CHAPTER III

RANDOM FUZZY NETWORK SCHEDULING

Construction projects are usually complex and risky. A large number of interrelated operations are involved. Durations of these operations are correlated because the durations are affected by factors such as unforeseeable underground condition, adverse weather, lacking practical experience, failing to distinguish site characteristics, varying from location to location, and being of poor management quality. Each of these factors might induce other factors to have the greater impact on the operation duration. A better understanding of the impacts of these factors would make the construction schedule a more reliable by providing a better estimate of uncertainty in duration of project activities and a project completion time and assisting in focusing attention on a set of correlated factors that have the considerable impact.

Simulation is considered as a useful tool to measure the effects of external factors, like those mentioned above, by introducing randomness into the analysis of construction processes. Probabilistic methods use marginal probability distributions to represent the characteristics of the scheduling variables being modeled. The application of probabilistic distributions implicitly assumes that past performance of the same activity has been investigated. A marginal distribution is then established from these investigations. Consequently, different conditions and random events found in new projects can be modeled by using the developed distributions. In practice, however, it is generally known that construction projects are unique. Working conditions for particular projects at activity level are different. Quantitative data related to affecting factors are usually unavailable. Scheduling experts have to use their own judgement based on their experience to assess risk factors and evaluate their impacts on duration of project activities. Results involve another type of uncertainties which cannot be addressed by using the probability theory. A fuzzy set and logic theory typically used to measure uncertainty stemming from systematic and unknown contributions is employed to overcome this limitation,. The applications of these two theories to network scheduling, risk assessment, and impact evaluation are explained in Chapter II.

This chapter is devoted to the presentation of a new network calculation method accounting for every uncertainty in the assessment of risk factors, the evaluation of their impacts on the activity duration and the construction scheduling. The proposed method employs both probability theory and fuzzy set theory to model uncertainties associated with the attributes of risk factors and duration of project activities which are brought about by the contributions (i.e., random, systematic, and unknown). A random–fuzzy variable is used to represent those uncertainties. Membership functions of durations of project activities are established by using a simple statistical function and alternatively a nonlinear function obtained from a neurofuzzy system used to capture the relationships between the attributes of risk factors and duration of project activities. The mathematics for random–fuzzy variables are used to propagate the underlying uncertainties through the project network. The created membership functions of durations of project activities are inputted into the network calculation performed on the mathematic for random–fuzzy variables. Components involved in the proposed method are presented in the form of a flowchart as shown in Figure 3.1. The proposed method calculates the project completion time by using these following steps:

(1) Identifying significant risk factors affecting each activity based on subjective judgement and experience.

(2) Providing definitions of likelihood and consequence of risk factors based on available data (i.e., quantitative and subjective data)

(3) Developing functions to transform information obtained from the risk assessment into the variation in activity duration which is represented by most likely and maximum values

(4) Acquiring data from multiple sources including historical records, questionnaire, interview, and direct observation.

(5) Performing pairwise comparisons of risk factors affecting each activity and computing weights of risk factors.

(6) Assessing the likelihood and consequence of risk factors affecting a particular activity based on the given definitions. Results are represented by fuzzy numbers.

(7) Computing a total score of risk factors based on the weights and assessed values of the likelihood and consequence of risk factors in consideration.

(8) Calculating variation (i.e., most likely and maximum values) in activity duration due to risk factors. Each portion of activity

durations representing the impact of a particular risk factor influencing a considered activity is given by using defined functions of L , C , and E .

(9) Estimating duration based on the optimistic estimates (made without determining impacts of risk factors) obtained from the as – planned database or the normal progress (made based on expected conditions) and remaining duration obtained from the as – built database

(10) Revising preliminary estimated optimistic duration and the normal progress and remaining duration by adding up different portions of durations of project activities obtained from step (8) with a consideration about points of times that the particular portions occur.

(11) Establishing probability distributions of duration of project activity, probability distribution of likelihood and consequence of a risk factor.

(12) Developing a simulation model capturing construction processes in consideration.

(13) Estimating the project completion time based on the simulated activity duration drawn randomly based on the derived probability distributions.

(14) Collecting simulation data of activity duration and likelihood and consequence of risk factors.

(15) Transforming the probability distribution of activity duration into the triangular membership function representing a random part of activity duration by using a Salicone's method or a neurofuzzy metamodel trained on the collected simulation data (i.e., activity duration and likelihood and consequence of risk factors) and new values of likelihood and consequence of risk factors obtained from a case study project

(16) Developing the triangular probability distribution of activity duration by using a neurofuzzy metamodel trained on the collected simulation data (i.e., activity duration and likelihood and

consequence of risk factors) and new values of likelihood and consequence of risk factors obtained from a case study project.

(17) Identifying factors affecting the subjective assessment including (i) the complexity of working and conditions for judgement and (ii) the level of education and experience.

(18) Assigning values of epistemic uncertainty due to factors affecting the subjective assessment

(19) Developing a membership function of a fuzzy part of activity duration based on the assigned values of the epistemic uncertainty

(20) Developing a final membership function of activity duration by inserting the membership function of the fuzzy part into the membership function of the random part

(21) Computing the project completion time by using mathematics for random–fuzzy variables in the forward and backward pass calculations such as those performed in the FNET and the credibility coefficients C_{gr}, C_{lo}, and C_{eq}

The following sections describe the fundamental concepts that are necessary for the development of the proposed method. Firstly, a definition of a random–fuzzy variable used to represent duration of project activities and attributes of risk factors affecting those activities is provided. Then, definitions of attributes of risk factors are offered to acquire data from several approaches to perform the risk assessment and the evaluation of impact of risk factors on duration of project activities. Next, the establishment of membership functions representing a random part and a nonrandom part of random–fuzzy variables is provided by using a statistical analysis method (i.e., Salicone's method) and a neurofuzzy metamodel. Steps used to compute a project completion time is thereafter described together with the mathematics applied to the network calculation and methods used to make a comparison among random–fuzzy durations.

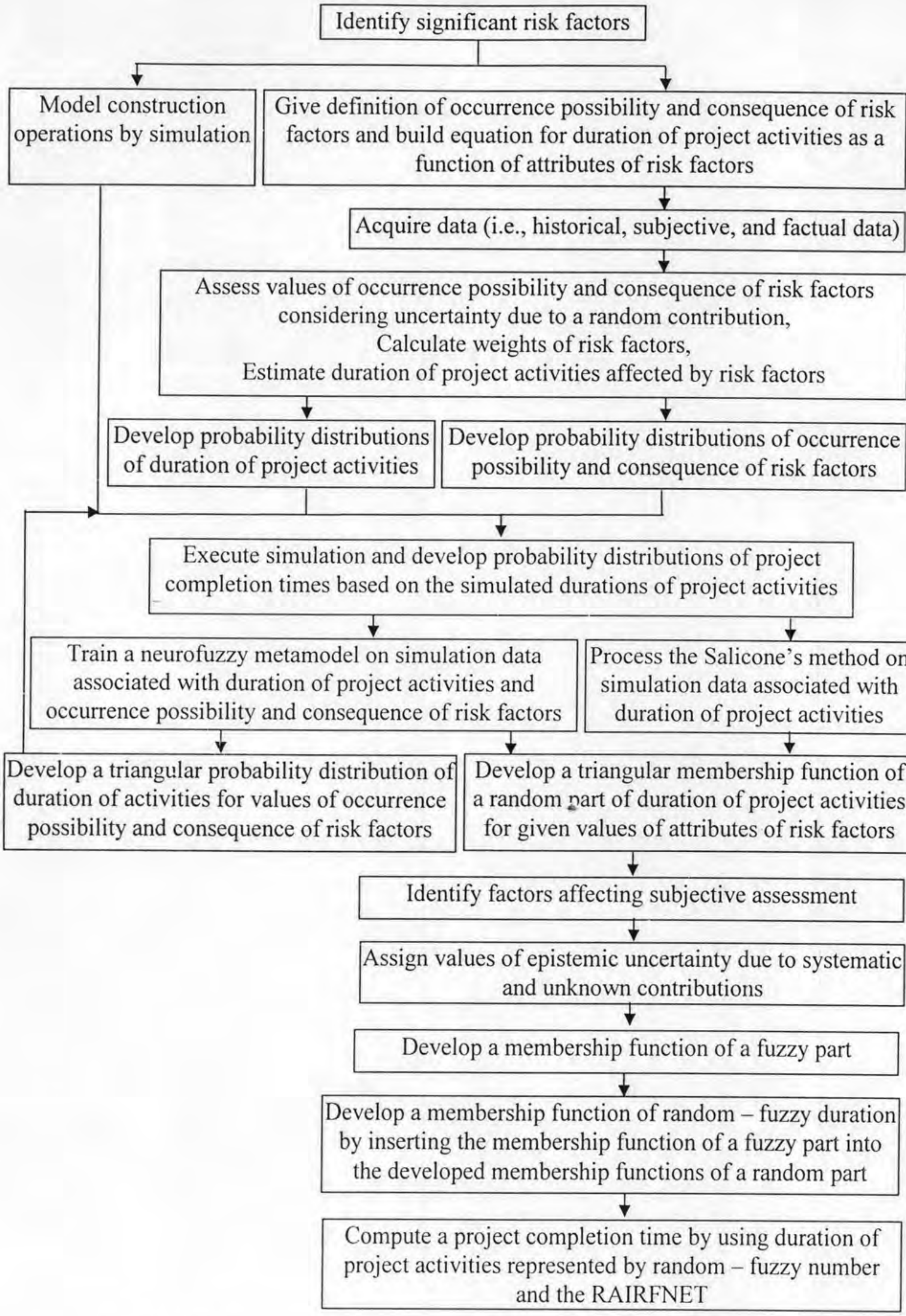


Figure 3.1 A flow chart indicating components involved in the proposed approach

3.1 Random–Fuzzy Variable

A random–fuzzy variable can generalize both concepts of a random contribution to aleatory uncertainty and systematic and unknown contributions to epistemic uncertainty. Thus, it is able to address every uncertainty. The random fuzzy variable is used to consider uncertainties of attributes (i.e., likelihood of occurrence L, adverse consequence C, and extended duration E for each level of adverse consequence) of a risk factor and duration of project activities on the basis of existing studies (Salicone, 2007; Wang and Demsetz, 2000; Oliveros and Fayek, 2005; and Cho et al, 2002).

To demonstrate the procedure for constructing a membership function of a random–fuzzy variable, there are four steps presenting as follows:

First, an internal membership function μ_I is constructed to represent the linguistic variables presenting the degree of epistemic uncertainties in accordance with different combinations of the subjective values of the factors that affect determining the linguistic variables for any uncertainty range. The factors are composed of the complexity of working condition and the level of the experiences of an assessor and the familiarity with the project in hand and projects of a similar nature which are related to the subjective judgments and the unreliable/insufficient data and the use of inadequate statistical analysis methods which are associated with the probabilistic parameter estimates.

Second, a membership function of a random–fuzzy variable that only represents the random contribution to aleatory uncertainty is transformed from a corresponding probability distribution function. In the proposed approach, the probability distribution of activity duration is established by performing the sum, product, and division operations to compute duration of project activities.

The proposed approach employs the basis of the production rate typically used in estimating activity duration. Therefore, duration of project activities is calculated in terms of work quantity and production rate: $D = W/P$, where D is activity duration, W is work quantity, P is production rate. Impact of risk factors is presented in terms of the extended duration, increased quantity of work, and decreased production rate. The basis of the production rate is then used to calculate the extended duration in terms of increased quantity of work and decreased production rate. The impacted variables (i.e., D, W, P) particularly vary according to the identified risk factors and their impacts.

A factor – based procedure facilitated by an activity duration model is employed to disaggregate the effects of uncertainty from a set of risk factors and the conditions of a particular risk factor. Therefore, a mathematical function of attributes of a risk factor, and a function of production rate and quantity of work are used to compute duration of project activities. As data related to risk factors are usually unavailable, the subjective risk assessment and impact evaluation are provided. The variations of duration resulting from a set of risk factors and particular risk factors are represented by probability distributions based on the results (i.e., mean μ_D and standard deviation σ_D) obtained from the developed mathematical functions. The production rate, quantity of work and attributes of a specific risk factor are also presented by the probability distributions based on given definitions.

Different types of probability distributions are used to capture data related to duration of project activities which include either a parametric distributions (i.e., normal distribution) or a non – parametric distributions (i.e., triangular and uniform distributions) that have been artificially created.

Third, the establishment of the membership function μ_2 representing a random part of a random–fuzzy variable is provided regarding the corresponding probability distribution. The probability distributions are transformed into the membership function μ_2 by applying the probability–possibility transformation method.

Forth, the combination of μ_1 and μ_2 is obtained by inserting the membership function μ_1 into the membership function μ_2. In this way, the final random–fuzzy variable will have an internal membership function which considers all systematic and unknown contributions, and an external membership function, which determines all kinds of contributions. In the other words, the internal membership function is μ_1, whereas the external membership function is built examining both μ_1 and μ_2.

3.2 Development of Probability Distribution

In the proposed method, the likelihood of occurrence L, adverse consequence C, and extended duration E for each level of adverse consequence of a risk factor, production rate P, quantity of work Q, and duration D of project activities are considered as random–fuzzy variables. The external membership function of a random–fuzzy variable is transformed from the probability distribution of a random–fuzzy variable.

This study applies the Monte Carlo simulation to find the probability distributions of duration of activities based on values of L, C, E, P, and Q drawn randomly based on the derived probability distribution functions. A number of data sets will be based on given simulation iterations. For example, if a value of 0.5 randomly provided by simulation from a distribution of the likelihood of occurrence, the likelihood of occurrence of a risk factor is 0.5 which indicates a 50 % chance that such risk factor will occur. The following sections explain the development of the probability distributions of likelihood and consequence of risk factors and duration of project activities.

3.2.1 Likelihood of Occurrence of a Risk Factor

For each activity, the quantitative likelihood and weight (Wt) or degree of significance of a risk factor are used to help establish the distribution of the likelihood $L_{i(j)}$ of a particular risk factor j occurring in an activity i. Risk factors are usually dependent on each other. Therefore, a certain likelihood of a specific risk factor causing the extended duration might be immeasurable. In addition, the historical data normally used to derive the reference likelihood of a specific risk factor might be unavailable or inapplicable to an existing project because construction projects are unique.

This study modifies a model for evaluating networks under correlated uncertainty (NETCOR) proposed by Wang and Demsetz (2000) and Wang, Liu and Chou (2006) in order to measure the likelihood of occurrence of a risk factor. The subjective assessment of the likelihood of occurrence of a particular risk factor is performed so as to adjust the distribution of likelihood of risk factors which is developed based on the historical data. It can overcome the problems related to immeasurable likelihood of a particular risk factor discussed above. To provide a stimulus for integrating information obtained from the risk assessment into the network scheduling, the subjective likelihood of a risk factor has to be transformed into the quantitative one.

