system is initiated, the first step performed is to capture the model inputs at T_Now. The dynamic inputs are updated at T_Now. The day of the week and the temperature are recorded at T_Now. When the day of the week at T_Now is, for example, "Monday" , the "day of the week" input is recorded as "Monday". In addition, when the "temperature generation model" at T_Now generates, for example, a 30.5° C temperature, the "temperature" input is recorded as "30.5° C". As for the static

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

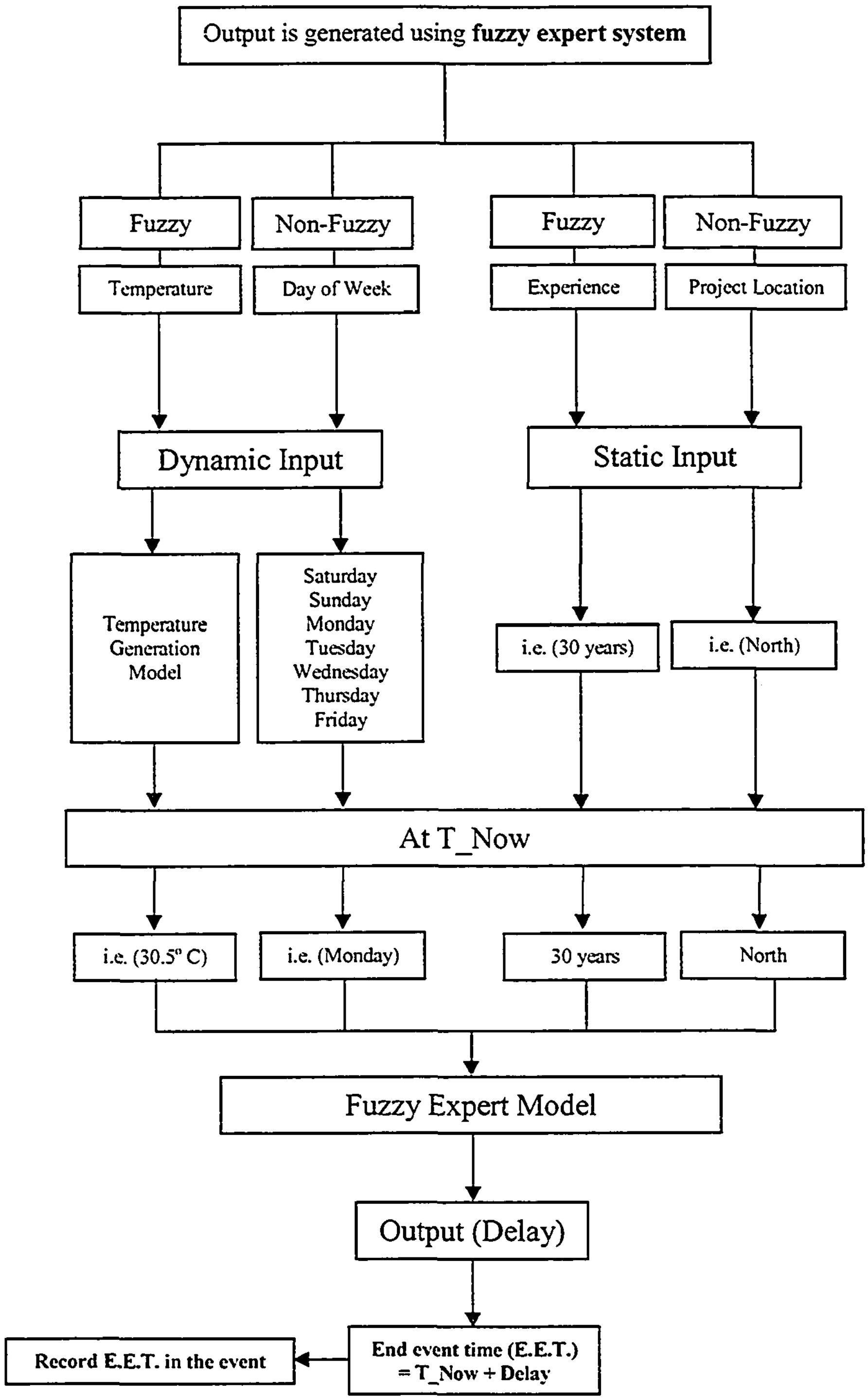


Figure 4.5: An Example of how Inputs are Handled in an Integrated Fuzzy Expert System and Discrete Event Simulation Model

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

inputs, they are delivered directly to the next step without updating. When all inputs are recorded at T_Now, the next step starts as described previously in Figure 4.5. Each time the fuzzy expert system is initiated, the dynamic input updating process continues until the simulation is terminated.

The following Sections will show how the integration between fuzzy expert systems and simulation models is performed using the TBM advance rate model developed in Chapter 3 as a case study.

4.4 Simulation Modeling of Tunneling Operations

In general, the information generated from the simulation modeling can be used in planning the tunneling projects before actual work starts. In addition, simulation modeling can be used to test different work scenarios and check their effect on the overall productivity and performance. Several researchers have studied tunneling operations in construction engineering using several simulation modeling tools. Some of the research achievements in the tunneling simulation are listed in Ruwanapura et al. (2001). For example, Touran and Asai (1998) predicted the advance rate of a small-diameter tunnel in soft rock using CYCLONE, a general purpose simulation modeling system developed by Halpin (1977). Another model using CYCLONE was developed by Tanaka (1993) to model the tunneling operation using shielded tunnel boring machines.

Simphony® is a simulation program developed to build special and general purpose simulation models (Simphony, 2000). Special purpose simulation is "a computer based environment built to enable a practitioner who is knowledgeable

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

in a given domain, but not necessarily in simulation, to model a project within that domain in a manner where symbolic representations, navigation schemes within the framework, creation of model specifications, and reporting are completed in a format native to the domain itself" (AbouRizk and Hajjar, 1998). In other words, the special purpose simulation (SPS) models are built for a specific domain and target group, which can only be used within the intended application domain (AbouRizk and Hajjar, 1998).

An example of a special purpose simulation model of a tunneling operation is the model developed by Ruwanapura et al. (2001). The developed tool is used for the analysis of tunnel construction projects. Two templates were developed: one for one-way tunneling using one TBM and the other for two-way tunneling using two TBM's. The template consists of several modeling elements. For example, the main tunnel parent element contains the general tunnel information such as tunnel length, TBM type, shift length and cost information on i.e. labor and equipment. The muck car element contains information about the number of trains and muck cars used and their capacities. In addition, the TBM element includes information on the TBM diameter and liner installation time. The information that the templates generate are the tunnel advance rate, productivity, cost, schedule and resource utilization based on the simulation analysis.

The second concept of simulation is the general purpose simulation. In Simphony®, common template is a general purpose simulation tool that enables the user to model a system using process interaction concepts (Simphony, 2000). The modeling elements in the general purpose simulation common template are

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

tools that can be used to model any type of problem which require the user to have background in simulation techniques. Some of the elements included in the common templates are entity creation and routing, recourses, statistics, activities, and tracing (Simphony, 2000). These modeling elements and many others can be used in conjunction with elements from other templates to add certain behaviors to the developed model. An example of a tunneling simulation model using common template was created by Mohamed and AbouRizk (2001). Figure 4.6 shows some snapshots of the major modeling elements of the tunneling template developed suing the common template in Simphony®. The major modeling elements in the template, as shown in Figure 4.6, are the shaft/undercut, the soil segment, and the removal shaft elements. The shaft/undercut element includes the operations that take place at the shaft and undercut locations. For example, the dirt hoisting using the crane, the dirt dumping, the material hoisting, and the train switching tracks at the undercut area are some of the operations modeled in the shaft/undercut element. In addition, the soil segment element includes the major tunneling operations taking place at the tunnel face. The excavation operation, the lining of concrete segments, the dirt removal and TBM resetting are some of the operations modeled in the soil segment element. The last element, removal shaft, is considered a minor element because it only models the TBM removal operation when the major tunneling work is terminated.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

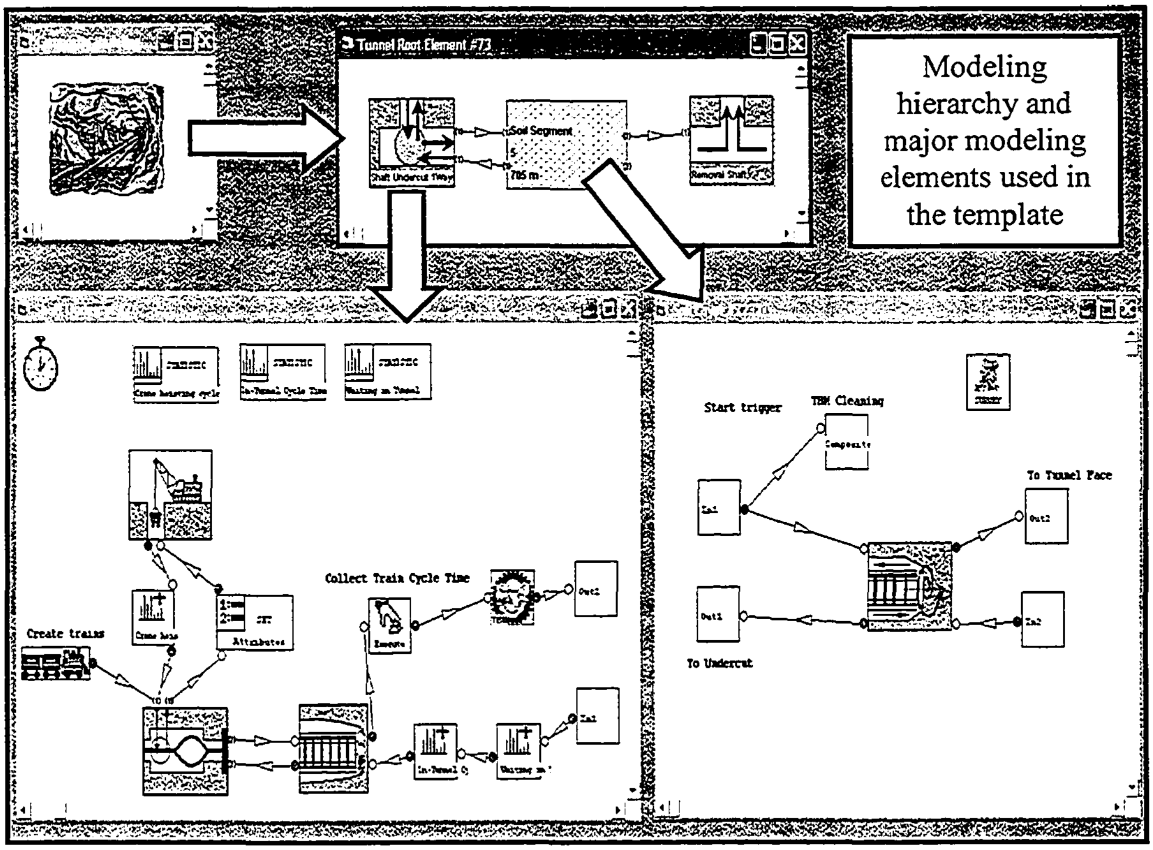


Figure 4.6: Snapshots of the Tunneling Model Using Common Template (Mohamed and AbouRizk S 2001)

AbouRizk et al. (2004) used the developed model in the analysis of one the major tunneling projects in the City of Edmonton. In their model, they investigated different modeling scenarios. One-train and two-train setups and 8-hour and 10-hour shifts were studied. Other scenarios included minimizing some of the delay times of some of the activities such as dirt removal and surveying. The overall tunnel advance rate and the total project durations are some of the important information generated from the tunneling model.

