Computer-Aided Civil and Infrastructure Engineering **25** (2010) 238–252

Fuzzy Monte Carlo Simulation and Risk Assessment in Construction

N. Sadeghi

Department of Civil & Environmental Engineering, University of Alberta, 3-133 Markin/CNRL Natural Resources Engineering Facility, Edmonton, Alberta, Canada T6G 2W2

A. R. Fayek*

Department of Civil & Environmental Engineering, University of Alberta, 3-013 Markin/CNRL Natural Resources Engineering Facility, Edmonton, Alberta, Canada T6G 2W2

&

W. Pedrycz

Department of Electrical & Computer Engineering, University of Alberta, W2-032 Electrical and Computer Engineering Research Facility, Edmonton, Alberta, Canada T6G 2V4

Abstract: *Monte Carlo simulation has been used extensively for addressing probabilistic uncertainty in range estimating for construction projects. However, subjective and linguistically expressed information results in added non-probabilistic uncertainty in construction management. Fuzzy logic has been used successfully for representing such uncertainties in construction projects. In practice, an approach that can handle both random and fuzzy uncertainties in a risk assessment model is necessary. This article discusses the deficiencies of the available methods and proposes a Fuzzy Monte Carlo Simulation (FMCS) framework for risk analysis of construction projects. In this framework, we construct a fuzzy cumulative distribution function as a novel way to represent uncertainty. To verify the feasibility of the FMCS framework and demonstrate its main features, the authors have developed a special purpose simulation template for cost range estimating. This template is employed to estimate the cost of a highway overpass project.*

*To whom correspondence should be addressed. E-mail: *aminah.robinson@ualberta.ca*.

DOI: 10.1111/j.1467-8667.2009.00632.x

1 INTRODUCTION

Performing risk analysis using Monte Carlo simulation is very common in construction project management; traditionally, probability theory is used to model uncertainty regarding simulation model inputs. In practice, the probability of an event can be estimated according to the frequency of that event occurring in a number of experiments (Pedrycz, 1998). However, if the number of experiments is not large enough to be significant, and more experiments cannot be performed, it is not possible to accurately estimate the event's probability. In these circumstances, we can engage human experts who are usually good at supplying the required information. Some researchers try to convert experts' knowledge into probabilistic distributions. This estimated probability may be used directly in a risk analysis problem (Ahuja et al., 1994), or it may be combined with available data using Bayesian methods to estimate a parameter that considers both subjective judgment and historical data (Garthwaite et al., 2005). However, there are some criticisms on performing probabilistic analysis on subjective and linguistically expressed data

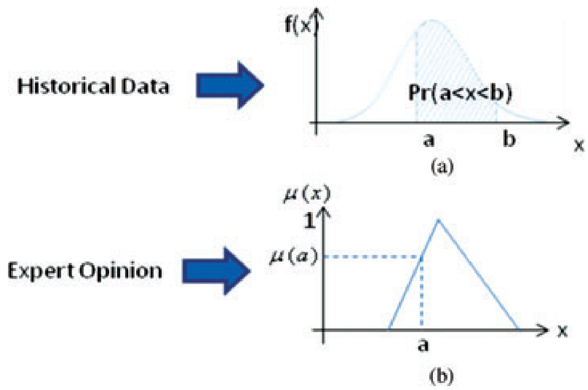


Fig. 1. (a) A probability density function (PDF) developed based on historical data. (b) A fuzzy set developed based on experts' judgment.

because subjective reasoning of individuals may not be appropriate for objective scientific conclusions (Goldstein, 2006). In other words, the information gained from experts is subjective and contains ambiguity, and there is a chance of introducing artificial knowledge that is not actually available to the model using probability values gained from experts (Guyonnet et al., 2003).

Fuzzy set theory (Zadeh, 1965) provides a methodology for handling subjective and linguistically expressed variables and representing uncertainty in the absence of complete and precise data. A fuzzy set A on the universal set X is defined by its membership function $\mu(x)$ and represents the degree that x belongs to the fuzzy set. In probability theory, a probability density function (PDF) is defined on continuous variables. The area under the curve of a PDF can be used to find the probability that the random variable falls into a particular interval (Figure 1a). In fuzzy set theory, instead of representing the probability value, the degree to which the objects are compatible with the properties of the fuzzy set is represented (Pedrycz and Gomide, 2007; Figure 1b). Therefore, although probability theory can handle the reach information gained from historical data, fuzzy set theory can represent the imprecise information of experts' judgments.