Firstly the definition of an occurrence of risk factors is provided in order to measure the occurrence likelihood of risk factors as frequencies of occurrences of risk factors can be used to approximate the occurrence likelihood of risk factors. It is helpful to compare the likelihood that the risk factor will occur with the length of the project appraisal. For example, considering a period of six months, a risk factor with a likelihood of occurrence of once in six months is relevant. As the length of the project appraisal is a function of units of a product constructed in and times required for complete each unit, a risk factor with a likelihood of occurrence of once in a number of units of a product built within a specific construction period is determined.

This study could count a number of occurrences of risk factors whenever such risk factors extend the initially estimated duration of an activity carried out for finishing each unit of a product. The quantitative likelihood of risk factors is defined as the total number of times that the risk factors extend duration of a considered activity regarding the initially estimated duration divided by the total number of times that a considered activity is carried out during the designed period (or the total number of times that a considered activity is carried out for finishing the required number of units of a product). If the relevant historical data are available, the quantitative likelihood of occurrence of risk factors is used to help examine the quantitative likelihood of a particular risk factor by considering its subjective value as a multiplier.

By nature of construction projects in general, significance of a risk factor j for an activity i is different from other risk factors. To reflect their differences, the experts were asked to perform pairwise comparisons between the significance of risk factors affecting each activity. The results were then used to compute weight or degree of significance of a risk factor. The weight of each risk factor remains the same for any project. Assuming that the historical data are available and applicable to any succeeding project, the likelihood of a particular risk factor is a function of the adjustment constants (i.e., quantitative likelihood of risk factors, weight of a particular risk factor) for any project, and the subjective likelihood of a considered risk factor or a multiplier which differs for each project.

$$I_{i(j)} = Wt \times P_{i(j)} \times F_{i(j)} \tag{3.1}$$

$$F_{i(j)} = \frac{\text{A total number of occurrences of extension of duration due to risk factor } j}{\text{A total number of times that an activity is carried out}} \tag{3.2}$$

$$P_{i(j)} = \frac{\text{Likelihood of occurrence of a risk factor } j}{\text{Likelihood of occurrence of all risk factors}} \tag{3.3}$$

where Wt_j is weight of a risk factor, $F_{i(j)}$ is quantitative likelihood of a set of risk factors, and $P_{i(j)}$ is subjective likelihood of a risk factor or a multiplier

Linguistic terms (i.e., very high, high, medium, low, very low) are used to perform the subjective assessment of the multiplier $P_{i(j)}$. The assessed linguistic terms are transformed into the cardinal value ranged from 0 to 100. As assessors might have different perceptions towards the linguistic terms of $P_{i(j)}$, the measurement of the individual perception in each term must be made before performing the subjective assessment. This study modifies a model used to calibrate linguistic terms proposed by Pipattanapiwong (2004). The scales of $P_{i(j)}$ are calibrated before carrying out the assessment. Examples of calibrating scales of linguistic expressions for assessing $P_{i(j)}$ are shown in Figure 3.2. The expected scale is mean between the lower bound and upper bound of the sub-range corresponding to the expression. Table 3.1 shows the example of the scale of $P_{i(j)}$.

Assessment scale (%) for $P_{i(j)}$

	0–10	10–20	20–30	30–40	40–50	50–60	60–70	70–80	80–90	90–100
Very high										
High										
Medium										
Low										
Very low										

Figure 3.2 Example of calibrating scale for assessment of probability of consequent risk

Table 3.1 Example of $P_{i(j)}$ assessment expression

Description	Scenario	Lower-upper possibility (%)	Mean (%)
Very high	Very frequent occurrence	80-100	90
High	More than even chance	60-80	70
Medium	Quite often occurs	40-60	50
Low	Small likelihood but could well happen	1-40	20
Very low	Not expected to happen	Less than 1	Less than 1

Three – point estimation is adopted to establish the distribution $l_{i(j)}$ of each risk factor. Then, the mean $M_{i(j)}$ and standard deviation $\sigma_{i(j)}$ of this distribution are computed. In this study, experts determine $l_{i(j)}$ based on available data of an existing project. The average, standard deviation, minimum and maximum values of $l_{i(j)}$ are later employed to help develop the distributions of $l_{i(j)}$.

The overall distribution of likelihood of risk factors $L_{i(j)}$ is disaggregated as the sum of distributions of likelihood of each risk factor. Mathematically, $L_{i(j)}$, a random variable, is shown as follows

$$L_{i(j)} = l_{i(1)} + l_{i(2)} + ... + l_{i(j)} + ... + l_{i(J)} \quad (3.4)$$

where $l_{i(j)}$ is the distribution of likelihood of risk factor j affecting activity i. The Monte Carlo simulation is applied to find $L_{i(j)}$ based on the probability distribution functions from which the $l_{i(j)}$ are derived. However, it should be realized that all risk factors actually occurring in the previous projects cannot be completely identified. Therefore, the reference likelihood of occurrence of the risk factors is unlikely equal to the overall likelihood obtained from simulation based on particular risk factors identified based on available information of an existing project.

3.2.2 Activity Duration Affected by Risk Factor

This study presents impact of risk factors in terms of duration extended from duration preliminarily estimated based on an expected condition. This section presents the development of the distribution of duration of project activity i affected by a risk factor j during the designed period which is denoted as $e_{i(j)}$. Typically, the distribution is established by using historical data acquired through completed projects. However, data related to duration of project activities affected by risk factors are not fully recorded. Duration of project activities affected by a particular risk factor also cannot be precisely estimated. If the historical data is statistically insufficient to create the distribution, the distribution is established from durations which are subjectively estimated.

This study modifies a model for evaluating networks under correlated uncertainty (NETCOR) proposed by Wang and Demsetz (2000) and Wang, Liu and Chou (2006) in order to develop the distribution of duration of project activity i affected by a risk factor and also the distribution of duration of project activity i affected by a set of risk factors. The subjective assessment of duration of project activities affected by a particular risk factor is performed so as to adjust the distribution developed based on the historical data.

As mentioned above, several risk factors simultaneously cause the extended activity duration. Generally, these risk factors are dependent. Thus, precise duration of a project activity affected by a particular risk factor cannot be measured. The subjective consequence evaluation has to be performed. The result can be also considered as a multiplier which is accordingly used to help estimate the duration of a project activity affected by a considered risk factor. Then, the distribution of duration of an activity is established with respect to the subjective consequence $C_{i(j)}$.

$$C_{i(j)} = \frac{\text{Extended duration due to a risk factor } j \text{ occurring for an activity } i}{\text{Extended duration due to every risk factor}} \quad (3.5)$$

The linguistic terms are used to assess consequences of a risk factor. When the consequence of a risk factor is not well defined, the assessed result might be inaccurate because of the linguistic imprecision. In addition, the risk assessors might impose their understanding and structuring of the risk factors on the experts. In order to mitigate the linguistic imprecision and the tendency to be led – effect, the descriptions of risk factors, consequence of a risk factors and linguistic terms used in the assessment are given unambiguously. The logic of the uncertainty in questions is also explained before performing the assessment. This study modifies a model used to calibrate linguistic terms proposed by Pipattanapiwong (2004). The calibration of the linguistic expressions for assessing $e_{i(j)}$ is performed as shown in Table 3.2 to describe the assessors' perception in $e_{i(j)}$. The measurement of the individual perception in each term is provided before performing the subjective consequence evaluation because assessors may have different perceptions towards the linguistic terms of $e_{i(j)}$.

To simplify the subjective risk assessment, the linguistic terms employed to describe the assessors' perception in $P_{i(j)}$ previously presented in Table 3.1 are used to assess $C_{i(j)}$. The measurement of the individual perception in each term must be made before performing the subjective consequence assessment because assessors may have different perceptions towards the linguistic terms of $C_{i(j)}$ and it might be different from $P_{i(j)}$. The calibration of the linguistic expressions for assessing $C_{i(j)}$ is analogous to the one for $P_{i(j)}$. Figure 3.4 is also used to present examples of calibrating scales of linguistic expressions for assessing $C_{i(j)}$ similar to the assessment of $P_{i(j)}$. Table 3.3 shows the example of the scale of $C_{i(j)}$.

The overall distribution of duration of a project activity affected by a particular risk factor, denoted as $E_{i(j)}$, is disaggregated as the sum of distributions of extended duration associated with each risk factor $e_{i(j)}$. Mathematically, $E_{i(j)}$, a random variable, is shown as follows

$$E_{i(j)} = e_{i(1)} + e_{i(2)} + ... + e_{i(j)} + ... + e_{i(J)} \tag{3.6}$$

where $e_{i(j)}$ is the distribution of extended duration for a level of consequence of risk factor j for activity i. This study applies the Monte Carlo simulation to develop $E_{i(j)}$ based on the probability distribution functions from which the $e_{i(j)}$ are derived.

Table 3.2 Example of $e_{i(j)}$ assessment expression

Description	Scenario	Lower-upper extended duration (%)	Mean (%)
Catastrophic	Risk can result in lengthy extended duration on a wide scale	90-100	95
Critical	Risk can result in serious extended duration	60-90	75
Marginal	Risk can cause extended duration, but the results would not be expected to be serious	30-60	45
Acceptable	Risk can cause extended duration, but the results would be within the range of expected level	1-30	15
Negligible	Risk will not result in serious extended duration	Less than1	Less than 1

Table 3.3 Example of $C_{i(j)}$ assessment expression

Description	Scenario	Lower-upper possibility (%)	Mean (%)
Very high	Very frequent occurrence	90-100	95
High	More than even chance	70-90	80
Medium	Quite often occurs	40-70	55
Low	Small likelihood but could well happen	1-40	20
Very low	Not expected to happen	Less than 1	Less than 1

Following the methodology proposed by Wang and Demsetz (2000) and Wang, Liu and Chou (2006), the overall distribution of duration of project activities $D_{i(j)}$ is considered as the sum of distributions associated with risk factors. The expected duration of a project activity affected by a particular risk factor is computed as follows

$$D_i = \sum_{j=1}^{J} D_{i(j)} = \sum_{j}^{J} l_{i(j)} \otimes c_{i(j)} \otimes e_{i(j)} \quad (3.7)$$

where $l_{i(j)}$, $c_{i(j)}$, and $e_{i(j)}$ are randomly drawn based on the corresponding distributions.

In the deterministic environment, the estimated activity duration $D_{i(estimated)}$ can be mathematically represented as follows (Pipattanapiwong, 2004)

$$D_{i(estimated)} = D_{i(0)} + \sum_{j=1}^{J} D_{i(j)} \quad (3.8)$$

where $D_{i(0)}$ is the deterministic duration initially estimated without determining risk factors or duration initially estimated under the expected conditions of all risk factors, and the random variable $D_{i(j)}$ represents a set of duration distributions of duration of a project activity i probably affected by a particular risk factors $j = 1,...,J$.

While the extended duration transformed from the impact of risks factors into duration of a project activity is described above, the transformation from the impact of risk factors represented by the increased quantity of work and decreased production rate into the duration of a project activity is achieved by considering the division w/p based on the concept of the production rate. As risk factors randomly occur, the selected mathematic must be able to represent any behavior of the random contributions. A suitable mathematic should be defined for the product operation as shown in Eq (3.8) and (3.6) and also the division operation (i.e., w/p). However, when product and division operations are considered, typically, a probability distribution to which the result tends is not exactly known. Simulation is usually used to generate the distribution (Salicone, 2008). By running simulation a number of times, a probability distribution to which the result tends can be created. The random variables in simulation are composed of risk variables and temporal variable. Simulation models the activity duration by using a following equation.

$$\overline{D} = \overline{D_d} + \overline{D_{w/p}} \quad (3.9)$$

where $\overline{D}$ is the probabilistic activity duration regarding every impact of risk factors, $\overline{D_d}$ is the probabilistic activity duration regarding impact of risk factors on the activity duration, $\overline{D_{w/p}}$ is the probabilistic activity duration regarding impact of risk factors on the work quantity and production rate.

3.3 Random–Fuzzy Concept

3.3.1 Random–Fuzzy Concept

This section is to describe the definition of a random–fuzzy variable. At first, the concept of interval confidence is introduced. An interval of confidences is a closed interval in R, which the possible values of an uncertain result can be located between its lower and upper bound. However, lower and upper bound might be uncertain. Thus, the determination of the uncertain intervals of confidence is demonstrated by a set of two intervals of confidence, which are referred to as the internal interval and the external interval. The properties of a random–fuzzy variable or a fuzzy variable of type 2 illustrated with the internal interval and the external interval are shown in Figure 3.2, where the intervals indicated by the equal arrows. The intervals corresponding to a greater value of α are included in those corresponding to smaller values, the maximum of the interval membership function is obtained for value α_m.

The internal interval of the random–fuzzy variables is used to represent the measurement results (e.g., activity duration, attributes of a risk factor) affected by systematic and unknown contributions to epistemic uncertainty. While the internal interval is used to define the minimum variation of the given confidence interval, the external interval is employed to express the maximum one. Therefore, the side intervals of each interval of confidence of the random–fuzzy variables intuitively represent the random contribution to aleatory uncertainty.

The available information associated with the internal interval leads to the establishment of the internal membership function of the random–fuzzy variable. The external membership function of the random–fuzzy variable is created regarding the probability distribution function representing the distribution of the random contribution to aleatory uncertainty. To construct a membership function of a random–fuzzy variable, the transformation from the probability distribution function into an equivalent possibility distribution function representing the external membership function is required. Different behaviors associated with each part of the random–fuzzy variables and their mathematics have to be considered before mathematically integrating the internal membership function with the external one.

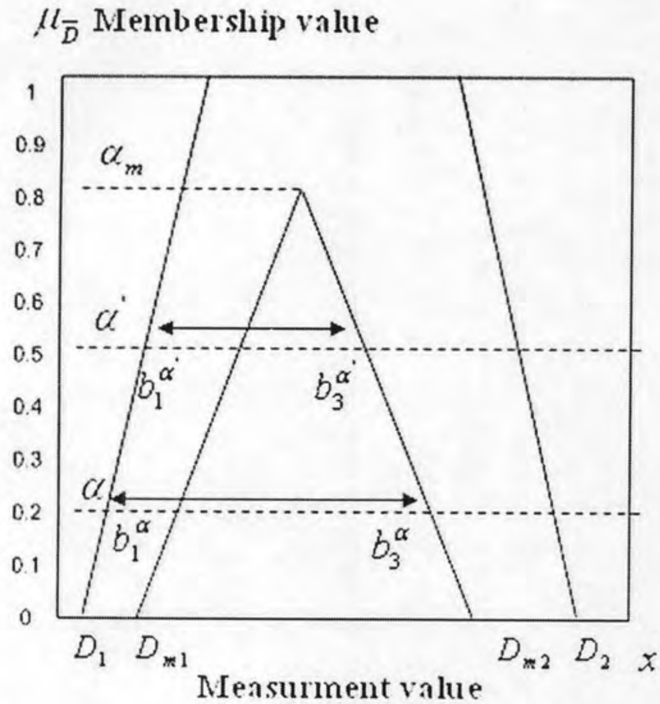


Figure 3.3 Properties for a fuzzy number of types 2

3.3.2 Probability–Possibility Transformation Method

There are several methods proposed to transform the probability distribution into the possibility distribution. The transformations between probabilities and possibilities are very different under different types of scales. There are five most common scale types including ratio, difference, interval, log – interval, and ordinal scales (Klir, 2005). Salicone (2007) provided the suitable way to transform the information contained in a probability distribution function into a possibility distribution function based on the concept of intervals of confidence which is presented in both the probability and possibility theories.