One of the major differences between the special purpose simulation tunneling model developed by Ruwanapura et al. (2001) and the tunnel template developed

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

using the common template modeling tools is the flexibility in adding new modeling constructs to the system. In both systems, adding new modeling elements require a good simulation backgrounds and skills. However, adding new modeling elements using the common template will not usually require changing the way the other elements are programmed or structured. In other words, adding new modeling elements in the common template will not require major editing or programming work. On the other hand, the special purpose simulation models require more changes and editing in order to accommodate any new modeling elements. A good example showing the flexibility of the common template modeling environment is discussed in Section 4.5.

4.5 Common Template Capabilities in Handling Several Modeling Scenarios

One of the advantages of the common template is its capabilities and flexibility in handling several modeling scenarios. In the common template modeling environment, several modeling scenarios can be created and investigated. This advantage helps the modeler investigate how effective and productive some scenarios are. As explained in Section 3.4, the tunneling operation is performed through a set of repetitive activities that involves major components such as excavation of shaft and undercut, dirt removal, and tunnel supporting. Sometimes the engineer desires to change, by adding or canceling, some activities to improve on the tunneling process. One of the most effective and least costly methods to improve on the tunneling operation and increase the productivity is to model the process using discrete event simulation. The powerful modeling environment of

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the discrete event simulation introduced in Section 4.4 can help achieve the goal of enhancing the tunneling process. To illustrate the capabilities of the discrete event simulation, a tunneling scenario provided by the City of Edmonton is modeled using the common template in Simphony®. In a study made by AbouRizk et al. (2004), the authors developed a common template model in Simphony® modeling the original tunneling operation described in Section 4.4. The new scenario suggests running the tunneling operation without the undercut as to minimize the total construction cost. Consequently, the dirt removal cycle will change:

1) Two trains will be used in the dirt removing process.
2) Due to space limitations underneath the shaft, only one train is allowed to travel back and forth from the shaft to the tunnel face.
3) When the train arrives at the shaft filled with dirt, the crane will be used to hoist the muck cars up one by one until all of them are located above ground. Then the material car will be hoisted up.
4) Finally, the tugger will be hoisted up in order to reposition it towards the tunnel face.
5) The filled material car is then hoisted down.
6) The other empty muck cars of the second train will be hoisted down one by one. When they are all down, the second train is ready to travel to the tunnel face.
7) The dirt removal cycle continues from point # 3 to point # 6.

Figure 4.7 shows snapshots of the developed model elements.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

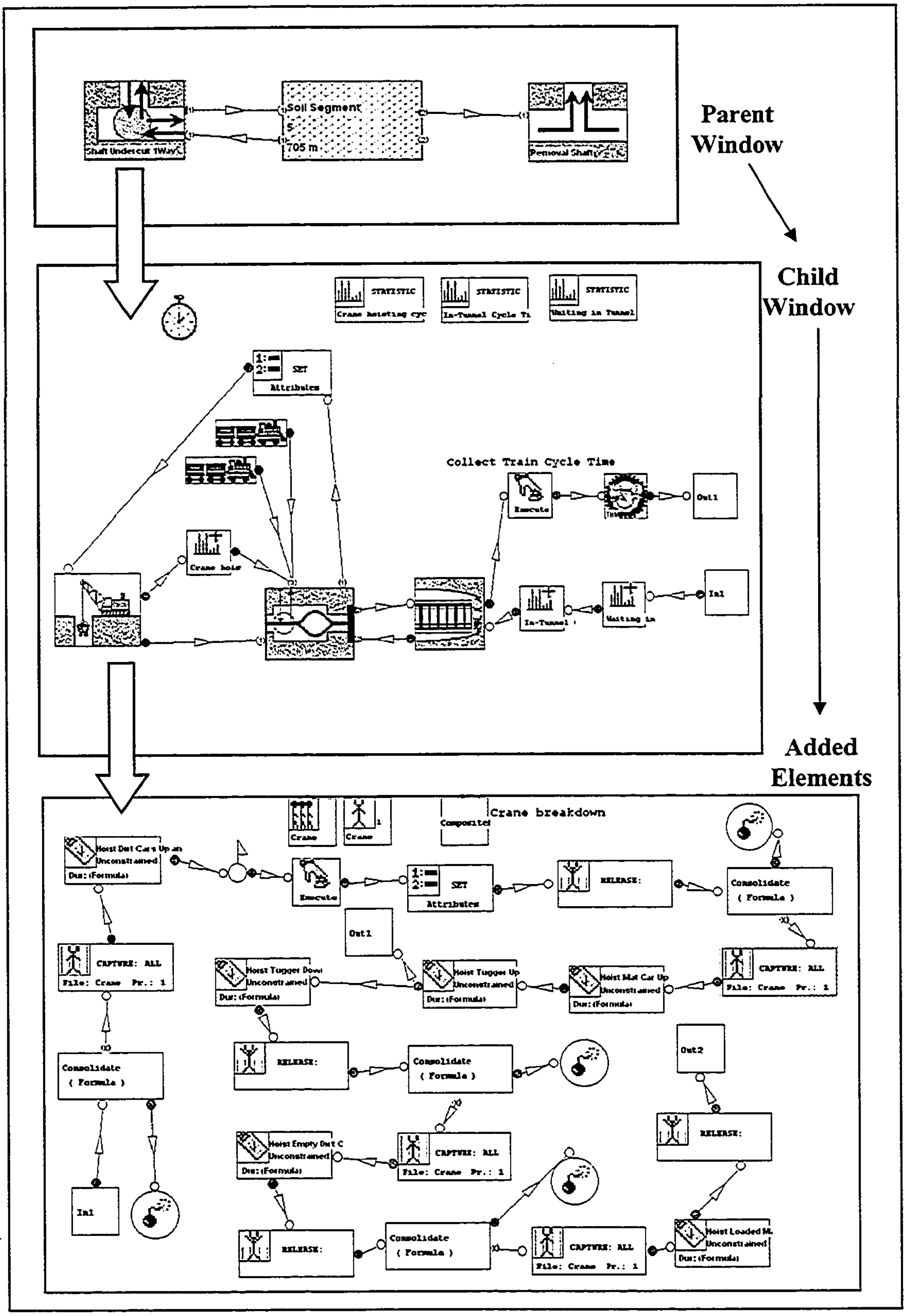


Figure 4.7 Modeling Elements Added to Model No-Undercut Scenario

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The results of the new scenario compared to the base case scenario developed by AbouRizk et al. (2004) are shown in Table 4.1. It can be noticed from Table 4.1 that the values of productivity and total duration scenario a compromise between the results of the base case scenarios.

The values of the new scenario lie between those of the base case scenario for one-train and two-train options and 8-hour and 10-hour shifts. Therefore, modeling the new scenarios revealed to the decision maker the effect of removing the undercut from the tunneling operation on the overall productivity and duration. The generated information will help the decision maker plan for future operations based in terms of time and cost.

This example shows the importance of discrete event simulation. Running models in discrete event simulation environments such as Simphony® saves time and effort. Without the aid of discrete event simulation, it would have been difficult to obtain the information listed in Table 4.1 without practically doing the work in reality to see what the outcome would be.

Table 4.1: Comparison between Base Case Model and New Scenario

Base Case Scenario (With Undercut)		
2 Trains		
Shift	Productivity (m/shift)	Project Duration (Days)
8 hrs	7.59	90
10 hrs	10.21	69
1 Train		
Shift	Productivity (m/shift)	Project Duration (Days)
8 hrs	5.38	127
10 hrs	7.45	95
New Scenario (No Undercut)		
2 Trains		
Shift	Productivity (m/shift)	Project Duration (Days)
8 hrs	6.53	108
10 hrs	8.83	80

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

4.6 The Need to Enhance the Discrete Event Simulation Capabilities

As discussed before in Section 4.4, discrete event simulation is a powerful modeling tool that models the randomness in the process under study. According to AbouRizk and Halpin (1992), modeling the random simulation inputs is considered the key factor behind the success of the simulation construction process. When a specific statistical distribution is used to model an activity, it incorporates all the elements of uncertainty in it. In addition, the conditions affecting the activity are modeled implicitly (AbouRizk and Sawhney, 1992). In case of data limitation, selecting the probability distribution that best represents the missing data is not as effective and easy. The difficulty in approximating a probability distribution is that experts do not think in probability values, but rather they think in linguistic terms such as much, very, high, etc. (Kim and Fishwick 1997). Therefore, the modeling capabilities of discrete event simulation need to be enhanced by incorporating more modeling techniques which will help model uncertainty explicitly and more effectively.

To illustrate the importance of providing more explicit and comprehensive modeling tools in discrete event simulation, the tunneling model discussed in Section 4.4 is taken as an example. One of the most important modeling parameters in the tunneling operation is the tunnel boring machine (TBM) rate or productivity. The modeler has to provide a value representing the TBM advance rate which is defined as the speed at which the TBM penetrates the different soil

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

layers encountered. In the current models, the TBM penetration rate is usually represented by a probabilistic distribution as shown in Figure 4.8.