Let λ (where $0 \le \lambda \le 1$) denote a particular level of confidence, the confidence interval of a probability distribution function $p(x)$ is the interval $[x_a, x_b]$ for which

$$\int_{x_a}^{x_b} p(x)dx = \lambda \tag{3.10}$$

For a fuzzy or a random–fuzzy variable X, the level of confidence associated with the generic α - cut or the generic confidence interval is given by the necessary measure defined over the α - cut itself, and is written as

$$Nec(X_\alpha) = 1 - \alpha \tag{3.11}$$

The probability–possibility transformation is based on Eq. (3.10) and (3.11). If interval $[x_a, x_b]$ and the associated level of confidence λ are known, the $[x_a, x_b]$ is the α - cut of a fuzzy variable at level:

$$\alpha = 1 - \lambda \tag{3.12}$$

When the interval $[x_a, x_b]$ to the confidence interval at level of confidence 1 is determined, the same interval is considered the support of the possibility distribution function corresponding to the given probability distribution function. The interval $[x_a, x_b]$ for $\int_{x_a}^{x_b} p(x)dx = 1$ is the α - cut at level $\alpha = 0$ of the membership function.

Let x_p denotes the absolute maximum value of the probability distribution function, the peak value of the desired membership function is assumed to fall at this value. The α - cut at level $\alpha = 1$ of the membership function is the degenerated interval $[x_a, x_b]$ for $\int_{x_a}^{x_b} p(x)dx = 0$ or the interval $[x_a, x_b]$ to the confidence interval at level of confidence 0. It means that a perfect agreement is obtained between the information given by the probability distribution function and the possibility distribution function. Both of them provide the same interval of confidence for each desired level of confidence, although the shape of the possibility distribution function is different from the shape of the probability distribution function.

To present uncertainty due to a random contribution in a simply manner, this study uses a triangular membership function to represent the random part of a random–fuzzy variable (i.e., risk variations and temporal variation). The confidence interval of the membership function is represented by the interval $[-3\sigma, 3\sigma]$ around the mean value which is the interval of confidence at level of confidence $0.997 \approx 1$. Hence, the obtained lower bound and upper bound of the membership function are spread at level $\alpha = 0$ of the membership function. The mean value of the probability distribution represents the peak value of the triangular membership function. After the α – cuts for $\alpha = 0$ and $\alpha = 1$ have been found, the intermediate α – cuts are interpolated from these two points in order to establish the triangular membership function. To approximate the way the random variables combine together and calculate their mean and standard deviation, Monte Carlo simulation is employed to generate data and estimate the probability distribution functions representing uncertainty involved in the random variables according to the requirement of a statistic analysis.

Once the membership functions representing a random part of random–fuzzy variables are established, the development of the internal membership function representing a nonrandom part of a random–fuzzy variable is provided to consider the uncertainties due to systematic and unknown contributions or uncertainties of subjective judgement. Based on a risk assessment methodology for incorporating uncertainties using fuzzy concepts developed by Cho et al. (2002), the factors that have influence on the uncertainties of subjective judgement are divided into two main categories: (1) the complexity of working and conditions for judgement and (2) the level of education and experience. The linguistic variables are also used to represent the degree of uncertainties according to different combinations of the two main factors. Then, values of these two factors are subjectively evaluated. The internal membership function is accordingly developed based on the assessed values. The membership function representing a random part is now called an external membership function.

As the meaning of the internal and external intervals is differently defined by the interval of confidence of a random–fuzzy variable, different methods are used to determine them. For simplicity, this study uses the method proposed by Salicone (2007) to intuitively combine two membership functions. Different methods are also used to determine them. The method to combine two membership functions is very intuitive. Let μ_1 be the membership function of a random–fuzzy variable when no random contribution affects the measurement result and μ_2 be the membership function of a random–fuzzy variable when only random contribution affects the measurement result, for which the internal membership function is nil or the width of all its α - cuts is zero. Then, a membership function of the random–fuzzy variable is obtained by inserting the membership function μ_1 into the membership function μ_2. In other words, the internal membership function is μ_1, while the external membership function is established by taking into account both μ_1 and μ_2. Let

$\left[x^{\alpha}_{L_1}, x^{\alpha}_{R_1}\right]$ and $\left[x^{\alpha}_{L_2}, x^{\alpha}_{R_2}\right]$ denote the generic $\alpha-$cuts of the two membership functions μ_1 and μ_2, the interval $\left[x^{\alpha}_{L_2}, x^{\alpha}_{R_2}\right]$ can be divided into the two intervals $\left[x^{\alpha}_{L_2}, x_p\right]$ and $\left[x_p, x^{\alpha}_{R_2}\right]$, where x_p is the peak value of μ_2. The $\alpha-$cut $X_{\alpha} = \left\{x^{\alpha}_a, x^{\alpha}_b, x^{\alpha}_c, x^{\alpha}_d\right\}$ of the random–fuzzy variable can be defined as

$$x^{\alpha}_b = x^{\alpha}_{L_1} \tag{3.13}$$

$$x^{\alpha}_c = x^{\alpha}_{R_1} \tag{3.14}$$

$$x^{\alpha}_a = x^{\alpha}_b - (x_p - x^{\alpha}_{L_2}) \tag{3.15}$$

$$x^{\alpha}_d = x^{\alpha}_c - (x^{\alpha}_{R_2} - x_p) \tag{3.16}$$

The random–fuzzy variable is established by applying these equations to all $\alpha-$cut. Figure 3.4 shows the combination of membership functions μ_1 and μ_2.

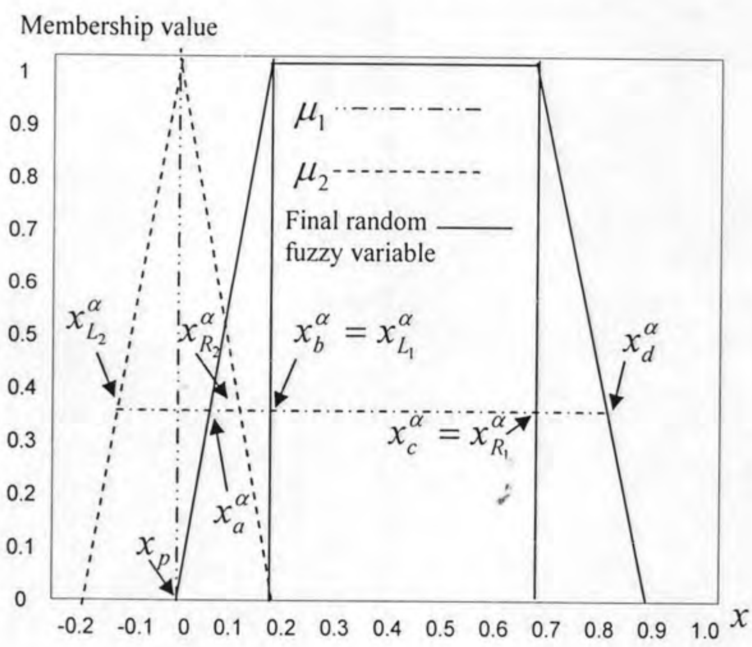


Figure 3.4 Combination of membership functions μ_1 and μ_2

The random–fuzzy set can be interpreted by linguistic terms such as "most likely about D_m or between D_{m1} and D_{m2}, but not less than D_1 and not greater that D_2. If the internal membership function is nil, then the random–fuzzy set is certainly a triangular membership function. If the internal membership function is presented by the definite internal interval (i.e., rectangular, trapezoidal, triangular), then the random–fuzzy set can be represented by trapezoidal and triangular membership functions. The triangular and trapezoidal membership functions are shown as follows:

$$\mu_{\overline{D}}(x)=\begin{cases}\dfrac{x-D_1}{D_{m1}-D_1}, D_1 \le x \le D_{m1} \text{ and } & D_1 \ne D_{m1}\\ 1 \qquad , D_{m1} \le x \le D_{m2} & \\ \dfrac{x-D_1}{D_{m2}-D_2}, D_{m2} \le x \le D_2 \text{ and } & D_{m2} \ne D_2 \\ 0, \qquad otherwise & \end{cases} \tag{3.17}$$

When $D_{m1} = D_{m2}$, a trapezoidal fuzzy set is transformed into a triangular fuzzy set. The trapezoidal fuzzy set is denoted by $(D_1, D_{m1}, D_{m2}, D_2)$, while the triangular fuzzy set is denoted by $(D_1, D_{m1}, D_{m1}, D_2)$ or (D_1, D_{m1}, D_2).

To take an advantage of the simulation data associated with duration of project activities derived from different types of probability distributions, the membership functions representing the random part of activity duration and risk variables are alternatively developed by using a neurofuzzy metamodel (a neurofuzzy trained on simulation data). The neurofuzzy metamodel is able to capture nonlinear relationships between duration of project activities and risk variables. The spreads of durations of project activities provided by using the knowledge network of the neurofuzzy metamodel depend on training data.

3.4 Neurofuzzy Metamodel for Probability–Possibility Transformation

This section is to describe the development of a membership function representing the random part of a random–fuzzy variable by using the neurofuzzy metamodel. The neurofuzzy metamodel is a neurofuzzy system trained on simulation data to assist in fuzzy modeling which is associated with three issues: (1) problem representation and data acquisition, (2) structure and parameter identification of the fuzzy rule base, and (3) interpretation of the resulting fuzzy model.

The use of the neurofuzzy metamodel is justified because simulation is able to create a number of experiments to investigate behaviors of random variables and relationships between random variables. Sufficient simulation data are evenly distributed which assists in effectively using the optimization algorithms for training a neurofuzzy system. Thus, the neurofuzzy metamodel reasonably become preferable to estimating duration of project activities affected by risk factors accounting for their randomness.

The neurofuzzy network can be regarded both as an adaptive fuzzy inference system with the capability of learning fuzzy rules from data, and as a connectionist architecture provided with linguistic meaning. It offers the interpretability and transparency in the resulting models which cannot be obtained from simulation. The neurofuzzy metamodel can be also considered as an approximation version of a simulation model, which can be displayed in the form of the fuzzy IF – THEN rules.

The neurofuzzy metamodel represents variables in a simulation model by membership functions. This study therefore uses the neurofuzzy metamodel as a probability–possibility transformation method to create the membership function presenting a random part of a random–fuzzy variable. In the developed neurofuzzy metamodel, risk variables are considered as input variables. An output variable is a temporal variable (i.e., duration of project activities and alternatively project completion time). The membership function of the random part is developed by inputting new values of risk variables into the trained neurofuzzy metamodel. This study determines the values of risk variables based on three scenarios (i.e., optimistic, most likely, and pessimistic) in order to develop the triangular membership function. Using Eq. (13.4) to (13.7), the membership function of a random–fuzzy variable is established by inserting an internal membership function representing uncertainties stemming from systematic and unknown contributions into the created membership function of a random part.

To improve the performance of the trained neurofuzzy metamodel, the initially obtained membership functions are replaced by the final membership functions. The neurofuzzy system having the final membership functions of its input variables can take into account every uncertainty resulting from the random contribution and systematic and unknown contributions. The improvement on interpretability of the resulting fuzzy model is achieved by using structure reduction and local tuning of parameters. Once generalization capability of the trained network is tested, the learned knowledge network is encoded in the form of fuzzy rules which are used to process data according to reasoning principles.

This study also uses other data collection methods (i.e., the expert's questionnaire, historical record, and direct observation) to acquire data available in each construction period. Then, adapting and optimizing the topography and parameters of the neurofuzzy network are defined based on available data. A developed neurofuzzy modeling approach using multiple sources of data as presented by Figure 3.5 is composed of eight steps as follows:

1) Acquire subjective data through the experts' questionnaire
2) Acquire quantitative data through the historical record
3) Establish the initial structure and initial parameters of a knowledge network by using a subtractive clustering method to determine the proper number of fuzzy rules and membership functions (size of the network topology), together with initial values of rule parameters (network weight) and construct a rule base by a self-organized method
4) Tune the parameters of the fuzzy rules by a gradient decent technique and a recursive least squares (RLS) method to enhance the accuracy of the derived fuzzy model by finding the best fit of the training data (historical data adjusted by subjective data)
5) Improve the performance and realization of the resulting fuzzy model by replacing the fuzzy sets by the random–fuzzy sets
6) Improve the interpretability of the resulting fuzzy model by reducing the number of rules in the fuzzy rule base and fuzzy sets in fuzzy rules by applying a trial and error method
7) Process new input values through the improved neurofuzzy model
8) Acquire simulation data through the executions of simulation
9) Repeat steps 3 to 8 by replacing the subjective and historical data with the simulation data; and
10) Repeat steps 3 to 8 by replacing the historical data with the factual data

The neurofuzzy metamodel can be used to estimate duration of project activities for any given values of risk variables. It is used to provide values of parameters of the probability distributions and establish the probability distributes of duration of project activities. In this way, parameters used to develop the membership function of a random part of duration of project activities described above are employed to develop the probability distributes in order for simulation to improve its performance with the realization that duration of project activities is estimated regarding impact of risk factors.