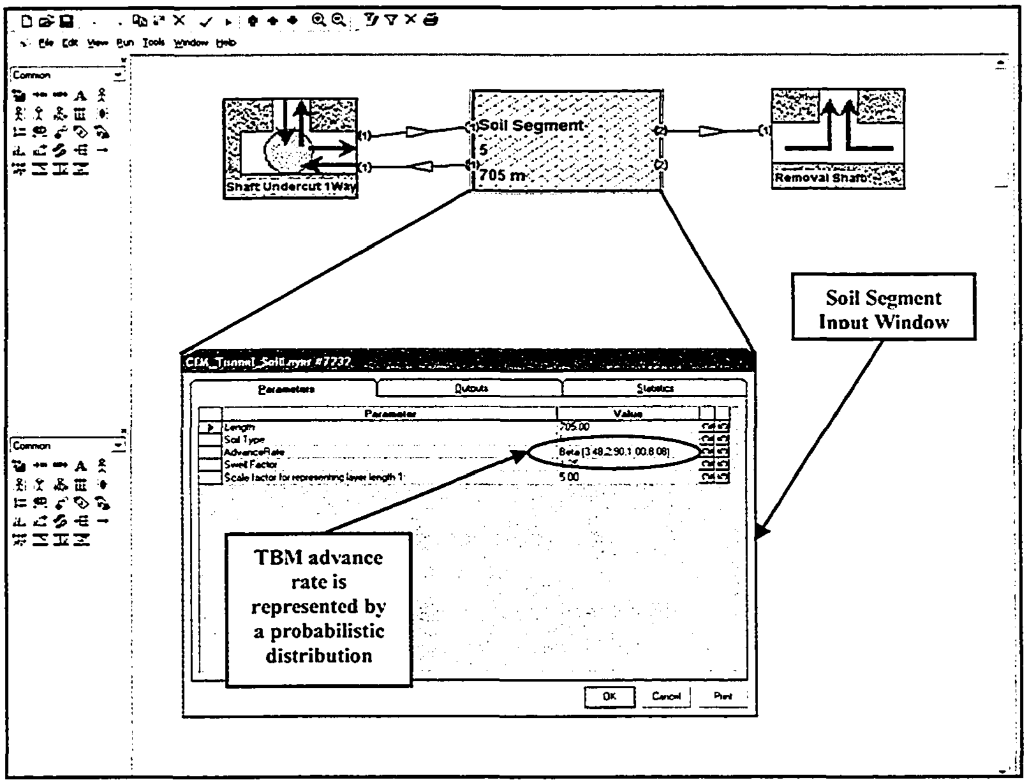


Figure 4.8: TBM Advance Rate Representation in the Current Simulation Model

The problem arises when there is not enough data to provide a more reliable estimate of the advance rate. Consequently, the modeler will subjectively estimate the probabilistic distribution of the advance rate. To illustrate the effect of selecting a specific advance rate value on the overall model output, a sensitivity analysis is made using the developed tunnel simulation model in Simphony® by AbouRizk et al. (2004). In this analysis, the model is run using different values of TBM advance rate starting from 1 meter per hour to 12 meters per hour. The final results of the analysis are shown in Figures 4.9 and 4.10.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

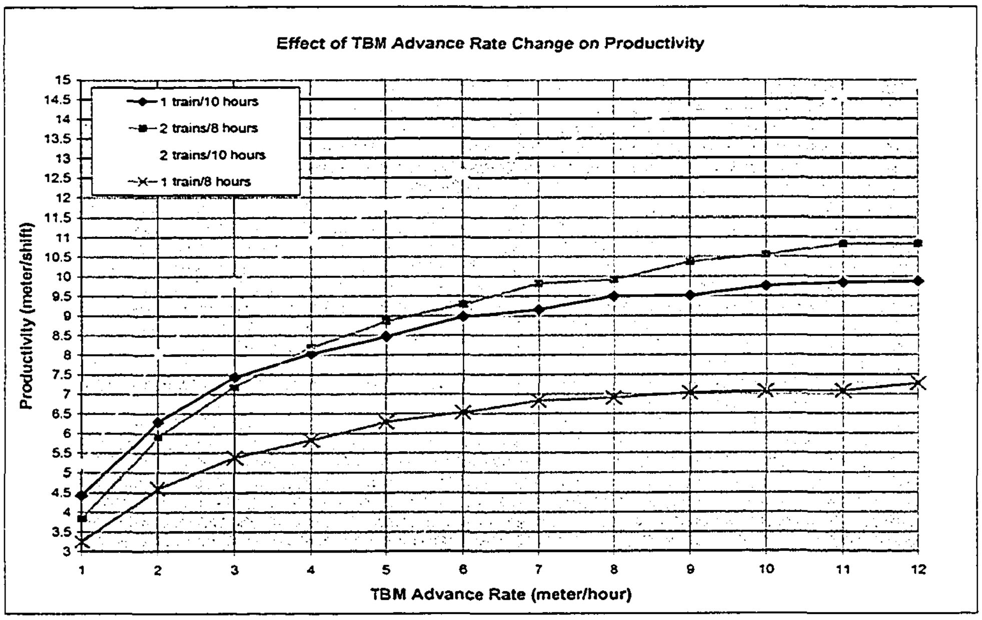


Figure 4.9: Sensitivity Analysis Productivity Outputs

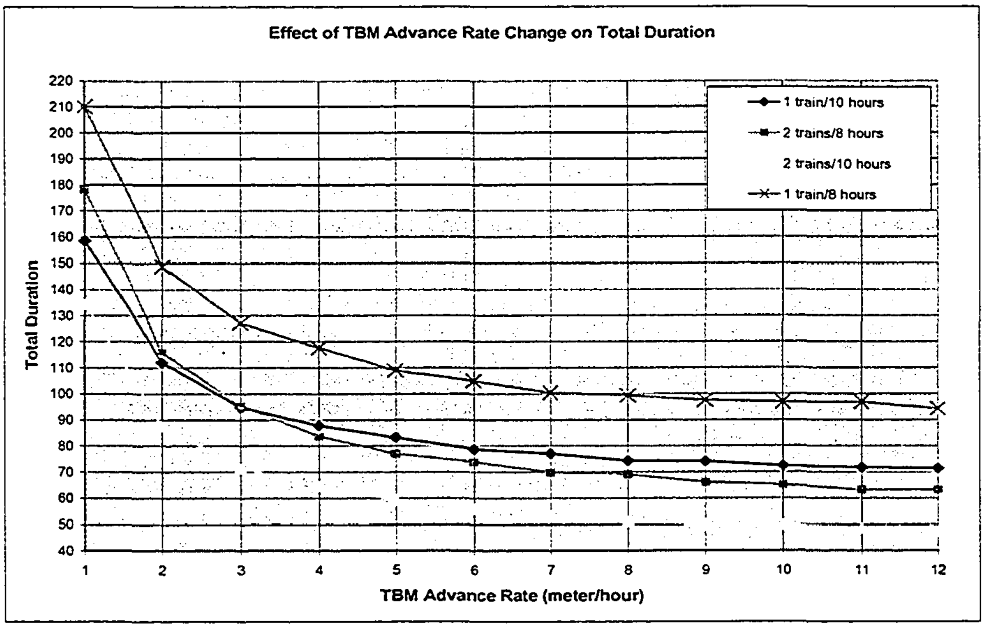


Figure 4.10: Sensitivity Analysis Duration Outputs

The curves shown in Figures 4.9 and 4.10 show the effect TBM advance rate change on the overall model productivity (meter /shift) and duration (days), respectively. It can be noticed from the generated curves that the value of TBM

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

advance rate can highly affect the overall model productivity and duration for the different scenarios investigated (one train versus two and 8-hour shift versus 10). Therefore, selecting the TBM advance rate value that best models the activity will help generate more realistic and reliable results.

Therefore, to achieve the goal of generating more reliable and realistic results in discrete event simulation, the following section will show how fuzzy expert system developed in Chapter 3 for predicting the tunnel boring machine advance rate can be and will be used to enhance the modeling process. The tunneling template developed by Mohamed and AbouRizk (2001) and analyzed by AbouRizk et al. (2004) and discussed in Section 4.4 will be used as the main research modeling reference in this chapter and the following chapters.

4.7 TBM Advance Rate Case Study

This section shows how the TBM advance rate model developed using fuzzy expert system in Chapter 4 will be integrated with the discrete event simulation tunneling model.

4.7.1 Fuzzy Expert System Inputs Representation in Tunneling Template

One of the advantages of integrating the fuzzy expert system mode is to allow the user to explicitly test different input conditions. Figure 4.11 shows the form that the user will use to input the conditions of different factors affecting the TBM advance rate.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

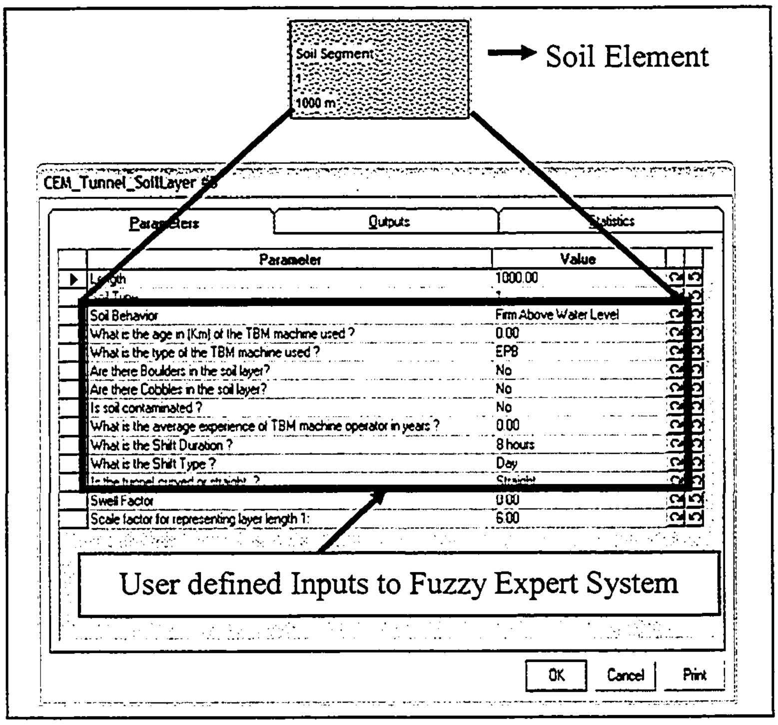


Figure 4.11: Fuzzy Expert System Input Form in Tunneling Template

The user will be able to change the conditions of the factors affecting the TBM advance rate to study the effect of these factors on the final simulation output. For the TBM advance rate model, the only input that needs to be updated as simulation time advances is the TBM age (in meter) which is a function of the total number of meters excavated.

4.7.2 Simulation Output Analysis Using Fuzzy Expert System to Predict Advance Rate

The study conducted by AbouRizk et al. (2005) on the simulation analysis of North Edmonton Sanitary Trunk (NEST) tunnel is used as a case study to show how the fuzzy expert system model is used to predict the TBM penetration rate in

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the simulation model. Table 4.2 shows the actual soil data used in the NEST tunnel.

Table 4.2: Actual Soil Data Used in Simulation Template

Soil	Length (m)	Soil Number	Penetration Rate (m/h)
Soil Segment 1	486	5	Beta (3.48,2.90,1.00,8.08)
Soil Segment 2	45	6	Triangular (2.82,5.24,8.20)
Soil Segment 3	213	7	Beta (2.89,2.41,0.90,7.97)
Soil Segment 4	132	4	Triangular (0.73,5.39,7.95)
Soil Segment 5	350	5	Beta (3.48,2.90,1.00,8.08)
Soil Segment 6	174	3	Beta (1.96,2.01,2.72,9.00)
Soil Segment 7	54	2	Beta (1.63,1.31,1.21,5.63)
Soil Segment 8	22	4	Triangular (0.73,5.39,7.95)

The shift duration adopted for the NEST project was 10 hours per shift. According to the actual data studied by AbouRizk et al. (2005), the actual construction productivity achieved for the NEST tunnel was 8.87 meters per shift and the simulation model was able to estimate productivity rate of 9.31 meters per shift (averaging 10 simulation runs). The fuzzy expert system model developed will be used to predict the TBM penetration rate as simulation time advances. The fuzzy expert system inputs that are common among all the soil elements based on the actual tunnel conditions are:

1) TBM Type = EPB
2) Shift Type = Day
3) Shift Duration = 10 hours
4) Contaminated Soil = No

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

5) Operator's experience = 12 years

6) Initial TBM age = 4.5 Km.

The TBM age will change as the excavation advances. Table 4.3 shows the inputs that are used for the different soil segments reflecting the actual conditions of the soil segments.

Table 4.3: Some of the Soil Inputs to the Fuzzy Expert System Model

Soil	Soil Behavior	Inclusion of Boulders	Inclusion of Cobbles	Tunnel Alignment	Length (m)
Soil Segment 1	0	No	Yes	Straight	84
Soil Segment 2	0	No	Yes	Curved	56
Soil Segment 3	1	Yes	No	Curved	45
Soil Segment 4	0	No	Yes	Curved	181
Soil Segment 5	0	No	Yes	Straight	378
Soil Segment 6	1	No	No	Straight	132
Soil Segment 7	0	Yes	No	Straight	433
Soil Segment 8	3	Yes	No	Straight	48
Soil Segment 9	3	No	No	Straight	124

Figure 4.12 shows the soil segments used in the simulation template.

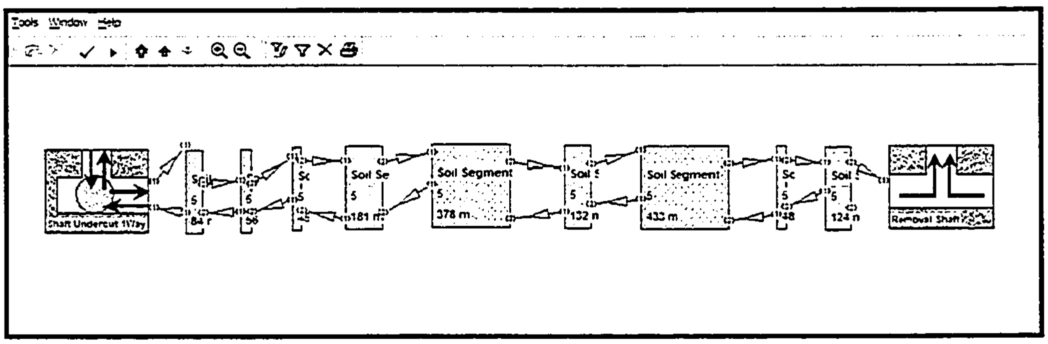


Figure 4.12: Soil Elements Used in the simulation template

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

After 10 simulation runs, the average productivity rate was 9.75 meters per shift. The estimated productivity rate is 4.7 % higher than the probabilistic approach which is relatively close. However, the productivity rate generated using the integrated fuzzy expert system and simulation model is still higher than the actual productivity rate (8.87 meter per shift). This can be attributed to the fact that the system is optimized using little data, and it needs more data to generate better results as indicated in Section 3.8. In addition, according to AbouRizk et al. (2005), the difference between the probabilistic simulation (9.31 meter /shift) and the actual one can be attributed to the combined effect of some insignificant delay factors that are not accounted for, such as muck cars and crane breakdown and TBM teeth changes. These delays can be modeled using fuzzy expert systems. However, modeling the delays is not within the scope of the study.

The comparable productivity rate that the integrated system generated indicated that the fuzzy expert system can be a very effective predictive tool within the simulation framework as long as it is well designed and optimized.

4.8 Conclusions

The integration of the fuzzy expert system predictive tool and discrete event simulation adds a lot of modeling features to the simulation modeling. Based on the analysis and discussions in Chapters 4, the following conclusions can be drawn:

1) The integration of fuzzy expert systems within the discrete event simulation framework will provide more input modeling features to the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

simulation models. The fuzzy expert system can be utilized to predict the behavior of some of the simulation elements (i.e. activity duration, productivity rates, etc.) explicitly utilizing the experts' knowledge and expertise in the problem domain.

2) Modeling the experts' knowledge and feedback within a specific problem domain while accounting for the uncertainty embedded within the knowledge provided is made possible by the powers of fuzzy expert system. Integrating the fuzzy expert system within the simulation modeling will make the input modeling process of some activities more realistic since it accounts for the different factors affecting its performance.

3) The integrated fuzzy expert system and discrete event simulation model provides an interactive modeling scheme that allows the activity duration to get updated with the change of simulation time. Adding the time dimension to the input modeling process makes it more realistic.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

CHAPTER 5 –FUZZY EXPERT SYSTEM AS A DECISION MAKING TOOL

5.1 Introduction

This Chapter discusses how fuzzy expert systems can be used as decision making tools. When fuzzy decision making tools are used in simulation models, they will add more modeling features to the simulation modeling.. Chapter 4 showed how fuzzy expert systems are used as predictive tools within the simulation frameworks. This Chapter will show how fuzzy expert systems can be used as decision making tools within the simulation framework, using the scheduling of a module assembly yard process as a case study.

5.2 Fuzzy Expert Systems as Decision Making Tools

5.2.1 The Need for Fuzzy Expert Systems in Decision Making

As indicated in Chapter 3, fuzzy expert systems were developed using the experts' knowledge and expertise in a specific problem domain. The design and development process of the predictive models is cyclic and requires the experts' involvement and verification throughout the different stages. In addition, predicting a specific phenomenon and behavior requires verification and validation in order to check whether the prediction is within the expected range. Also, the predictive model may require some optimization in order to fine-tune the parameters of the model for better and more accurate prediction. All of these requirements were fully outlined and illustrated in Chapter 3. On the other hand, using the fuzzy rule based systems for decision making purposes has different

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

requirements and limitations. In construction engineering, some problems require that the experts make a fast and reliable decision to meet a specific set of objectives. Some of the decisions required in construction engineering are complex, multi-objective, and non linear in nature. In addition, the domain within which the decision is made, is highly uncertain and adds to the complexity of the decision making process. Therefore, to help reach fast and reliable decisions within a specific problem domain, the experts' way of thinking towards the problem under study and the uncertainty surrounding this decision need to be captured and modeled. To meet all of these requirements, fuzzy expert systems can be utilized as effective decision making tools. In fuzzy expert systems, the experts' way of thinking is captured and modeled using rules. The uncertainty is accounted for using fuzzy set theory which accounts for the uncertainty embedded with the linguistic terms normally used by experts to describe some attributes such as "HIGH temperature" and "LOW productivity".

Figure 5.1 shows the characteristics of the decision making process in the domain of construction engineering.