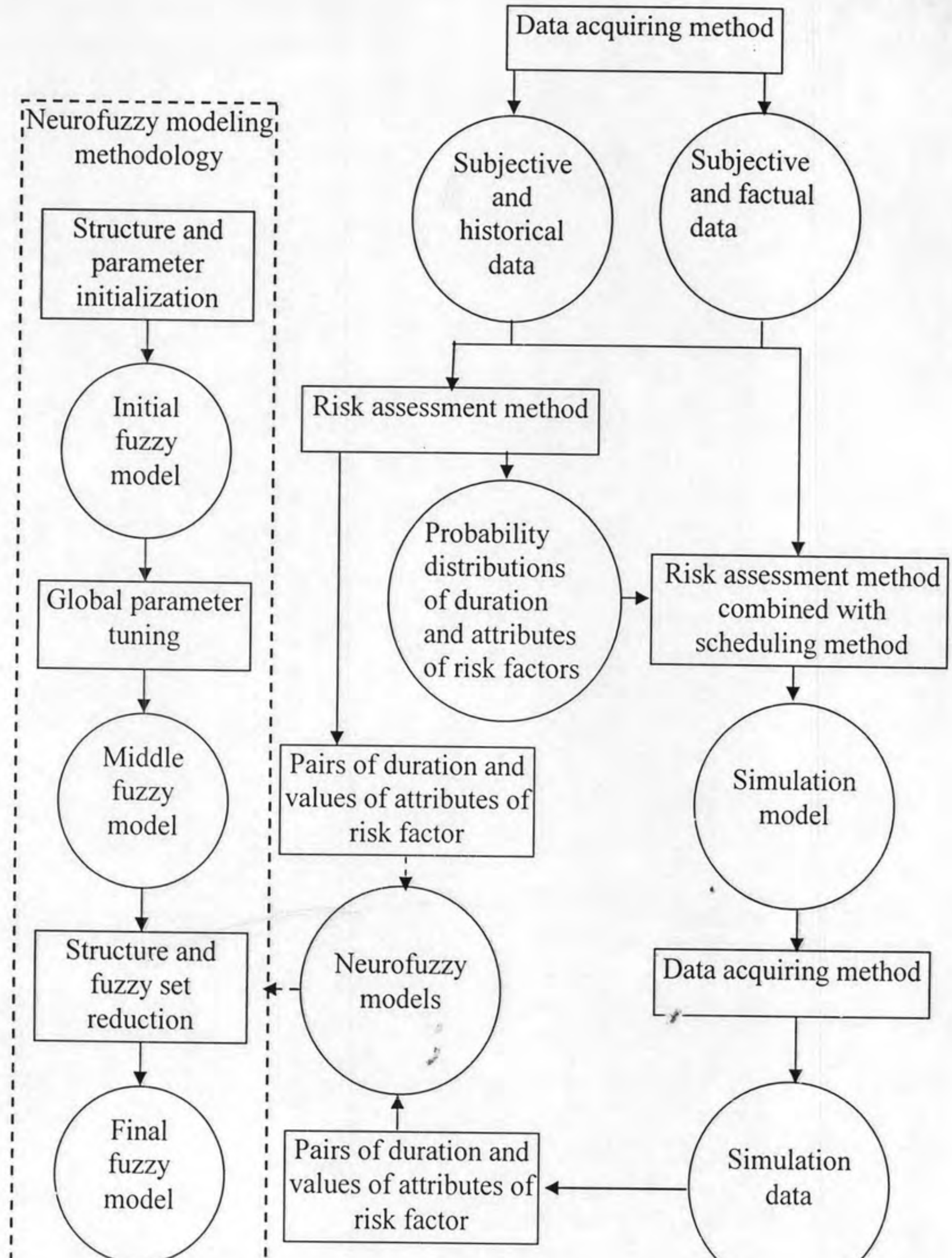


Figure 3.5 A neurofuzzy modeling approach consisting of data acquiring methods, neurofuzzy models and simulation models

3.5 Fuzzy Analytic Hierarchy Process (FAHP)

As described in previous section, the likelihood of occurrence of a risk factor is a function of weight or degree of significance of each risk factor. Typically, the weight is determined by using a multicriteria decision making (MCDM) tool (e.g., AHP). The discrete scale has a merit of simplicity and ease of the application. However, the uncertainty associated with the mapping of a decision maker's perception (or judgement) to a number cannot be precisely presented by the discrete scale. It is even more difficult for decision makers to make a comparison between risk factors of which characteristics are different. Risk factors are also affected by various factors which consequently have different influences on their category in the next higher level. When the comparison between a pair of risk factors is provided regarding the corresponding risk category in the next higher level, impacts of factors affecting each risk factor on the risk category might be quantitatively immeasurable. As a result, the decision makers cannot provide the complex and imprecise comparison by the discrete scale (Georgy, 2005).

The expert's opinions about the relative significance of risk factors are usually described by linguistic terms. The decision aids for pairwise comparisons of the AHP (data from Saaty 1980 and Büyüközkan, 2007) are modified by using the following linguistic weighting set to evaluate the significance of risk factors; {ALS; VSLS; SLS; WLS; ES; WMS; SMS; VSMS; AMS}, where ALS: absolutely less significant, VSLS: very strongly less significant, SLS: strongly less significant, WLS: weakly less significant, ES: equally significant, WMS: weakly more significant, SMS: strongly more significant, VSMS: very strongly more significant, and AMS: absolutely more significant.

This study uses a fuzzy analytic hierarchy process (FAHP) to address uncertainty in measuring weights of risk factors. The fuzzy pairwise comparison considers favorable and adverse effects of uncertainty. Fuzzy numbers are used to present uncertainty involved in the pairwise comparisons by using parameters l and u to present favorable and adverse effects of uncertainty, respectively. For example, experts may consider that the risk factor i is "Very strongly more significant" as compared with the risk factor j regarding a particular activity affected by such risk factors. The expert may set $a_{ij} = (l, m, u) = \left(2, \frac{5}{2}, 3\right)$. If the risk factor j is thought to be "Very strongly less significant" than the risk factor i, the pairwise comparison between j and i could be presented by using the fuzzy number,

$$a_{ij} = \left(\frac{1}{u}, \frac{1}{m}, \frac{1}{l}\right) = \left(\frac{1}{3}, \frac{2}{5}, \frac{1}{2}\right).$$

The decision makers are allowed to provide the comparing results by the interval judgement instead of crisp value judgement which makes the decision makers feel more convenient and confident. The FAHP integrates these individual effects of uncertainty into the pairwise comparison by combining the calculated ratio-score local priorities according to the requirement of the synthesis of priorities.

In fuzzy pairwise comparison, the relative weight is a ratio between-group sum to total sum. The final score of the pairwise comparison is represented by fuzzy numbers. In order to prioritize risk factors by their weights, a method for comparing fuzzy numbers is required. Büyüközkan's extent analysis method is applied to derive the significance weights of the risk factors. It is the integration of Chang's extent analysis method (Chang, 1996) and Zhu et al's improvement method (Zhu, 1999). This study uses this method because it provides the computational simplicity and effectiveness.

The extend analysis method is adopted to measure the extent of an object to satisfy a goal. A fuzzy number is used to represent such extent. The extent analysis of each object is performed according to definition provided based on the fuzzy principle. A fuzzy synthetic degree value which is a result of the extent analysis is derived from the following definition.

Let $X = \{x_1, x_2, ..., x_n\}$ be an object set, and $G = \{g_1, g_2, ..., g_m\}$ be a goal set. Chang's extent analysis method analyzes each object in X regarding each goal set (Chang, 1996). Therefore, each object has m extent analysis values, with the following signs:

$$\widetilde{M}_{i1}, \widetilde{M}_{i2}, ..., \widetilde{M}_{im}, i = 1,2,...,n \quad (3.18)$$

In this study, triangular fuzzy numbers are used to represent all the $\widetilde{M}_{ij}\ (i = 1,2,...,m)$. Steps in Büyüközkan's extent analysis method are given as follows:

Step 1. For m goals in a defined goal set, the value of fuzzy synthetic extent with respect to the i^{th} object is defined as

$$\widetilde{S}_i = \frac{between-group\ sum}{total\ sum} = \sum_{j=1}^{m} \widetilde{M}_{ij} \left[\sum_{i=1}^{n} \sum_{j=1}^{m} \widetilde{M}_{ij}\right]^{-1} \tag{3.19}$$

The fuzzy addition operation of m extent analysis values for a particular matrix is performed to obtain $\sum_{j=1}^{m} \widetilde{M}_{ij}$ for a given i such that

$$\sum_{j=1}^{m} \widetilde{M}_{ij} = \left(\sum_{j=1}^{m} l_j, \sum_{j=1}^{m} m_j, \sum_{j=1}^{m} u_j\right) \tag{3.20}$$

The fuzzy addition operation of $\widetilde{M}_{ij}\ (i = 1,2,...,n : j = 1,2,...m)$ values is performed to obtain $[\sum_{i=1}^{n} \sum_{j=1}^{m} \widetilde{M}_{ij}]^{-1}$ such that

$$\sum_{i=1}^{n} \sum_{j=1}^{m} \widetilde{M}_{ij} = \left(\sum_{i=1}^{n} \sum_{j=1}^{m} l_j, \sum_{i=1}^{n} \sum_{j=1}^{m} m_j, \sum_{i=1}^{n} \sum_{j=1}^{m} u_j\right) \tag{3.21}$$

and then the inverse of the vector in (3.22) is calculated such that

$$[\sum_{i=1}^{n} \sum_{j=1}^{m} \widetilde{M}_{ij}]^{-1} = \left(1/\sum_{i=1}^{n} \sum_{j=1}^{m} l_j, 1/\sum_{i=1}^{n} \sum_{j=1}^{m} m_j, 1/\sum_{i=1}^{n} \sum_{j=1}^{m} u_j\right) \tag{3.22}$$

The derived fuzzy synthetic values are used to calculate the weight vectors of all elements for each level of the hierarchy by applying the principles for the comparison of fuzzy numbers.

Step 2. Using the extension principle, the degree of possibility of $\widetilde{S}_1 \geq \widetilde{S}_2$ is defined as

$$V(\widetilde{S}_1 \geq \widetilde{S}_2) = \sup_{x \geq y} \left[\min\left\{\mu_{\widetilde{S}_1}(x), \mu_{\widetilde{S}_2}(y)\right\}\right] \tag{3.23}$$

when a pair (x, y) exists such that $x \geq y$ and $\mu_{\widetilde{S}_1}(x) = \mu_{\widetilde{S}_2}(y) = 1$, then $V(\widetilde{S}_1 \geq \widetilde{S}_2) = 1$. Since $\widetilde{S}_1 = (l_1, m_1, u_1)$ and $\widetilde{S}_2 = (l_2, m_2, u_2)$ are convex fuzzy numbers so that

$$V(\widetilde{S}_1 \geq \widetilde{S}_2) = \mu_{\widetilde{S}_1}(d) = \begin{cases} 1, m_1 \geq m_2 \\ 0, l_2 \geq \mu_1 \\ \dfrac{l_2 - \mu_1}{(m_1 - u_1) - (m_2 - l_2)}, otherwise \end{cases} \tag{3.24}$$

where d is the ordinate of the highest intersection point between $\mu_{\tilde{S}_1}$ and $\mu_{\tilde{S}_2}$. The values of $V(\tilde{S}_1 \geq \tilde{S}_2)$ and $V(\tilde{S}_2 \geq \tilde{S}_1)$ are used to compare $\tilde{S}_1$ and $\tilde{S}_2$

Step 3. The degree of possibility for a convex fuzzy number $\tilde{S}$ to be greater than k convex fuzzy numbers $\tilde{S}_i (i = 1,..., k)$ can be defined by

$$V(\tilde{S} \geq \tilde{S}_1, \tilde{S}_2,..., \tilde{S}_k) = V((\tilde{S} \geq \tilde{S}_1) and (\tilde{S} \geq \tilde{S}_2) and ... and (\tilde{S} \geq \tilde{S}_k)) \qquad (3.25)$$

$$= \min_k \{V(\tilde{S} \geq \tilde{S}_k)\}$$

Assume that

$$d'(x_i) = \min_k V(\tilde{S}_i \geq \tilde{S}_k) \qquad (3.26)$$

For $k = 1,2,..., n; k \neq i$. the weight vector is given by

$$W' = (d'(x_1), d'(x_2),..., d'(x_n))^T \qquad (3.27)$$

where $x_i (i = 1,2,..., n)$ are n objects.

Step 4. Via normalization, the normalized weight vectors are,

$$W = (d(x_1), d(x_2),..., d(x_n))^T \qquad (3.28)$$

where W is not a fuzzy number.

3.6 Risk Assessment Integrated within Random–Fuzzy Network Scheduling Method (RAIRFNET)

This study develops a method called a risk assessment integrated within random-fuzzy network scheduling method (RAIRFNET) to address the nondeterministic risk based scheduling problem. The random–fuzzy variables are used to represent the random and fuzzy parts of duration of project activities. The activity duration is presented in the form of either triangular or trapezoidal fuzzy numbers which provide the stimulus for applying the RIARFNET modified from the FNET proposed by Lorterapong and Moselhi (1996). Instead of computing the schedule based only on the internal mechanism of the fuzzy arithmetic algorithms used in the FNET, a mathematic for combining the random contribution to aleatory uncertainty based only on the fundamentals of probability theory is used to recalculate the schedule. In RAIRFNET, appropriate mathematics are used to address the corresponding uncertainties. Specifically, the random part of a random–fuzzy variable is analyzed by using the probability theory, while the fuzzy part is analyzed by the fuzzy theory. After performing the risk assessment and representing risk and temporal variable values by fuzzy numbers, the activity duration preliminary estimated based on the optimistic forecasts (made without determining impacts of risk factors) or the normal progress (made based on expected conditions) and remaining duration of the activity duration is revised to reflect the impacts of risk factors. Each portion of activity durations representing the impacts of a particular risk factor influencing a considered activity is given by using defined equations of L , C , and E as explained in the previous section. The proposed RAIRFNET provides project managers with a method that is capable of measuring the uncertainties inherent in the nature of the risk analysis process. The project managers are allowed to perform the risk assessments using linguistic terms, which can be appropriately put into the actual practices of carrying out the construction. Thus, different portions of durations of project activities are summed up by applying a fuzzy addition operation with the consideration about points of times that the particular portions occur and activity's start constraints, which depend on the technical dependencies among project activities. Next, project duration is computed by applying forward and backward path calculations such as those performed in FNET, but they are modified by using mathematics for random–fuzzy variables.

To get the start time of activity i or $T^S(i)$, the pairwise comparisons for each and every pair of fuzzy available times $\overline{T}^A(i,j) = (T_1^A(i,j), T_{m1}^A(i,j), T_{m2}^A(i,j), T_2^A(i,j))$ of entities (i.e., logical dependencies) of activity i are performed by computing the credibility coefficients C_{gr}, C_{lo}, and C_{eq}. The calculation of fuzzy start times in the forward pass is determined by comparing fuzzy early finish times of precedence activities. The pairwise comparisons among all finish fuzzy durations $\overline{T}^{EE}(k) = (T_1^{EE}(k), T_{m1}^{EE}(k), T_{m2}^{EE}(k), T_2^{EE}(k))$ of preceding activities are performed by computing the credibility coefficients C_{gr}, C_{lo}, and C_{eq}. The early finish time is determined by adding duration of a considered activity to its early start time as shown in Eq (3.29) to (3.30). Finally, the fuzzy project completion time

$\overline{T}^{EE}(e) = (T_1^{EE}(e), T_{m1}^{EE}(e), T_{m2}^{EE}(e), T_2^{EE}(e))$ is calculated, where e is the last activity of a project.

$$FES_x = \underset{p \in P}{\tilde{\max}}(FEF_p) \tag{3.29}$$

$$FEF_x = FES_x + D_x \tag{3.30}$$

where FES_x is the fuzzy early start of activity x; p is the predecessor of x; P is the set of predecessors, FEF_x is the fuzzy early finish; D_x is the fuzzy duration.