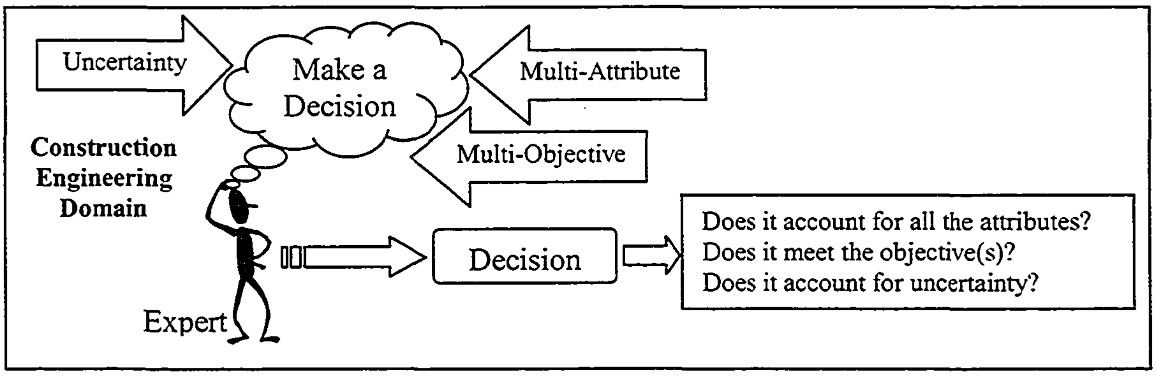
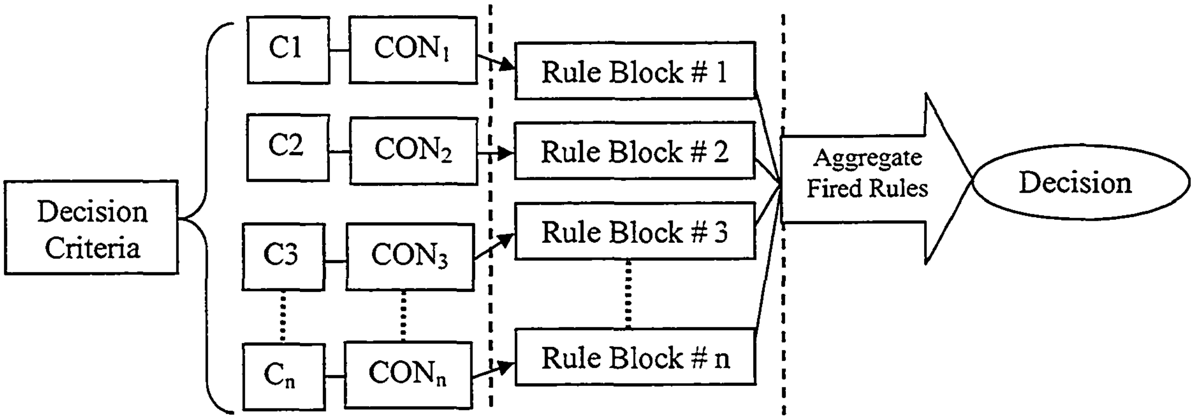


Figure 5.1: Decision Making Constraints in Construction Engineering

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

5.2.2 Structure of Fuzzy Decision Making System

Chapter 3 introduced fuzzy expert systems. The fuzzy decision making system shares the same guidelines as the fuzzy expert system but differs in the structure of the model, which is explained in the following paragraphs. Figure 5.2 shows a structure of a typical fuzzy decision making system.



C_n = decision criterion (n)
CON_n = contribution of Criterion (n)

Figure 5.2: Structure of Fuzzy Decision Making System

The system consists of the following components:

1) The first stage is to identify the different decision criteria that contribute to the decision objective(s). In this stage, uncertainty is accounted for by representing the criteria (inputs) and the decision (output) by membership functions (fuzzification).

2) The next step is to evaluate the contribution of each of the criterion to the decision objective(s). The contribution or weight of each criterion represents the strength and effect of a specific criterion on the final decision. If the expert decides to reach a specific objective, he/she will give more weight to the criteria that they believe contribute more toward achieving this objective. In a multi-objective decision making situation,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the weights of the criteria represent the expert's preference toward realizing a specific objective by adding more weight to the criteria contributing to the desired objective to make this objective more dominant. Evaluating the contribution of the criteria is done using a scale from 0 (no contribution) to 10 (highest contribution). It is important to note that giving (0) contribution to a specific criterion means that the criterion is excluded from the decision making process. The weights or contributions are then normalized by dividing by the maximum weight.

3) The next step is to develop the rule base. The experts will determine the effect of each of the criteria being at a specific state (i.e. high or low) on the final decision. The rule base represents the core of the decision making process because it captures the way the experts think about a specific problem. If more than on decision criterion is included within a rule block, the methodology developed in Section 3.7 can be adopted to develop the rule base.

4) The final step is the output aggregation and the decision making stage. Similar to what had been explained in detail in Chapter 3 for fuzzy expert systems, the fired rules are aggregated using an output aggregation (i.e. Maximum). The final decision is made by defuzzifying the fuzzy output using a proper defuzzification method. As shown in Chapter 3, Center of Area (COA) defuzzification method was used in the predictive model developed. According to Altrock (1995), Center of Area (COA) and Center of Maximum (COM) defuzzification methods are very applicable

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

to control and decision making problems. However, COM is easier in calculations compared to the (COA) method.

The following sections will show how the fuzzy rule based systems can be effectively used as decision making tools in simulation models within the construction engineering domain.

5.2.3 Fuzzy Decision Making in Simulation Models

As mentioned before, construction engineering related problems require decisions that may be multi-criteria and complex in nature. As introduced in Chapter 4, simulation is an abstraction of a real-world system. Therefore, many decisions have to be modeled in simulation mimicking the real-world system. Chapter 4 showed how fuzzy expert systems can be integrated with the simulation framework to predict behaviors such as activity durations using the expertise and knowledge of the experts. The TBM advance rate prediction was taken as a case study. The integration of fuzzy decision making system with simulation modeling can help incorporate the experts' knowledge and way of thinking towards a specific decision within the simulation framework. The integration between the fuzzy decision making system and simulation framework will be similar to the integration methodology adopted between the fuzzy expert system and simulation, except for some minor differences. Figure 5.3 shows the integration between the fuzzy decision making system and the simulation framework.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

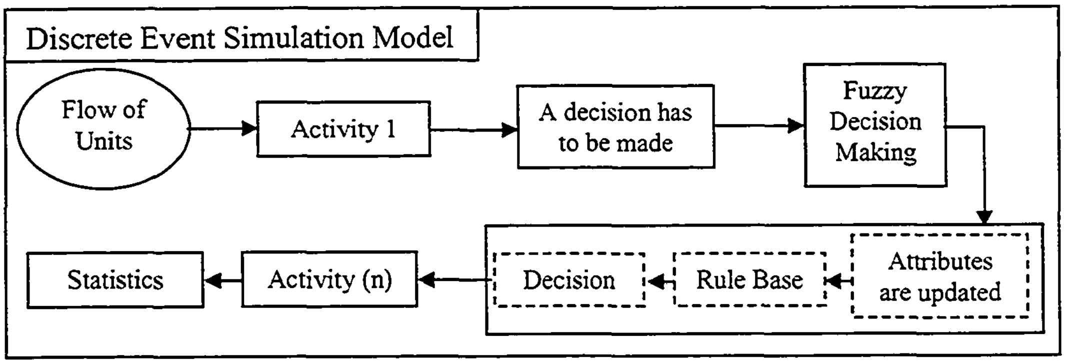


Figure 5.3: Sample Integrated Fuzzy Decision Making System and Simulation Model

When a decision is modeled using fuzzy decision making system in simulation it is run when the model is triggered. Once the model is triggered, it first checks the status of the inputs and updates them if they are "dynamic". The system activates the rule base and the fired rules are used to generate the final decision. The applications of the fuzzy decision making system within the simulation framework are many. Some of the applications can be:

1) Prioritization of simulation units. The system can be used to prioritize units competing for the same resource or waiting to be processed.
2) Studying the effect of certain actions. The system can be used to test the effect of some decisions that are usually controlled and determined by experts such as resource allocations.

One of the applications of the fuzzy decision making system that will be illustrated in this Chapter is the prioritization of module assembly yard scheduling using simulation, which is discussed in the following section.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

5.3 Decision Making in Module Assembly Yard Scheduling

5.3.1 Introduction

The fuzzy decision making system presented is integrated with the simulation model developed by Borrego (2004) to schedule module assembly yards. Borrego's simulation model is used to schedule pipe spool modules assembly operation in assembly yards. The simulation based scheduling was adopted in Borrego's model to account for the different physical and logical constraints controlling the scheduling process. The simulation based scheduling will also allow the user to experiment with certain scheduling scenarios by changing some of the control parameters. The yard is divided into different bays that are grouped in four main bay areas. The capacity of these bays differs based on the sizes of the modules assembled. Modules are classified into 5 types (cable tray, equipment, pipe rack, structural, and miscellaneous). The modules are routed to the areas based on their types (for example, all cable tray modules are routed to area A). The priority of module assembly is based on the module float (F) and it is calculated using the following equation:

[5.1] $$P\ (\text{priority}) = 500 - (F + \text{abs}(\min(F)))$$

Where F (float) = PSD –ESD-D , where PSD = planned shipment date and ESD = early start date , and D = duration.

Once routed to a specific bay, the module undergoes a set of processing activities (structure, piping, cable tray, electrical heat tracing. insulation, and fireproof). After the module is finished, it can only be shipped when the space in front of it in the bay is empty. The finished modules can wait a maximum of 5 days for

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

shipment and the total number of shipments per day is six modules only. These constraints and limitations are all captured using the simulation based scheduling developed by Borrego (2004). The main components of the models are:

1) The database that contains the attributes of the modules (type, size, durations, early start, planned shipment date and calculated priorities).
2) Simulation model using Simphony that reads the module attributes from the database, simulates the assembly process, then generates the results in a tabular form.

More details on the model and how it is developed using Simphony can be found in Borrego (2004).

In spite of the tremendous advantages that the simulation based scheduling provide including evaluations of different scheduling scenarios by controlling the modeled constraints, it still lacks the transparency and flexibility in developing the desired scenarios. The flexibility in changing and experimenting with the model constraints is limited to what the developer outlined. For example, the prioritization of the modules is a very important scheduling aspect and needs to be well accounted for to achieve the desired objective(s). The current simulation-based scheduling model only accounts for one prioritization method or dispatching rule, which is based on the least float. Therefore, a decision making system needs to be incorporated within the simulation model to provide the user with the flexibility of utilizing different dispatching rules and accomplishing a multi-objective scheduling process. The following sections will show how the fuzzy decision making system can be incorporated with the simulation model to

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

provide more transparent and flexible controlling of the module prioritization process.

5.3.2 Background on Dispatching/ Priority Rules

There are different priority and dispatching rules to control the job shop scheduling process. Some of these rules are designed to meet a specific scheduling objective. According to Kiran (1998), some of the priority rules are as follows:

1) Critical ratio (CR): it is based on the critical ratio index calculated as lead time (due date –time now) / remaining operation time;
2) Earliest due date (EDD): it processes the jobs with the earliest due date;
3) First come, first served (FCFS): priority is given to the job that arrives into particular queue earliest;
4) Shortest processing time (SPT): the job with the longest processing time is selected.

Other priority rules can be found in Kiran (1998). Each priority rule meets a specific set of objectives. The shortest processing time (SPT) tries to reduce the average flow time (completion time- release time) of the jobs. (SPT) is also good at reducing the average lateness (completion time –due date) of the jobs processed. The critical ratio (CR) and the earliest due date (EDD) rules minimize the average tardiness (number of late jobs). More details on the different objectives that different priority rules meet can be found in Kiran (1998).

According to Kiran (1998), the function categorization of the different rules can be as follows:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

1) Simple priority rules: used when one priority rule is used to prioritize all of the jobs.
2) Combinations of rules: used when two or more rules are utilized at a specific time being a function of the queue and order characteristics.
3) Weighted priority rules: involve the utilization of rules in (1) and/or (2) combined with different weights.

Extensive research work has been conducted on the priority or dispatching rules studying the different function categorization of rules. Barman (1997) studied the combination of simple priority rules to improve the flow time and tardiness in a flow shop with three work centers. Barman studied the combination of 4 priority rules at the three work centers of the flow shop generating 64 priority rule combination schemes. The rules studied are first in queue rule, shortest processing time, earliest due date, and critical ratio. Barman used the mean flow time, mean tardiness, and percentage of tardy jobs performance criteria to study the relative advantage of the rule combinations. He concluded that the combining (SPT) and (EDD) rules generated excellent results. Subramaniam et al. (2000) proposed an analytic hierarchy process (AHP) that dynamically selects a dispatching rule from several candidate rules based on job shop conditions. In their system, they used (AHP) to help select the appropriate dispatching rule to use from three candidates which are shortest processing time (SPT), most work remaining (MWKR), and most operations remaining (MOPNR). These rules are selected because they minimize the makespan of the system, which is the main objective of the study. The selection is made based on the existing job shop conditions or performance metrics, which include the availability of machines,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

number of jobs in the queue, and type of jobs remaining. With the use of AHP, the user updates the status of the performance metrics suggesting the appropriate dispatching rule based on the dominant performance.

Canbolat and Gundogar (2004) proposed a fuzzy priority rule for job shop scheduling. In their study, they proposed a fuzzy rule based system that accounts for multi-objective scheduling. In their model, they developed a fuzzy logic system that calculates priority by considering the shortest processing time (SPT), critical ratio (CR), and next machine load (NML) priority rules. The inputs and outputs are represented by fuzzy membership functions. The rule base is developed by studying the effect of fuzzy input on the final priority. The inputs are weighted to provide more importance to the criterion that meets the required objective. The developed system is very effective in satisfying multi-criteria prioritization process.

The previously discussed studies motivated the utilization of a multi-attribute decision making system using fuzzy rule based modeling in prioritizing the module assembly yard scheduling within the simulation framework. The following section explains the details of the proposed system.

5.3.3 Using Fuzzy Decision Making System in Simulation-Based Scheduling

Figure 5.4 shows the details of the proposed fuzzy decision making system within the simulation framework. The components of the model follow the structure of the model explained in Section 5.2.2, which has the following components:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

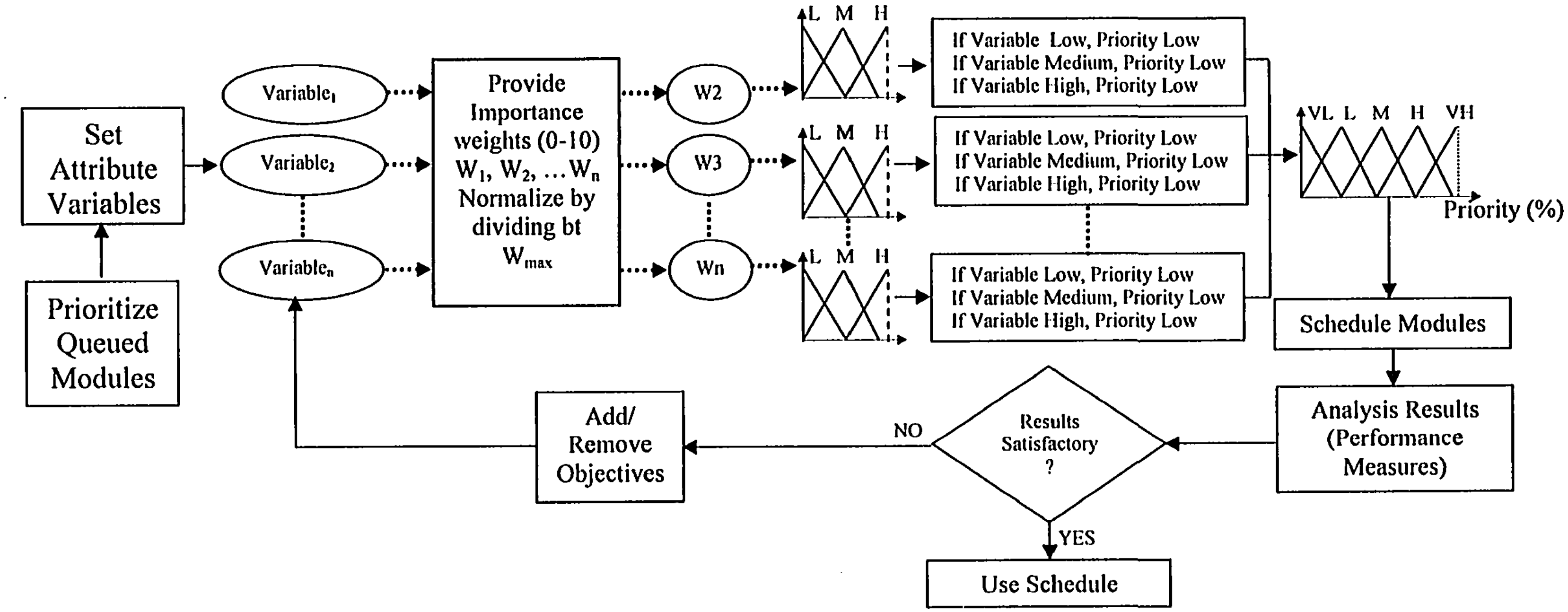


Figure 5.4: Fuzzy Decision Making Model Used in Module Assembly Yard Simulation Model

1) The user starts by adding the desired attribute/input variables that contribute to the final decision.

2) The expert provides the contribution (weight/importance) of each attribute to the final decision. The contribution scale is from 0 (no contribution) to 10 (highest contribution). Each input then has a normalized contribution after dividing by the maximum contribution.

3) The fuzzification process of input variables is done based on the nature of the variable and the user's preference. For the module yard assembly process, the inputs are represented by normalizing their values for all the modules awaiting using the following equation:

[5.2] $$(Xi)_{nomalized} = ((Xi) - (Xi)_{max}) / = ((Xi)_{max} - (Xi)_{min})$$

Where (Xi) is the value of the i^{th} input.

After normalization, the membership values of these normalized values are calculated using the membership functions shown in Figure 5.5. This fuzzification scheme is adopted because the nature of the inputs is comparative, which can be compared using fuzzified normalization.

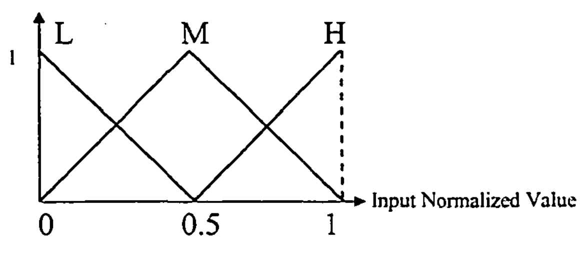


Figure 5.5: Sample Membership Function for the Fuzzy Decision Making Model

The fuzzy membership functions provide a better representation of the decision criteria because they capture the gradual membership of an input

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

to a specific state (i.e. high processing time) rather than representing it as a crisp value being either a full member or not a member to a specific state. This is how fuzzy membership functions deal with the subjectivity and fuzziness of the decision criteria. The fuzzy membership functions are more realistic representation of the inputs because they capture the way experts think toward a specific behavior.

4) The rules are developed for each input mapping the effect of the inputs on the final priority output.

5) The output is represented by the membership functions shown in Figure 5. 6.

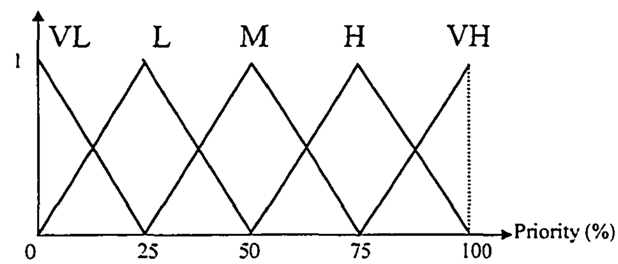


Figure 5.6: Fuzzy Membership Functions of the Fuzzy Decision Making Output

6) The output aggregation of the model is selected to be the Maximum.

7) As discussed in Section 5.2.2, the center of maximum defuzzification method is used for the decision making purpose.

8) The defuzzified output represents the priority of the module, which is used in the scheduling process using the developed simulation model by Borrego (2004).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Using Simphony® (Simphony, 2000), a fuzzy decision making model template is developed to be used in the decision making process and linked to simulation models. Figure 5.7 shows the elements of the fuzzy decision making template developed in Simphony®.

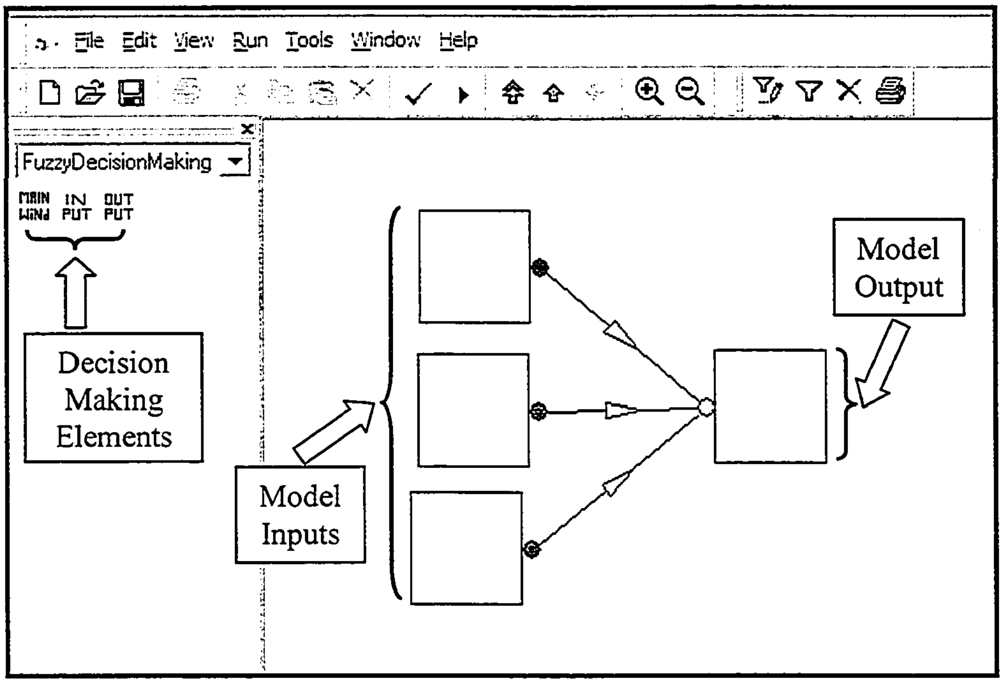


Figure 5.7: Elements of the Fuzzy Decision Making Template In Simphony

Three main elements are included in the fuzzy decision making which are:

1) The "MainWind" element refers to the main window element which is the parent element that contains the different model components. The input window of the "MainWind" element prompts the user to specify the number of inputs that the user will be creating and the source of the database in which the data of the inputs are stored.