As reviewed, realistic fuzzy results must adhere to the following conditions: 1) scheduling constraints generated from the calculations of forward and backward passes must be satisfied, 2) no negative quadruple for the activity late times is permitted, 3) the value of each element in the calculated quadruple does not exceed its successor (i.e., for a quadruple $(D_1, D_{m1}, D_{m2}, D_2)$; $D_1 \le D_{m1} \le D_{m2} \le D_2$). The RAIRFNET employs fuzzy bounds to facilitate the calculations to satisfy these three conditions simultaneously. It is reasonable to utilize fuzzy bounds because of its flexibility with regard to the greater allowable time – window for calculating fuzzy late times. In addition, a violation of the constraints produced from both forward and backward pass calculations is not found by using this time – window.

As duration of project activities are represented by random–fuzzy numbers, the project completion time is accordingly expressed by a random–fuzzy number. The project completion time is calculated by using the fuzzy bound to meet the required three – conditions. The mathematics of sum and difference operations for random–fuzzy variables are determined in the following section as they are used in the forward and backward pass calculations.

The upper and lower bound of a fuzzy number are provided in a simply manner. For example, a quadruple $(D_1, D_{m1}, D_{m2}, D_2)$ characterizes a trapezoidal fuzzy number of duration D_x of which the upper bound named by D^u can be obtained by subtracting an interval $[0, \infty)$, expressed in terms of the quadruple as $(0, 0, \infty, \infty)$ from D_x. Fuzzy lower bound of D_x represented by D^l can be obtained by adding the quadruple $(0, 0, \infty, \infty)$ to D_x.

The calculation of fuzzy late times in the backward pass is facilitated by fuzzy upper and lower bounds. At first, the preliminary late finish (PLF) is determined by the computations of fuzzy late times of each activity. The PLF can be computed for any activity x as:

$$PLF_x = \underset{s \in S}{\tilde{\min}}(FLS_s) \tag{3.31}$$

where FLS_s is the fuzzy late start time of its successor s; S is the set of the successors of activity x.

The PLF is converted to the upper bound for the late finish (FLF^u) by subtracting an interval $[0, \infty)$ from PLF. Then, the fuzzy late finish time (FLF^u) can be calculated based on the assumption previously mentioned. Let the FEF be represented by (a,b,c,d), and the FLF^u be characterized by $(-\infty,-\infty,e,f)$, the fuzzy late times for project activities can be calculated by taking steps explained below.

- Determination of a greater uncertainty (i.e., the larger right spread) of the two fuzzy quantities by comparing the results of ($f-e$) and ($d-c$).
- Calculation for Y is performed to find the largest fuzzy number that meets the following condition.
- Calculation for FLF of project activities is carried out as follows:

- $FLF = FEF \oplus Y$ (3.32)

- Computation of the fuzzy late start FLS is performed by substituting FLF and D into the following equation.

- $FLS \oplus D = FLF$ (3.33)

In addition, fuzzy bounds are used to calculate the fuzzy start (FS) and fuzzy finish (FF) times for each activity.

$$FS_x = FES_x^l \cap FLS_x^u \tag{3.34}$$

$$FF_x = FEF_x^l \cap FLF_x^u \tag{3.35}$$

The RAIRFNET is different from that of FNET and probabilistic methods (i.e., simulation). The FNET completely ignores the fact that duration of project activities influenced by uncertainty due to the random contribution. The random contribution should be analyzed by using the probability theory, but the mathematics for the random contribution are not applied to the calculation of the project duration by using FNET. The probabilistic methods on the other hand neglect to examine uncertainty due to systematic and unknown contributions. The difference between the RAIRFNET and PERT is that PERT determines the expected time values without considering the standard deviation. Either expected time values or their standard deviations are determined in the RAIRFNET. Therefore, the RAIRFNET conceptually and practically provides a more realistic solution than that of FNET, probabilistic methods and PERT.

3.7 Mathematics for Random–Fuzzy Activity Duration

This section is to describe the sum and difference operation of the random–fuzzy variables used in the network calculation. $T^S(i), D(i)$ denote the random–fuzzy variable and $\lfloor T_1^{S,\alpha}, T_2^{S,\alpha}, T_3^{S,\alpha}, T_4^{S,\alpha} \rfloor$ and $\lfloor D_1^{\alpha}, D_2^{\alpha}, D_3^{\alpha}, D_4^{\alpha} \rfloor$ are the generic $\alpha-$ cuts and C denotes the result referring $\lfloor T_1^{EE,\alpha}, T_2^{EE,\alpha}, T_3^{EE,\alpha}, T_4^{EE,\alpha} \rfloor$ as the generic $\alpha-$ cuts. $T^S(i), D(i)$, and $T^{EE}(i)$ represent FES_x, D_x, and FEF_x discussed above. As mentioned previously, the product operation as shown in Eq (3.15) and (3.13) and the division operation (i.e., w/p) are used in the RAIRFNET. The $l_{i(j)}$, $c_{i(j)}$, and $e_{i(j)}$ are considered as the random–fuzzy variable. $\lfloor l_1^{\alpha}, l_2^{\alpha}, l_3^{\alpha}, l_4^{\alpha} \rfloor$, $\lfloor c_1^{\alpha}, c_2^{\alpha}, c_3^{\alpha}, c_4^{\alpha} \rfloor$ and $\lfloor e_1^{\alpha}, e_2^{\alpha}, e_3^{\alpha}, e_4^{\alpha} \rfloor$ are the generic $\alpha-$ cuts.

Effects of systematic and unknown contributions to epistemic uncertainty represented by an internal membership function are determined through four categories: (1) nil internal membership function, (2) rectangular internal membership function, (3) symmetrical trapezoidal internal membership function, and (4) symmetrical triangular internal membership function. For the first and forth categories, the external membership functions are represented by triangular fuzzy sets, while the trapezoidal fuzzy sets are used to present the external membership function of the second and third categories. The triangular fuzzy set can be presented in the form of a quadruple $(D_1, D_{m1}, D_2) = (D_1, D_{m1}, D_{m2}, D_2)$. If the activity durations can be certainly estimated, they are represented by crisp numbers, and the activity durations are represented by triangular or trapezoidal fuzzy sets, that is $D_1 = (D_1, D_1, D_1, D_1)$, which provides the ability to apply the mathematics for random–fuzzy variables.

To present the mathematics for random–fuzzy variable, there are three situations in association with the shapes of membership functions. The first situation considers only the random contributions of random–fuzzy variables for every level of the α - cut in the range $[0,1]$,

$$T_2^{S,\alpha} = T_3^{S,\alpha} \tag{3.36}$$

$$D_2^{\alpha} = D_3^{\alpha} \tag{3.37}$$

The combination of all values in interval $\lfloor T_1^{S,\alpha}, T_4^{S,\alpha} \rfloor$ and all values in interval $\lfloor D_1^{\alpha}, D_4^{\alpha} \rfloor$ can be randomly made. The support of the final distribution is then given by

$$\lfloor T_1^{S,\alpha} + D_1^{\alpha}, T_4^{S,\alpha} + D_4^{\alpha} \rfloor \tag{3.38}$$

The membership function μ_I is established, where I represents interval (determined from intervals). The normal probability distribution is transformed into the membership function by applying the probability–possibility transformation. Therefore, the mean (m) and standard deviation (σ) of the membership functions which are referred to as the normal possibility distributions $T^S(i), D(i)$ are given by

$$m = \frac{\left(T_1^{S,\alpha=1} + D_1^{\alpha=1}\right) + \left(T_4^{S,\alpha=1} + D_4^{\alpha=1}\right)}{2} \tag{3.39}$$

$$\sigma = \frac{1}{3}\min\left\{m - T_1^{S,\alpha=0} - D_1^{\alpha=0}; T_4^{S,\alpha=0} + D_4^{\alpha=0} - m\right\} \tag{3.40}$$

where the confidence interval at level of confidence 1 is supposed to be the $\pm 3\sigma$ interval around the mean value. The membership function μ_N is built from Eq. (3.39) and (3.40), where N represents normal. A normal probability distribution can be obtained from the sum of probability distributions. Thus, the sum of possibility distributions which represent the external membership functions of the corresponding random–fuzzy variables much tend to the normal possibility distribution. The membership function of the result $T^S(i) + D(i)$ is consequently determined by using the average operators to $\mu_I(x)$ and $\mu_N(x)$ for every value x belonging to the universal set. The membership grade c of the final membership function $T^S(i)+D(i)$ is defined as

$$T^{EE} = k \cdot ext(\mu_N, \mu_i) + (k-1) \cdot \text{int}(\mu_N, \mu_i) \tag{3.41}$$

where $ext(\mu_N, \mu_i) = ext(k \cdot \mu_i + (k-1) \cdot \mu_N)$, $\text{int}(\mu_N, \mu_i) = \text{int}(k \cdot \mu_i + (k-1) \cdot \mu_N)$, and $k = \frac{1}{\sqrt{2}}$.

The differences between all values in interval $\lfloor T_1^{S,\alpha}, T_4^{S,\alpha} \rfloor$ and all values in interval $\lfloor D_1^{\alpha}, D_4^{\alpha} \rfloor$ are drawn randomly so as to find the support of the final distribution which is shown as follows

$$\lfloor T_1^{S,\alpha} - D_1^{\alpha}, T_4^{S,\alpha} - D_4^{\alpha} \rfloor \tag{3.42}$$

The development of a membership function μ_I is determined, where I represents interval (considered from intervals). The membership function is transformed from the normal probability distribution by applying the probability–possibility transformation. Therefore, the normal possibility distributions $T^S(i)$ and $D(i)$ controlled by the mean and standard deviation are given by

$$m = \frac{\left(T_1^{S,\alpha=1} - D_1^{\alpha=1}\right) + \left(T_4^{S,\alpha=1} - D_4^{\alpha=1}\right)}{2} \tag{3.43}$$

$$\sigma = \frac{1}{3} \min\left\{m - T_1^{S,\alpha=0} + D_4^{\alpha=0}; T_4^{S,\alpha=0} + D_1^{\alpha=0} - m\right\} \tag{3.44}$$

where the confidence interval at level of confidence 1 is supposed to be the $\pm 3\sigma$ interval around the mean value. The membership function μ_N is built from Eq. (3.43) and (3.40), where N represents normal. The difference of probability distributions is performed for establishing a normal probability distribution. The difference of possibility distributions presenting the external membership functions of the corresponding random–fuzzy variables, therefore, tends to the normal possibility distribution. The membership function of the result $T^S(i) - D(i)$ can be obtained by using the average operators to $\mu_I(x)$ and $\mu_N(x)$ for every value x belonging to the universal set. At last, the membership function of the result $T^S(i) - D(i)$ is determined by applying Eq. (3.41) to these two – defined membership function $\mu_I(x)$ and $\mu_N(x)$. This leads to the membership function of μ_{T-D}

The second situation considers only nonrandom contribution for every level of the α - cut in the range $[0,1]$, one or both random–fuzzy variables

$$T_1^{S,\alpha} = T_2^{S,\alpha} \tag{3.45}$$

$$D_1^{\alpha} = D_2^{\alpha} \tag{3.46}$$

and/or

$$T_3^{S,\alpha} = T_4^{S,\alpha} \tag{3.47}$$

$$D_3^{\alpha} = D_4^{\alpha} \tag{3.48}$$

The mathematics of sum is expressed as

$$T_1^{EE,\alpha} = T_1^{S,\alpha} + D_1^{\alpha} \tag{3.49}$$

$$T_2^{EE,\alpha} = T_2^{S,\alpha} + D_2^{\alpha} \tag{3.50}$$

$$T_3^{EE,\alpha} = T_3^{S,\alpha} + D_3^{\alpha} \tag{3.51}$$

$$T_4^{EE,\alpha} = T_4^{S,\alpha} + D_4^{\alpha} \tag{3.52}$$

The mathematics of difference is expressed as

$$T_1^{EE,\alpha} = T_1^{S,\alpha} - D_1^{\alpha} \tag{3.53}$$

$$T_2^{EE,\alpha} = T_2^{S,\alpha} - D_2^{\alpha} \tag{3.54}$$

$$T_3^{EE,\alpha} = T_3^{S,\alpha} - D_3^{\alpha} \tag{3.55}$$

$$T_4^{EE,\alpha} = T_4^{S,\alpha} - D_4^{\alpha} \tag{3.56}$$

The mathematics of product is expressed as

$$T_1^{EE,\alpha} = \min(T_{r1}^{S,\alpha} D_{r1}^{\alpha}; T_{r1}^{S,\alpha} D_{r4}^{\alpha}; T_{r4}^{S,\alpha} D_{r1}^{\alpha}; T_{r4}^{S,\alpha} D_{r4}^{\alpha}) \tag{3.57}$$

$$T_2^{EE,\alpha} = \min(T_2^{S,\alpha} D_2^{\alpha}; T_2^{S,\alpha} D_3^{\alpha}; T_3^{S,\alpha} D_2^{\alpha}; T_3^{S,\alpha} D_3^{\alpha}) \tag{3.58}$$

$$T_3^{EE,\alpha} = \max(T_2^{S,\alpha} D_2^{\alpha}; T_2^{S,\alpha} D_3^{\alpha}; T_3^{S,\alpha} D_2^{\alpha}; T_3^{S,\alpha} D_3^{\alpha}) \tag{3.59}$$

$$T_4^{EE,\alpha} = \max(T_{r1}^{S,\alpha} D_{r1}^{\alpha}; T_{r1}^{S,\alpha} D_{r4}^{\alpha}; T_{r4}^{S,\alpha} D_{r1}^{\alpha}; T_{r4}^{S,\alpha} D_{r4}^{\alpha}) \tag{3.60}$$

The third situation determines either the random part or the systematic and unknown part is presented in the considered random–fuzzy variables. The third situation is divided into two conditions including (1) the random–fuzzy variables of which the internal membership function is not rectangular and (2) the random–fuzzy variables of which the internal membership function is rectangular having the following properties:

$$T_2^{S,\alpha} = T_2^{S,\alpha} \tag{3.61}$$

$$D_2^{\alpha} = D_2^{\alpha'} \tag{3.62}$$

and/or

$$T_3^{S,\alpha} = T_3^{S,\alpha'} \tag{3.63}$$

$$D_3^{\alpha} = D_3^{\alpha'} \tag{3.64}$$

for every α and α' in the range $[0,1]$.