2) The "Input" element allows the user to create the different decision criteria. Figure 5.8 shows the input window for the "Input" element in which the user inputs the different membership parameters for a decision

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

criterion. In addition, the user is prompted to input the "contribution" or weight of the decision input.

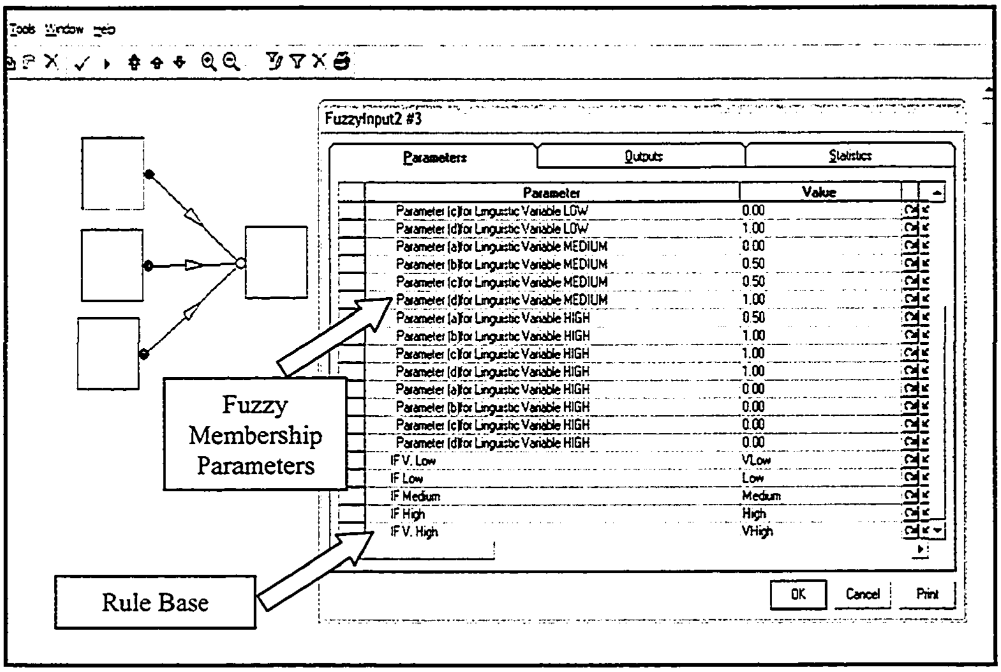


Figure 5.8: Input Window for the "Input" Element in the Fuzzy Decision Making Template

3) The "Output" element allows the user to create the decision output. The input window for the "Output" element is similar to the one shown in Figure 5.8 in which the user inputs the different membership parameters for the output.

4) The rule base is developed using the input window of the "Input" element. Each "Input" element forms a separate rule block. The number of rules will be matching the number of membership functions of the decision criterion assuming that there is no interaction between the different decision criteria.

The current fuzzy decision making template has some limitations that can be resolved in future improvement of the template. The template is designed to be

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

employed in the prioritization of the module assembly yard scheduling problem which will be illustrated in Section 5.3.4. However, it can be used in other decision making applications within the simulation framework provided that the following template limitations are accounted for:

1) Inputs are created separately assuming no interaction or rule combination between inputs.
2) Only one output can be created.
3) The maximum number of the membership functions of each input is 5 (very low, low, medium, high, and very high).
4) The maximum number of the output membership functions is 5 (very low, low, medium, high, and very high).
5) The input aggregation operator adopted is "Maximum".
6) The defuzzification method adopted is the "Center of Maximum".
7) The template is designed to read the different data of the decision criteria created from a database in which the different decision making outputs are saved.

Section 5.3.4 shows how the decision making template is utilized within the simulation framework using the module assembly yard problem as a case study

5.3.4 Running the Fuzzy Decision Making Template in Simulation

In this section, the developed fuzzy decision making template is used to prioritized the modules scheduled using the simulation model developed by Borrego (2004). A case study is designed to illustrate how the fuzzy decision

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

making template is used. The following assumptions are made for the illustrative case study:

1) Fifty modules are generated. Detailed information on all the modules is listed in Appendix C. The different module parameters listed in Appendix are randomly generated.
2) There are 4 types of modules (1,2,3 and 4) that are routed based on type to four different locations as follows:
    a) Modules of type 1 (cable tray) are routed to bay area 1 (total capacity of 16 modules)
    b) Modules of type 2 (equipment) are routed to bay area 2 (total capacity of 16 modules)
    c) Modules of type 3 (pipe rack) are routed to bay area 3 (total capacity of 16 modules)
    d) Modules of type 4 (structural) are routed to bay area 4 (total capacity of 16 modules)
3) It is assumed that the finished modules are shipped directly without waiting.
4) Each module undergoes one type of activity whose duration is listed in Appendix C.
5) The prioritization of the modules is based on the following criteria:
    a) Shortest processing time (SPT)
    b) Earliest due date (EDD)
    c) Maximum module cost (MMC)
    d) Combination of 1, 2, and 3 using fuzzy decision making.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Three rule blocks are created representing the three main criteria or decision attributes. Table 5.1 shows the rule base within each rule block and the fuzzy decision making parameters.

Table 5.1: Fuzzy Decision Making Parameters

Rule Block	Rule Base	Attribute Contribut.	Output Aggreg.	Deffuz. Method
Rule Block (1) SPT	IF SPT Low THEN High IF SPT Medium THEN Medium IF SPT High THEN Low	(0 to 1)	Max.	COM
Rule Block (2) EDD	IF EDD Low THEN Very High IF EDD Medium THEN Medium IF EDD High THEN Very Low	(0 to 1)	Max.	COM
Rule Block (3) MMC	IF MMC Low THEN Low IF MMC Medium THEN Medium IF MMC High THEN High	(0 to 1)	Max.	COM

The contributions of the attributes are changed based on the user's preference. Different scenarios are tested. Each attribute is first used as the main prioritization criterion by setting the other two attributes to zero. Then, combinations of all attributes are tested to show how the fuzzy decision making tool can be used to provide a multi-attribute and multi-objective decision. Table 5.2 shows the results of the different scenarios tested.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 5.2 Different Prioritization Scenarios Tested

Prioritization Criterion	Number of tardy Modules	% Tardiness	Average Flow Time	Average Waiting Time	Average Lateness	Average Earliness	Total Cost of Delayed Modules
Minimum Float	28	56	79.56	16.16	14.86	-6.08	962857
Earliest Due Date	23	46	76.28	12.58	12	-6.6	767854
SPT	26	52	81.2	17.5	16.52	-6.2	869222
MMC	26	52	81.64	17.94	17.3	-6.54	822726
Contribution = SPT =1.0 EDD =1.0 MMC = 1.0	24	48	80.16	16.46	15.86	-6.58	756780
Contribution = SPT =0.5 EDD =1.0 MMC = 0.5	23	46	78.26	14.56	14.02	-6.64	726319
Contribution = SPT =1 EDD =0.5 MMC = 0.	25	50	84.74	21.04	20.18	-6.32	816104
Contribution = SPT =0.5 EDD =0.5 MMC = 1.0	24	48	81.02	17.32	16.7	-6.56	756780

Different performance criteria were used to evaluate the effect of the prioritization criteria on the final schedule. Criteria based on due dates are calculated as:

a) The number of tardy jobs and percentage of tardiness, where tardiness = lateness if lateness > 0.

b) Average lateness, where lateness = completion time – due date.

c) Average earliness, where earliness = lateness if lateness < 0

Criteria based on module completion times are calculated as follows:

a) Average flow time, where flow time = completion time – release time

b) Average waiting time, where waiting time = completion time –release time – process time.

Criterion based on module cost is the calculated as follows:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

a) Total cost of delayed modules which is calculated by summing up the costs of the late modules.

Table 5.2 shows a comparison between the base case scenario where prioritization is done by processing modules of least float as introduced in Section 5.3.1 and other prioritization criteria. The following observations are made based on the results of Table 5.2:

1) When one prioritization criterion is utilized, it usually meets a single scheduling objective. For example, EDD and SPT rules minimized the number of tardy jobs. In addition, MMC rule generated one of the lowest total costs of delayed modules.
2) When scheduling rules or prioritization criteria are combined together by providing different contribution values to each, results generated illustrate some kind of compromise between more than one scheduling objective. For example, when all criteria are weighted one (means all receive 10 /10 in the contribution scale), the number of tardy modules decreased and the total cost of delayed modules also decreased.
3) The best performance was achieved when contribution of EDD was set to 1 (10/10) and when the other two criteria were set to 0.5 (5/10). One of the lowest numbers of tardy modules was generated and the lowest cost of delayed modules was achieved. The other performance criteria also achieved good results.

The fuzzy decision making tool provides a great deal of control flexibility and explicitness of the prioritization process of the module assembly operation. It is

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

important to note that the optimization of the prioritization criteria is beyond the scope of this study. The study shows, however, the importance of providing a decision making tool within the simulation framework. The case study provided shows the importance of having such a decision making system that accounts for both multi-attributes and multi-objectives within the decision making process, while accounting for the uncertainty caused by the fuzziness of the decision making criteria or attributes.

5.4 Conclusions

This chapter illustrates how a fuzzy rule based system can be used as a decision making tool within the simulation framework. The following points are concluded based on the results generated:

1) The fuzzy decision making model helped provide explicit control and decision making tool within the simulation model using the module assembly yard scheduling as a case study. In the case study, the user uses the fuzzy decision making tool to control the prioritization process of the modules awaiting assembly in the yard.
2) The integration and incorporation of a multi-attribute and multi-objective fuzzy decision making tool within the simulation framework will help make the control and decision making process within simulation models very explicit and more effective. Using this tool, the experts can test numerous decision scenarios within the simulation framework, which enhances the capabilities of the simulation modeling.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

3) The explicit fuzzy decision making tool will also help in minimizing the user's need to make changes within the simulation framework in order to account for his/her decision preferences.
4) The application of the fuzzy decision making tool to other decision making points within the simulation framework is very promising and can enhance the capabilities of the simulation models.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

CHAPTER 6 –INTEGRATED FUZZY MODELING AND SIMPHONY OBJECT MODEL FRAMEWORK

6.1 Introduction

Chapters 4 and 5 showed how some of the fuzzy set theory tools can be integrated with discrete event simulation modeling tools to provide better modeling capabilities. As demonstrated in Figure 6.1, the two major fuzzy tools integrated with the discrete event simulation model are the fuzzy predictive tool and fuzzy decision making tool.

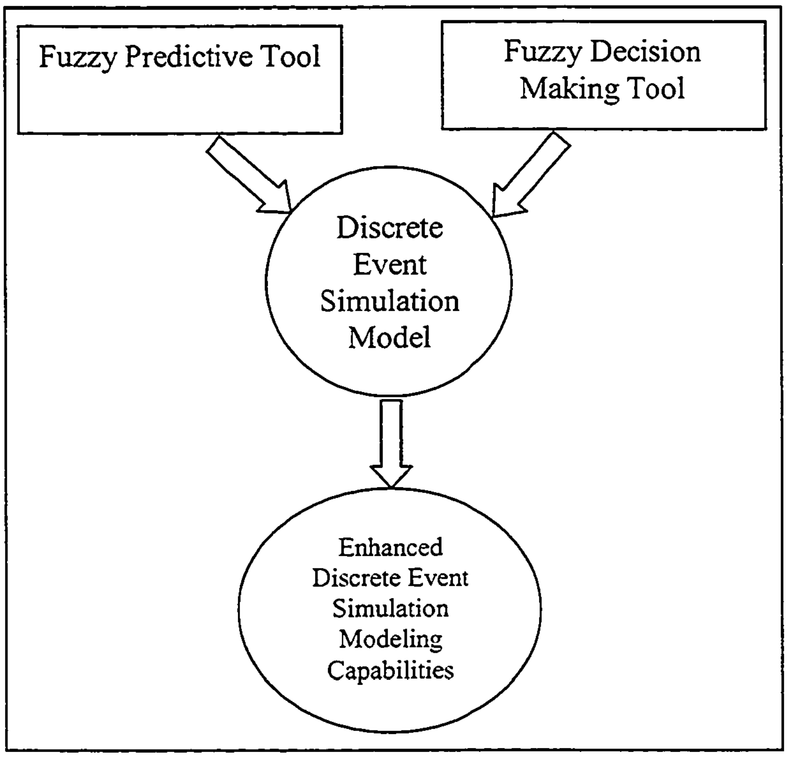


Figure 6.1: Integrated Fuzzy Tools with Discrete Event Simulation

The fuzzy predictive tool was integrated with the discrete event simulation model in Chapter 4 using the tunneling simulation template and fuzzy expert system for TBM advance rate prediction as a case study. The fuzzy decision making tool

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

was integrated with the simulation framework in Chapter 5 using the prioritization of module assembly yard scheduling simulation as a case study.

Chapter 2 presented fuzzy numbers as a comparable uncertainty modeling tool to the range estimating element using the probabilistic approach. This study was not integrated with the discrete event simulation framework. However, it was compared to the uncertainty range estimating tool which is part of Simphony Object Model (SOM) as introduced and presented by Hajjar (1999). In addition, the fuzzy numbers in range estimating is a component of an integrated fuzzy modeling framework which will be introduced in this chapter. All of the studies conducted throughout the thesis were performed within SOM developed by Hajjar (1999). Therefore, these studies initiated the need to present an integrated modeling that links the fuzzy modeling tools and the SOM. In order to integrate both systems, a fuzzy modeling framework is formulated and integrated within the framework of SOM. Section 6.2 provides an overview of the object-oriented frameworks development. In addition, the components of the fuzzy modeling framework and the proposed integration within the SOM are introduced and outlined in Section 6.3.

6.2 Overview of Object-Oriented Frameworks

The design and development of object oriented application frameworks are discussed and outlined by Froehlich et al. (1998a and 1998b). The development of the fuzzy modeling framework in this thesis is based on the work of Froehlich et al. (1998a and 1998b). A framework is defined as a combination of multiple objects used in conjunction to perform one or more tasks. Frameworks are

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

developed to provide a generic solution for a set of related problems within a specific domain. The following sections explain the different steps for designing and developing a framework.

6.2.1 Analysis

The first step of designing a framework is the analysis and definition of the problem domain. Defining the size of the domain will help define the scope and the applications that the framework will cover. The analysis of the framework domain will determine the key abstraction comprising the core of the framework. One approach of domain analysis is to examine existing applications within the problem domain and identify the different abstractions. In addition, developing different scenarios for the framework will help define the requirements of the framework and the primary abstractions and interaction patterns. The analysis will help identify the general areas of variability within the framework, referred to as "hot spots", which are the places in the framework that can be customized. In addition, the areas that the intended user has little or no control over, referred to as "frozen spots", can also be identified.

6.2.2 Design and Implementation

Constructing the abstractions, frozen spots, and hot spots is part of the design stage. In addition, the "hooks" which are connected to the "hot spots" and through which the framework is adapted and extended are designed and specified. Frameworks contain a core of abstract classes which embody the main architecture and interaction among the different framework classes. The

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

frameworks also contain a number of "concrete classes" forming the framework library. The concrete classes inherit from the abstract classes and provide specific functionality without modification. Hooks are provided into the framework either by composition or inheritance. In composition, the framework classes are adapted by filling in parameters that provide specific functionality. Inheritance involves adding functionality to an abstract in which the developer can easily add completely new functionality to a subclass of an existing class. Composition is usually utilized when framework interfaces and uses are well defined. On the other hand, inheritance is used when the full range of framework functionality is not fully anticipated.

6.2.3 Testing

Framework testing can be done by developing applications to help identify the areas where the framework requires modifications and enhancement. Another method is to test the framework by itself without any application extensions in order to avoid the errors that might be caused by the applications and help detect the defects within the framework itself.

6.2.4 Deployment

Once the framework is designed and tested, it needs to be deployed by the application developers. Some issues need to be considered when deploying a framework. First, the framework should be introduced gradually to the users for better learning of its basic components and requirements. This objective can be achieved through sessions, example applications and reference documentation.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Second, the distribution form that the framework will take has to be defined (i.e. source code). Third, collecting the users' feedback, comments, and requests is very important for future change and enhancement.

6.3 Fuzzy Modeling Application Framework Design

6.3.1 Framework Scope and Purpose

The domain of the designed framework is modeling construction engineering problems. The fuzzy modeling framework is used in modeling uncertainty in construction cost using fuzzy numbers, predicting construction related behaviors, and incorporating expert decisions in construction related problems. Chapters 2, 3 and 4 provided examples of the previously mentioned uses of fuzzy modeling framework. The fuzzy modeling framework will work as a support framework for SOM framework developed by Hajjar (1999). Figure 6.2 shows the domain of the fuzzy modeling framework.

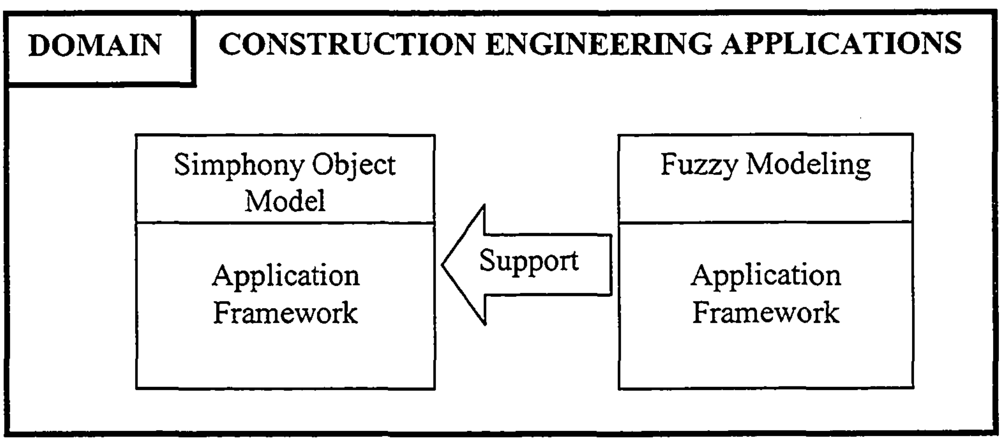


Figure 6.2: Fuzzy Modeling Application Framework Domain

The SOM is a simulation object-oriented application framework based on a unified modeling methodology achieved through the formalization of a generic

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

base modeling element. The generic base modeling element is used in the development of a special purpose simulation template in a specific construction domain. The generic base modeling element was implemented as a parameterized class which is customized through composition in which developers create objects based on the generic class (Hajjar, 1999). One way of customizing a specific modeling element is by defining its attributes. This methodology simplifies the development of new construction simulation tools.

6.3.2 Fuzzy Modeling Framework Components

Figures 6.3 and 6.4 show the different views of the fuzzy modeling framework. The framework is divided into three main unified components:

1) Fuzzy numbers in cost range estimating.
2) Fuzzy prediction.
3) Fuzzy decision making.

The last two components of the framework are directly integrated with SOM simulation framework. The output of the fuzzy modeling framework is the input to the SOM. The following paragraphs explain the details of the framework components.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.