In the difference of two random–fuzzy variables, the same kind of contributions composes together obeying to their own composition rules. Firstly, the random part of the random fuzzy variables must be separated from the internal part as the two parts are processed differently. The pure random part is determined starting from the given random–fuzzy variable as shown in Figure 3.6

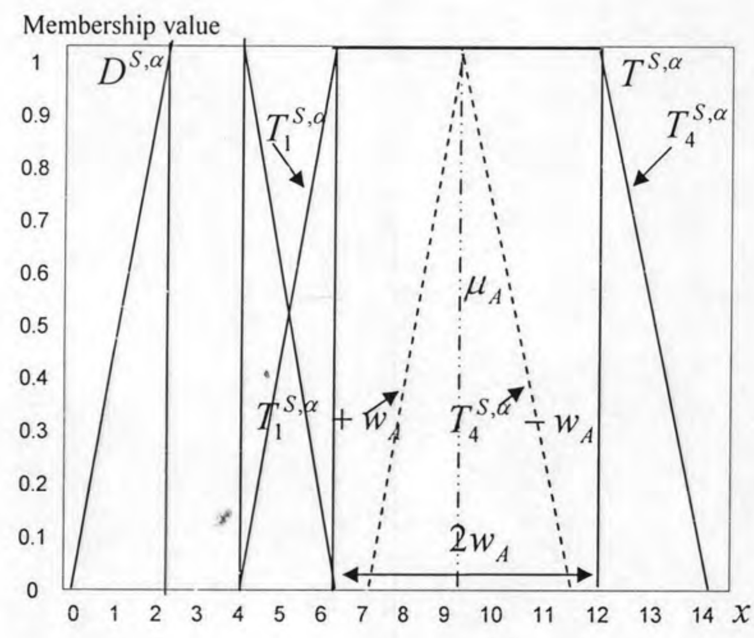


Figure 3.6 Determination of the pure random part

By definition, the mean value of the random–fuzzy variable $T^S(i)$ is determined as the mean value of the α – cut at level $\alpha = 1$:

$$\mu_A = \frac{T_2^{S,\alpha=1} + T_3^{S,\alpha=1}}{2} \tag{3.65}$$

Under the hypothesis of rectangular internal membership function, for each α – cut which varies with level α, the semi – width of the left part is

$$\omega_{Al}^{\alpha} = \mu_A - T_2^{S,\alpha} \tag{3.66}$$

and the semi – width of the right part is

$$\omega_{Ar}^{\alpha} = T_3^{S,\alpha} - \mu_A \tag{3.67}$$

and for each $\alpha -$cut which is constant for every level α, the semi – width of the left and right parts is

$$\omega_A = \frac{T_3^{S,\alpha=1} - T_2^{S,\alpha=1}}{2} \tag{3.68}$$

Hence, for each $\alpha -$cut which varies according to level α, the pure random part of the given random–fuzzy variable whose $\alpha -$cuts are

$$\left[T_1^{S,\alpha} + \omega_{Al}^{\alpha}; T_4^{S,\alpha} - \omega_{Ar}^{\alpha}\right] \tag{3.69}$$

and the pure random part of the given random–fuzzy variable whose $\alpha -$cuts for each $\alpha -$cut which is constant for every level α are

$$\left[T_1^{S,\alpha} + \omega_A; T_4^{S,\alpha} - \omega_A\right] \tag{3.70}$$

Let ω_A and ω_B be the semi – widths of the internal fuzzy variables of $T^S(i), D(i)$, respectively:

The mathematics of sum is expressed as

$$\omega_A = \frac{T_3^{S,\alpha=1} - T_2^{S,\alpha=1}}{2} \tag{3.71}$$

$$\omega_B = \frac{D_3^{\alpha=1} - D_2^{\alpha=1}}{2} \tag{3.72}$$

$$T_{r1}^{S,\alpha} = T_1^{S,\alpha} + \omega_A \tag{3.73}$$

$$T_{r4}^{S,\alpha} = T_4^{S,\alpha} - \omega_A \tag{3.74}$$

$$D_{r1}^{\alpha} = D_1^{\alpha} + \omega_B \tag{3.75}$$

$$D_{r4}^{\alpha} = D_4^{\alpha} - \omega_B \tag{3.76}$$

$$T_1^{EE,\alpha} = T_2^{EE,\alpha} - \mu_c + k \cdot ext\left(T_{r1}^{S,\alpha} + D_{r1}^{\alpha}, g_1^{\alpha}\right) + (1-k)\cdot \mathrm{int}(T_{r1}^{S,\alpha} + D_{r1}^{\alpha}, g_1^{\alpha}) \tag{3.77}$$

$$T_2^{EE,\alpha} = T_2^{S,\alpha} + D_2^{\alpha} \tag{3.78}$$

$$T_3^{EE,\alpha} = T_3^{S,\alpha} + D_3^{\alpha} \tag{3.79}$$

$$T_4^{EE,\alpha} = T_3^{EE,\alpha} - \mu_c + k \cdot ext\left(T_{r4}^{S,\alpha} + D_{r4}^{\alpha}, g_4^{\alpha}\right) + (1-k)\cdot \mathrm{int}(T_{r4}^{S,\alpha} + D_{r4}^{\alpha}, g_4^{\alpha}) \tag{3.80}$$

where μ_c is the mean value of the sum:

$$\mu_c = \frac{T_2^{EE,\alpha=1} + T_3^{EE,\alpha=1}}{2} \tag{3.81}$$

k is a constant

$$k = \frac{1}{\sqrt{2}} \quad (3.82)$$

and $\lfloor g_1^{\alpha}, g_4^{\alpha} \rfloor$ is the generic $\alpha -$ cut of the normal possibility distribution, having a mean value μ_c and standard deviation:

$$\sigma = \frac{1}{3}\min\left\{\mu_c - T_{r1}^{S,\alpha=0} - D_{r1}^{\alpha=0}; T_{r4}^{S,\alpha=0} + D_{r4}^{\alpha=0} - \mu_c\right\} \quad (3.83)$$

The mathematics of difference is expressed as

$$T_1^{EE,\alpha} = T_2^{EE,\alpha} - \mu_c + k \cdot ext\left(T_{r1}^{S,\alpha} - D_{r1}^{\alpha}, g_1^{\alpha}\right) + (1-k)\cdot \text{int}(T_{r1}^{S,\alpha} - D_{r1}^{\alpha}, g_1^{\alpha}) \quad (3.84)$$

$$T_2^{EE,\alpha} = T_2^{S,\alpha} - D_2^{\alpha} \quad (3.85)$$

$$T_3^{EE,\alpha} = T_3^{S,\alpha} - D_3^{\alpha} \quad (3.86)$$

$$T_4^{EE,\alpha} = T_3^{EE,\alpha} - \mu_c + k \cdot ext\left(T_{r4}^{S,\alpha} - D_{r4}^{\alpha}, g_4^{\alpha}\right) + (1-k)\cdot \text{int}(T_{r4}^{S,\alpha} - D_{r4}^{\alpha}, g_4^{\alpha}) \quad (3.87)$$

Where μ_c is the mean value of the difference:

$$\mu_c = \frac{T_2^{EE,\alpha=1} + T_3^{EE,\alpha=1}}{2} \quad (3.88)$$

k is a constant

$$k = \frac{1}{\sqrt{2}} \quad (3.89)$$

and $\lfloor g_1^{\alpha}, g_4^{\alpha} \rfloor$ is the generic $\alpha -$ cut of the normal possibility distribution, having a mean value μ_c and standard deviation:

$$\sigma = \frac{1}{3}\min\left\{\mu_c - T_{r1}^{S,\alpha=0} + D_{r1}^{\alpha=0}; T_{r4}^{S,\alpha=0} - D_{r4}^{\alpha=0} - \mu_c\right\} \quad (3.90)$$

The mathematics of product is expressed as

$$T_1^{EE,\alpha} = T_2^{EE,\alpha} - \mu_r + T_{r1}^{EE,\alpha} \quad (3.91)$$

$$T_2^{EE,\alpha} = \min(T_2^{S,\alpha} D_2^{\alpha}; T_2^{S,\alpha} D_3^{\alpha}; T_3^{S,\alpha} D_2^{\alpha}; T_3^{S,\alpha} D_3^{\alpha}) \quad (3.92)$$

$$T_3^{EE,\alpha} = \max(T_2^{S,\alpha} D_2^{\alpha}; T_2^{S,\alpha} D_3^{\alpha}; T_3^{S,\alpha} D_2^{\alpha}; T_3^{S,\alpha} D_3^{\alpha}) \quad (3.93)$$

$$T_4^{EE,\alpha} = T_3^{EE,\alpha} - \mu_r + T_{r4}^{EE,\alpha} \quad (3.94)$$

where

$$T_{r1}^{EE,\alpha} = \min(T_{r1}^{S,\alpha} D_{r1}^{\alpha}; T_{r1}^{S,\alpha} D_{r4}^{\alpha}; T_{r4}^{S,\alpha} D_{r1}^{\alpha}; T_{r4}^{S,\alpha} D_{r4}^{\alpha}) \quad (3.95)$$

$$T_{r4}^{EE,\alpha} = \max(T_{r1}^{S,\alpha} D_{r1}^{\alpha}; T_{r1}^{S,\alpha} D_{r4}^{\alpha}; T_{r4}^{S,\alpha} D_{r1}^{\alpha}; T_{r4}^{S,\alpha} D_{r4}^{\alpha}) \quad (3.96)$$

and

$$\mu_r = \frac{T_{r1}^{EE,\alpha=1} + T_{r4}^{EE,\alpha=1}}{2} \tag{3.97}$$

When the internal membership function of the random–fuzzy variables is not rectangular, the width of the internal intervals varies regarding level α. In the more complicated situation when the internal fuzzy variable could be asymmetric, the width of the left side of the interval is consequently different from the width of the right side. For the random–fuzzy variable $T^S(i)$, the width of the left and right part are

$$\omega_{Al}^{\alpha} = \mu_A - T_2^{S,\alpha} \tag{3.98}$$

$$\omega_{Ar}^{\alpha} = T_3^{S,\alpha} - \mu_A \tag{3.99}$$

The pure random part of the random–fuzzy variable then becomes a fuzzy variable having

$$T_{r1}^{S,\alpha} = T_1^{S,\alpha} + \omega_A \tag{3.100}$$

$$T_{r4}^{S,\alpha} = T_4^{S,\alpha} - \omega_A \tag{3.101}$$

In this case, the calculation of the pure random part of other random–fuzzy variables is made following above steps. The sum of the random–fuzzy variables $T^S(i)$ and $D(i)$ is analogous to those given in the previous situations.

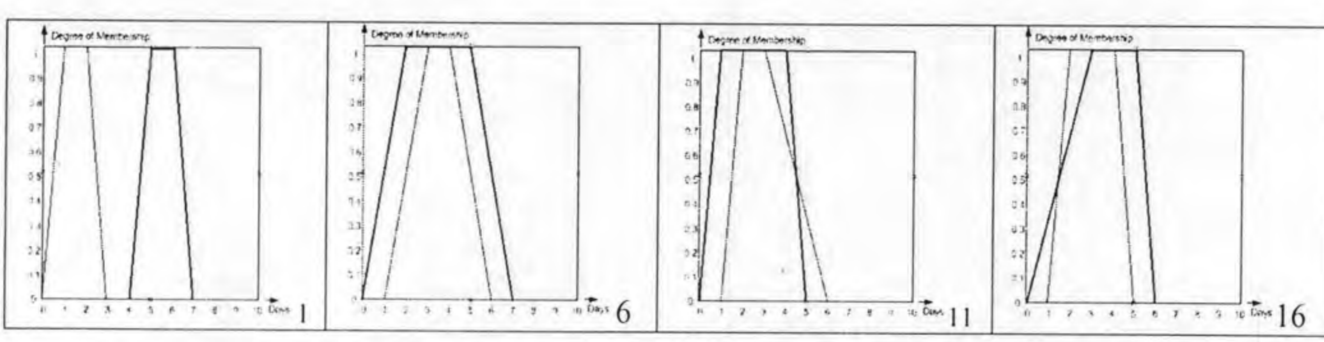
3.8 Comparison between Fuzzy–Random Variables

This section is to determine the $\max_{j=1,\ldots,n}\{T^A(i,j)\}$ in association with Eq (3.30) and $\min_{j=1,\ldots,n}\{T^A(i,j)\}$ in association with Eq (3.32) or maximum and minimum available time of entity j at activity i in association with Eq (3.30) and (3.32), respectively. Many different methods are available in the literature to compare fuzzy variables. The Yager method, Kerre method, and Nakamura method are known as suitable fuzzy ranking methods proposed essentially to solve problems in the measurement field (Klir, 2005).

Normally, in the fuzzy ranking method, the fuzzy numbers are converted into a series of crisp numbers and then the transformed crisp numbers are compared (Chen, 1985). However, some information (e.g., subjectivity or vagueness) presented in the form of a fuzzy number might lose in the transforming process (Wang, 1997).

When the fuzzy sets are overlapped or intersected as presented in Figure 3.7 (2 to 18), the fuzzy ranking problems are even more serious. Other disadvantages of the existing fuzzy ranking methods are presented in Wang (1997), such as inconsistency with human intuitive, indiscrimination and difficulty of interpretation, which in turn shows that the best fuzzy ranking method is not existent. Salicone (2007) also compared several fuzzy ranking methods. It was found out that some methods cannot succeed in quantifying the grade of belief related to the selected statement. Consequently, they do not succeed in providing an unambiguous rule for ordering fuzzy variables.

This study determines the intersection area between the fuzzy variable A and B as shown in Figure 3.8 and uses three ratios $\frac{d_1 + d_4}{Un(A,B)}$, $\frac{d_2 + d_3}{Un(A,B)}$, and $\frac{Int(A,B)}{Un(A,B)}$ proposed by Salicone (2007) to assess that fuzzy variable A is lower or greater than B. The union area is obtained by $Un(A,B) = d_1 + d_2 + d_3 + d_4 + Int(A,B)$. These ratios are referred to as credibility coefficients $C_{lo}(A,B)$, $C_{gr}(A,B)$, and $C_{eq}(A,B)$, respectively, where $C_{lo}(A,B) + C_{gr}(A,B) + C_{eq}(A,B) = 1$. Coefficient $C_{lo}(A,B)$ contains area d_1 and d_4 which provide information about the credibility that A is lower than B. Thus, it shows how much A is lower than B. On the other hand, coefficient $C_{gr}(A,B)$ contains area d_2 and d_3 which provide information about the credibility that A is greater than B. Thus, it shows how much A is greater than B. Coefficient $C_{eq}(A,B)$ contains intersection area d_1 and d_4 which provide information about how much two fuzzy variable overlap.



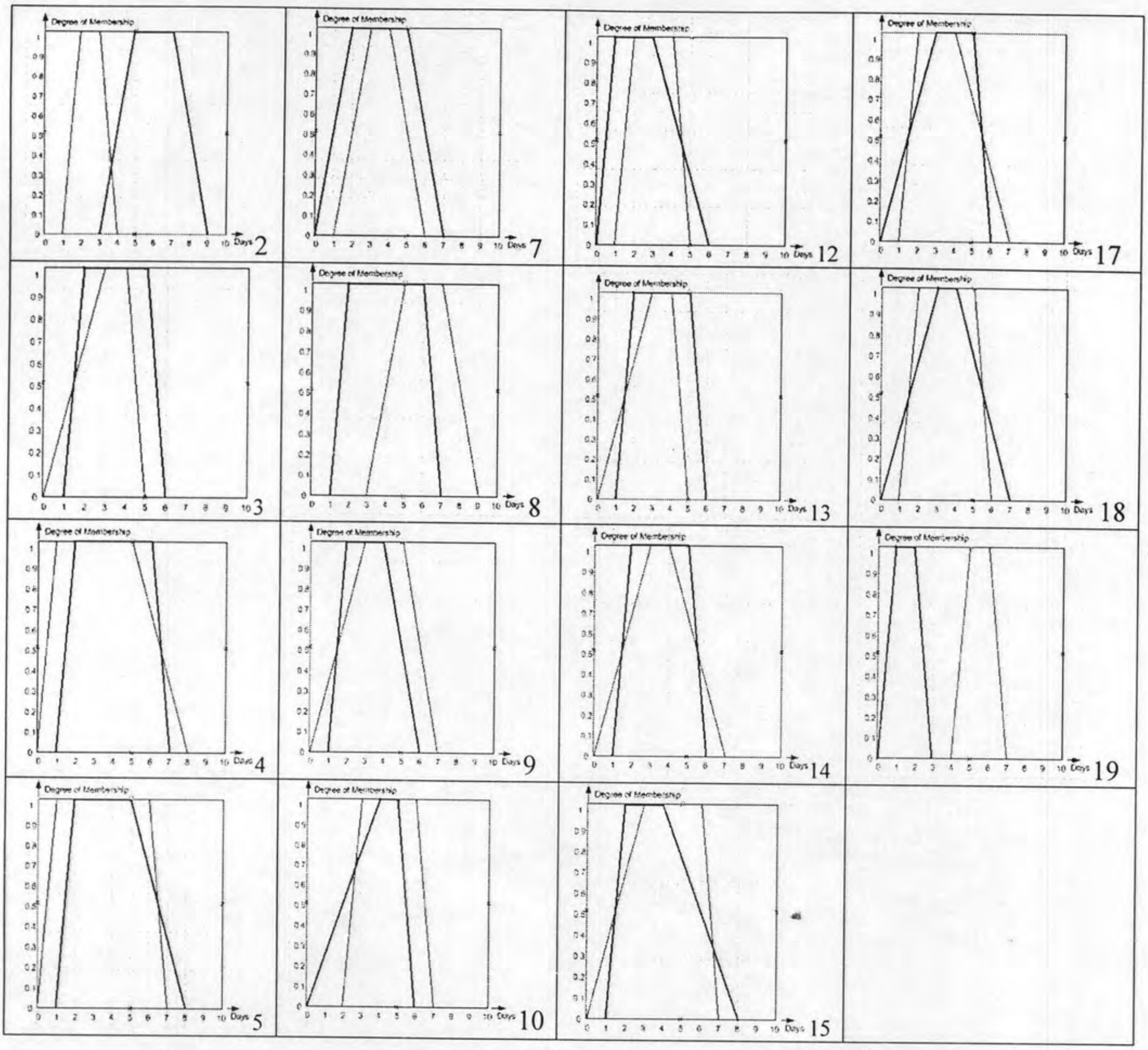
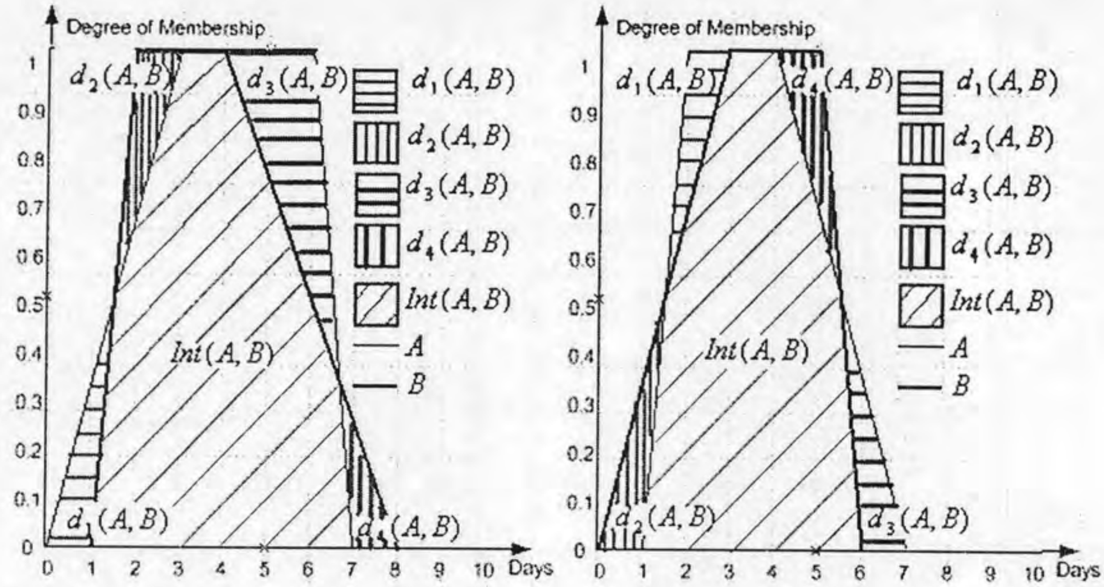


Figure 3.7 Comparisons between fuzzy set A and B



(a) (b)

Figure 3.8 (a) The Hamming distance of the two fuzzy variables A and B is numerically equal to the sum of the four areas d_1, d_2, d_3, and d_4, (b) Area d_1, d_2, d_3, and d_4 for the fuzzy variable A and B. The variables are the same as those of (a), but their names have been interchanged

This study compares two fuzzy variables based on these three credibility coefficients by following these decision rules:

(1) If $C_{gr}(A,B) > C_{lo}(A,B)$, then the final decision is that $A > B$. This decision is taken with a credibility factor equal to $C_{gr}(A,B)$.

(2) If $C_{lo}(A,B) > C_{gr}(A,B)$, then the final decision is that $A < B$. This decision is taken with a credibility factor equal to $C_{lo}(A,B)$.

(3) If $C_{gr}(A,B) = C_{lo}(A,B)$, then the final decision is that $A = B$. This decision is taken with a credibility factor equal to $C_{eq}(A,B)$.

The use of these three credibility coefficients increases the flexibility in providing a comparison between any shape (e.g., triangular, trapezoidal) of fuzzy sets which offers such a choice of differently representing the random–fuzzy variables of activity durations. Moreover, consistency of results can be obtained from the comparisons among more than two fuzzy numbers which are usually found in the network scheduling when multiple fuzzy activity durations are simultaneously ordered.

3.9 Determination of Significant Risk Factors

To identify significant risk factors having impact on activity duration, the extended activity duration resulting from a particular risk factor or a set of risk factors is computed by using a function of $l_{i(j)}$, $E_{i(j)}$, and initially estimated activity duration, where $l_{i(j)} = Wt \times P_{i(j)} \times$ quantitative likelihood and $E_{i(j)} = C_{i(j)} \times \%$ extended duration. In the proposed methods, the risk assessment is performed by experts. The average value vector represents the most likely values for the assessed risk factors is considered as the reference value of the considered risk factor. The result is then used to depict the most likely values of the risk factors of the available dataset.

Normally, risk factors are dependent. Effect of a particular risk factor may influence other risk factors and increase their impact on activity duration. The proposed method attempts to determine dependencies between risk factors. The more sophisticated method is developed to examine the extent to which the values of risk variables of a particular risk factor and a set of risk factors and a corresponding temporal variable are assigned. This method is modified from the sensitivity analysis by using simulation to provide data related to values of risk variables of a particular risk factor or a set of risk factors and the relevant activity duration affected by such risk factor(s). As a result, the correlations and relationships between the risk factors are considered. The differences between the actual durations and the estimated durations give an indication of the significance of a risk factor or a set of risk factors. The relative significance of the risk factors on the activity duration is prioritized by ordering the calculated differences. The risk factor or the set of risk factors causing the big difference is considered as the significant one, while the risk factor or the set of risk factors causing the small difference can be used to enhance the improvement of the project scheduling as it can provide an accurate estimate (the result is close the actual duration) and reflect reality of a project in consideration.

This research attempts to provide the more meaningful and communicative risk analysis results. The assessed significance of a risk factor or set of risk factors is represented by using linguistic term. The linguistic terms are assigned by applying the boundaries set for each proposed method, for example: very high is more than about 5%, high is 3% to 5%, medium is 2% to 3%, low is 1% to 2%, and very low is less than 1% which are used together with the application of RAIRFNET using the Salicone' s method to develop membership functions for a particular type of probability distributions. For example, the percent deviation of random–fuzzy duration from actual duration is 10% for triangular distribution for a set of risk factors. The calculated percent difference is interpreted as a considered set of risk factors having impact on an activity j is rated "very high".

In summary, this study presents risk variables and temporal variable by random–fuzzy variables. The mathematics for random–fuzzy variables is used in the network calculation to compute the fuzzy project completion time. The credibility coefficients C_{gr}, C_{lo}, and C_{eq} used in the decision – making rules with random–fuzzy variables are applied to order fuzzy durations of project activities. While Figure 3.9 presents components involved in the proposed approach and their relationships, Figure 3.10 shows the fuzzy network scheduling algorithm which is presented in the flowchart.

3.10 Conceptual and Theoretical Distinction

The integration between the risk assessment and the random–fuzzy network scheduling method (RAIRFNET) is explained in previous sections. This section is to explain distinction among RAIRFNET, the fuzzy network scheduling method (FNET), the probability – based method (Monte Carlo simulation). The advantages and disadvantages of each method are presented with several aspects. The comparison criteria are composed of (1) conceptual and theoretical distinctions, (2) assumptions of each method, and (3) obtained scheduling information from the completion of the analysis.

3.10.1 Types of Uncertainty

This section is to present distinctions among the proposed method (RAIRFNET), fuzzy – based method (FNET), and probability – based method used to address uncertainty in the network scheduling. Firstly, types of uncertainty are described. Uncertainties can be classified into two main categories based on information deficiencies. The first type of uncertainty is referred to as the variability uncertainty. It is a result of inherent fluctuations of the activity duration. The major sources of the variability uncertainty are the dynamic nature of construction projects, tight restriction on construction duration, multi – party and multi – task of construction projects.

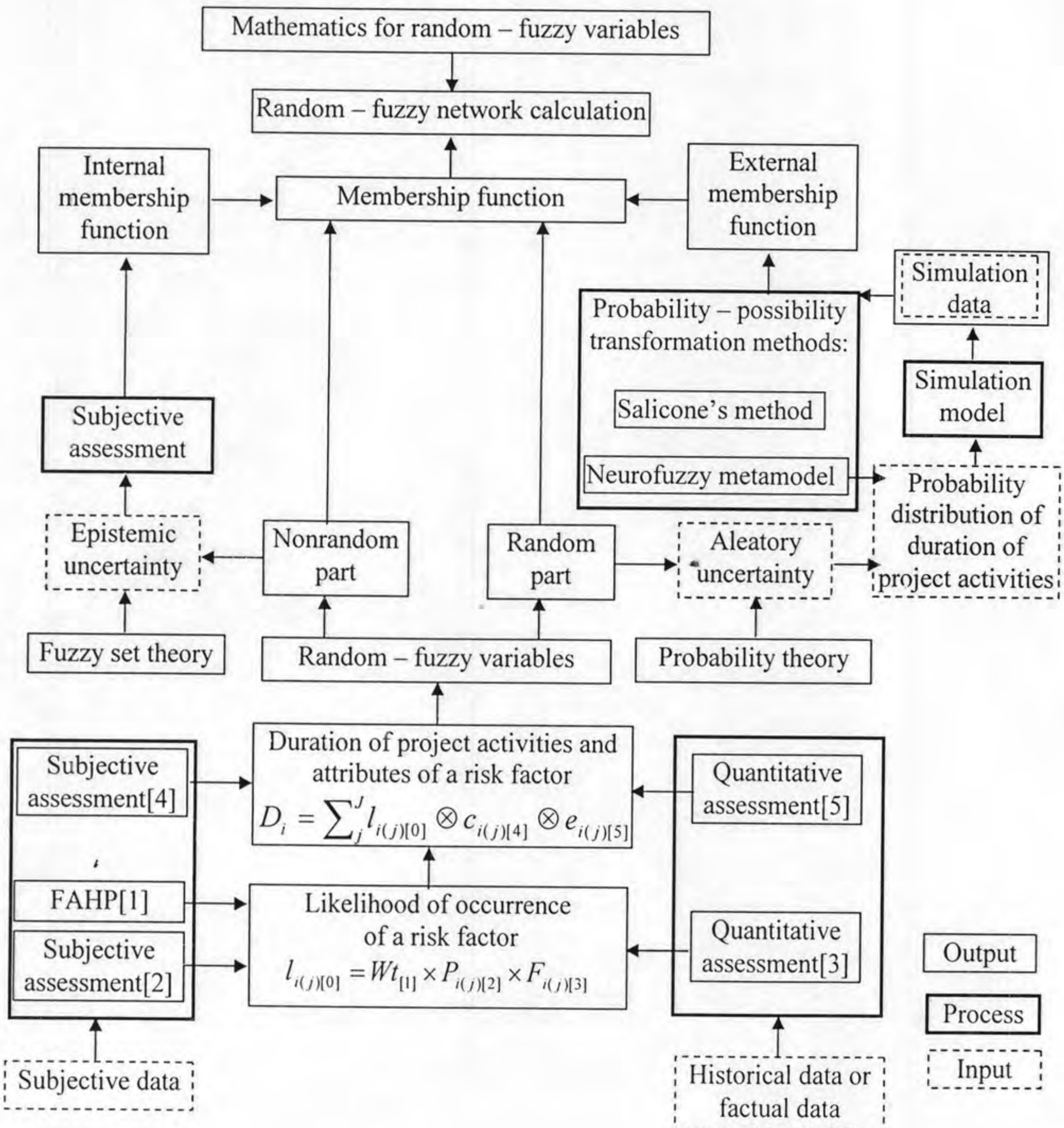


Figure 3.9 Components involved in the proposed approach and their relationships

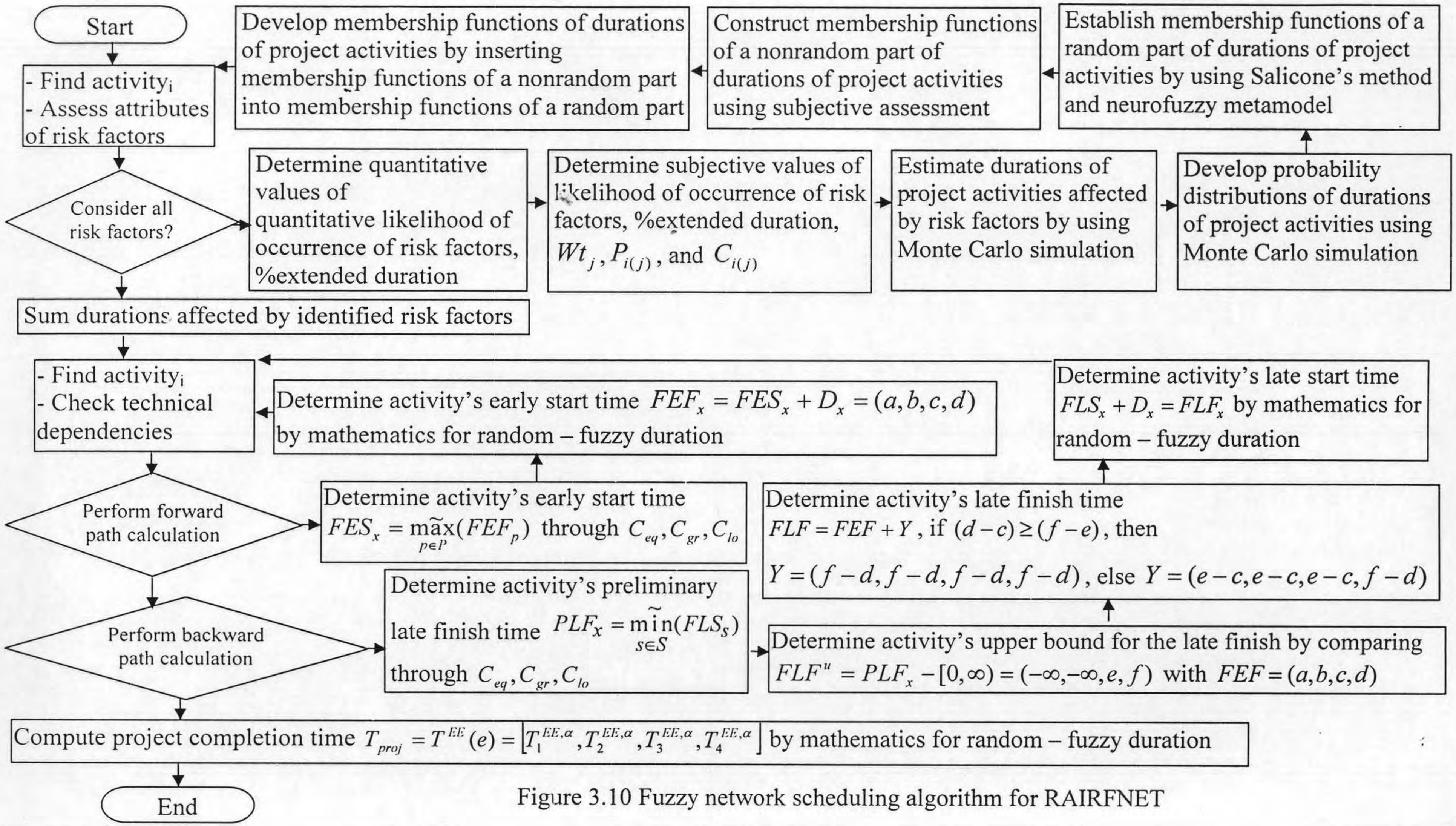


Figure 3.10 Fuzzy network scheduling algorithm for RAIRFNET

The second type of uncertainty which is called the knowledge uncertainty is brought about by incomplete knowledge. The main sources for uncertainty due to knowledge are model uncertainty, parameter uncertainty, and decision uncertainty. The model uncertainty stems from surrogate variables, excluded variables, abnormal situations, approximation uncertainty, incorrect form, and disagreement among experts. The parameter uncertainty can be classified into five types including 1) random error in direct measurement, 2) systematic error, 3) sampling error, 4) unpredictability, and 5) linguistic imprecision. The decision uncertainty results from risk measurement, social cost of risk, and quantification of social cost.

The quantification of uncertainty is a major task which is not much accomplished thorough existing researches. To be able to deal with every uncertainty, four issues should be concerned which include (1) an appropriate mathematical formalization of the conceived type of uncertainty (e.g., function of probability measure), (2) a calculus by which this type of uncertainty can be properly manipulated, (3) a meaningful way of measuring the amount of relevant uncertainty in any situation that is formalizable in the theory, and (4) a methodology developed for dealing with the various problems in which the uncertainty theory is involved. The quantification of uncertainty should first go through the selection of the suitable theories for addressing particular sources of uncertainty.

The next section presents the suitability of the probability theory, fuzzy theory, and Theory of Evidence, which encompasses the probability and possibility theories as particular cases, to represent uncertainty in construction scheduling in terms of random variables, fuzzy variables, and random–fuzzy variables, respectively.

3.10.2 Suitability of the Theory

Uncertainty in estimating activity duration which is only affected by random effects can be mathematically represented in the frame of the probability theory, by a probability density function. The establishment of a probability density function to estimate the uncertainty depends on an experimental method. The duration forecasting procedure is repeated several times, theoretically under identical circumstances and without mutual dependence. To capture characteristics of a random variable (i.e., duration of project activities in the case of scheduling), this procedure takes time and cannot be always followed for practical and economical reasons. There are two ways for establishing the probability density function, called type A evaluation and type B evaluation of uncertainty. The former is obtained by means of a statistical analysis of a series of observations, while the latter is based on judgement using all relevant information on the possible variability of measured quantity. Theoretically, the probability density function must be particularly constructed for all activities in the network. If the simulation is required, the determination of correlation coefficients among project activities is necessary.

In practice, however, the durations of project activities are not systematically recorded for statistical analysis, which brings about incomplete knowledge. Thus, the activity durations might not be all random according to the requirement of the probability theory. Although duration of project activities is affected by uncertainty due to random, systematic, and unknown contributions, theoretically it compensates and eliminates the systematic and unknown contributions. As the systematic and unknown effects do not propagate in a statistical way, uncertainties resulting from the systematic and unknown contributions can be classified as nonstatistical, which is considered as fuzzy variables rather than random ones and the probabilistic approach should not be applied. In order to estimate durations of project activities accounting for all kinds of contributions to the uncertainties inherent in duration of project activities including the random contribution (if the probability theory is considered as the case of PERT and Monte Carlo simulation) and the systematic and unknown contributions (if the fuzzy theory is considered as the case of FNET), random–fuzzy variables, which are defined within the Theory of Evidence (encompassing both the probability theory and the possibility theory as the case of RAIRFNET), can be employed to represent the scheduling variables in a unique mathematical object. Random–fuzzy variables can be defined as particular cases of fuzzy variables of type 2, while fuzzy variables are defined as particular cases of fuzzy variables of type 1. Thus, the use of random – fuzzy variables is practical and effective to follow the two approaches simultaneously. It is even more appropriate especially for construction projects that are unique as they employ new construction techniques and/or new resources (i.e., materials equipment) where the relevant historical data are usually insufficient and inapplicable.

3.10.3 Assumptions of Scheduling Methods

To enable the comparison among the probability based – methods (i.e., PERT and Monte Carlo simulation), fuzzy based – method (i.e., FNET), and the proposed RAIRFNET, the previous section presents the conceptual and theoretical distinctions of these methods and suitability of the theories on which each method is based, this section is to describe assumptions in order for each method to perform the network calculation. These methods are in association with three basic assumptions (Elmaghraby, 1997; Hendrickson and Au 1989, and Badiru 1993): (1) activities are independent, (2) the critical path is substantially longer than other paths, and (3) the critical path contains a sufficiently large number of activities. While PERT, FNET, and RAIRFNET are provided following the first assumption, Monte Carlo simulation is employed regarding the first assumption when the analysis accounts for the correlations among project activities. Practically, data related to dependencies among activities are unavailable which in turn lead to incomplete knowledge.

For the second assumptions, PERT provides optimistic project durations (the sum of the durations of the critical activities identified), when there are several near – critical paths contained in the project networks (Hendrickson and Au 1989, and Badiru 1993). However, the project completion time could be extended, if the noncritical paths merging into the critical path delay the critical path. While the expected early start time of a joint node, in the forward pass calculations of CPM and PERT, is determined from the largest early finish of all the activities leading to that node, the forward pass calculations of FNET use the $\widetilde{\max}$ operation to perform paired – wise comparisons for each and every element in the quadruples involved (i.e., the FEF of the preceding activities). The maximum value of each pair is selected to represent their respective elements for the new quadruples (i.e., the FES of the joint activity). Thus, the drawback associated with PERT can be overcome by using the FNET. The possibility measure (PM) and agreement index (AI) are used to calculate the criticality of project activities. However, PM does not provide an insightful assessment of the compatibility between the two fuzzy events. Besides, the AI is also inapplicable to the assessment of the single – valued type event and it also lacks an important characteristic called "the valuation". Even though the valuation can be given by PM, these two parameters cannot provide enough evidence to establish an ordering of the two fuzzy variables.

Unlike FNET, RAIRFNET employs the $\widetilde{\max}$ operation and credibility coefficients $C_{lo}(A,B)$, $C_{gr}(A,B)$, and $C_{eq}(A,B)$ to determine the earliest possible start time of the joint activities. A pitfall associated with FNET is undertaken. This procedure is not only able to represent incomplete knowledge about the ordering, but it also maps from an empirical to a symbolic relational analysis by preserving the scalar relation of the empirical analysis into the symbolic one.

In PERT and Monte Carlo simulation, the third assumption enables the use of the central limit theorem which controls a particular characteristic of the calculated project completion time by making the normal distribution for the project duration disregarding the distribution assumed for duration of project activities (Badiru, 1993). The third assumption is therefore considered as an advantage of FNET method of which an output is produced in the form of simple linear approximation, a membership function. The shape of the distribution of duration of project activities accordingly imposes the shape of the distribution of duration of a project.

The proposed RAIRFNET method, on the other hand, provides a random–fuzzy output consisting of a random part on which the characteristic is imposed by the central limit theorem and a nonrandom part which is produced in the form of simple linear approximation or a membership function. In this way, the sum of uncorrelated possibility distributions which represent the external membership functions of the corresponding random–fuzzy variables must tend to the normal possibility distribution based on the theoretical bound given by the central limit theorem. For the sum of uncorrelated possibility distributions representing the internal membership functions, the possibility distributions are supposed to never compensate with each other. Thus, the average operators providing a membership function within the sum of internal membership functions and the sum of external membership functions are used to produce the random–fuzzy output. The random–fuzzy output therefore represents an approximation of the normal possibility distribution.

However, the average operators cannot be applied, if the probability distribution to which the result (i.e., activity duration in the case of scheduling) should tend is not known (it is also not known the possibility distribution to which the result should tend). This problem can be observed in the estimation of duration of project activities affected by risk factors as the product operation used in a function of attributes of risk factors is utilized to calculate duration of project activities and accordingly the distribution to which the duration should tend cannot be predetermined. To approximate the way random variables (i.e., attributes of risk factors and activity duration) combine together, Monte Carlo simulation is employed to numerically estimate the probability distribution function representing uncertainty involved in the random variables. As simulation is able to provide sufficient data for establishing the probability distribution function of random variables, it can be appropriately used to predetermine the shape of the possibility distributions later used to represent the external membership functions.

As linguistic interpretations are used to perform the subjective risk assessment in the more direct and natural way, the used of fuzzy numbers is more appropriate to represent incomplete knowledge causing systematic and unknown contributions to uncertainty rather than deterministic and random numbers.

3.11 Summary

This chapter is to describe components including in a risk assessment integrated within random–fuzzy network scheduling method (RAIRFNET). The purpose of this chapter is to introduce a random–fuzzy concept into the risk assessment and network scheduling. Definitions of risk variables are provided in order for an assessor to collect the required data. The proposed method employs four data acquiring methods consisting of 1) experts' questionnaire, 2) historical record, 3) simulation, and 4) direct observation. The key feature of the proposed method is to develop the distributions of duration of project activities from data attainable at any construction period. Methods including the Salicone's method and neurofuzzy metamodel are used to establish membership functions of random–fuzzy variables (i.e., risk variables and temporal variable) regarding their probability distribution functions. As weight or degree of significance of each risk factor is different, FAHP is used to measure the weights of risk factors. Then, the activity durations represented by random–fuzzy numbers are inputted into the RAIRFNET and a project completion time is computed by using network calculation based on mathematics for random–fuzzy variables.

Alternatively, this research also attempts to integrate the risk assessment method within the scheduling method by using simulation as a combiner and use the neurofuzzy metamodel trained on simulation data to estimate the project completion time based on the inputted values of risk variables without any calculation required by the network scheduling method.

The main feature of the proposed method is summarized:

1) It is a hybrid modeling approach which combines different techniques, namely neural networks, fuzzy systems, simulation, and data acquiring approaches
2) It represents a general frameworks for random–fuzzy network scheduling where different data acquiring approaches (e.g., experts' questionnaire, historical record, simulation, and direct observation), probability–possibility transformation methods (e.g., Salicone's method and neurofuzzy metamodel), together with resulting model interpretability improving methods (e.g., significant input selecting and expert judgement) could be integrated.
3) Information associated with construction process design, working conditions, and technique and resource constrains are included in the training data acquired through either direct observation or simulation
4) Historical data are adjusted by subjective data for properly representing behaviors of an existing construction operation
5) A model accuracy and readability can be simultaneously achieved, and
6) It is an adaptable approach as the resulting fuzzy model can be modified when new data sets, additional data, and new input-output variables are included in the fuzzy model

In this way, the proposed method provides a more appropriate tool to model every uncertainty involved in the construction processes and the network analysis. In addition to handling a random contribution, it is able overcome problems in modeling systematic and unknown contributions affecting the risk assessment and activity duration estimate. Secondly, the analysis of uncertainty of several types is carried out by applying appropriate mathematics. Thirdly, it combines simulation and the $\alpha-$cut concept so that it can reflect the degrees of uncertainty differently influencing the impacted variables (i.e. D, W, P). The comparisons among models using random–fuzzy variables and the mathematics for random–fuzzy variables, traditional heuristic models (i.e., FNET), and models based probability theories are provided. Like models based on the fuzzy theory, when the possibility distribution is known, the proposed method requires only one calculation to achieve the outputs. This is one of the advantages over the probabilistic based method (i.e., simulation) which requires not only multiple cycles of the simulation experiments, but also a specifically designed number of simulation runs in order to perform the meaningful analysis of construction process. In these cases, the application of the proposed method will be very useful for the systematic and rational probabilistic estimate and risk analysis of construction projects.