Fuzzy logic methods have been used successfully in various types of construction applications. For example, fuzzy logic approaches have been implemented for project scheduling (Lorterapong and Moselhi, 1996), construction bidding (Chao, 2007), modeling risk allocation in privately financed infrastructure projects (Jin and Doloi, 2009), predicting industrial construction labor productivity (Fayek and Oduba, 2005), saving energy in domestic heating systems (Villar et al., 2009), cost and life-cycle cost optimization of steel structures (Sarma and Adeli, 2000a, 2000b, 2002), freeway work zone capacity estimation (Adeli and Jiang, 2003), traffic flow forecasting (Stathopoulos et al., 2008), structural health monitoring (Carden and Brownjohn, 2008), and nonlinear control of high-rise building structures (Nomura et al., 2007; Jiang and Adeli, 2008), just to name a few representative examples.

As fuzzy and probabilistic methods are appropriate for representing the uncertainty of different parameters, we usually face problems in which some of the variables are fuzzy and some are random (Guyonnet et al., 2003). These situations are common in the simulation of construction projects, where each project is unique and only limited data are available for many factors affecting a project. Information gained from experienced personnel is an excellent source of data in these projects. Furthermore, some factors, such as worker's skill and complexity of the work are subjective. In current construction simulation frameworks, a PDF should be provided for all uncertain variables. Having a simulation framework that can handle both fuzzy and probabilistic uncertainty is very essential in the risk analysis of construction projects.

This article proposes a Fuzzy Monte Carlo Simulation (FMCS) framework that, for the first time, provides the capability of considering fuzzy and probabilistic uncertainty simultaneously for the risk analysis of construction projects. The output of FMCS has been modeled using fuzzy random variables and represented using a cumulative distribution function (CDF). Fuzzy CDF is a novel approach for risk analysis that is capable of considering both types of uncertainties in a single representation. We develop a cost range estimating simulation template based on the FMCS method. Finally, an example involving cost range estimating of a highway overpass project is provided to illustrate the feasibility of the proposed FMCS framework and fuzzy CDF method.

2 LITERATURE REVIEW

The Monte Carlo simulation method is used for estimating the output Y of a function (M) with random input variables ($R_1, R_2, \ldots, R_n$) (Figure 2). We run various experiments for inputs by sampling from the input probability distributions and collecting the model outputs (Ahuja et al., 1994). In general, the generated random samples in Monte Carlo simulation are statistically independent for each input variable; therefore,

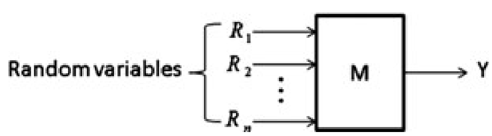


Fig. 2. The output Y of a function M with random inputs can be calculated using Monte Carlo simulation.

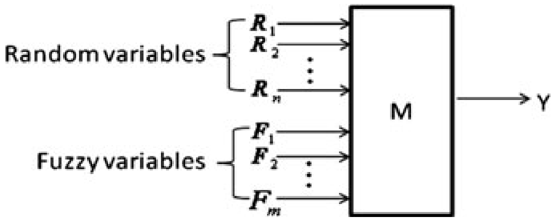


Fig. 3. A function M with both fuzzy and random inputs.

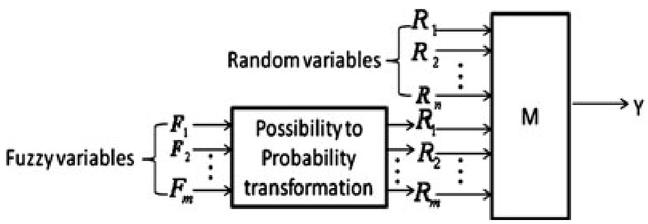


Fig. 4. Converting fuzzy sets to PDF before performing Monte Carlo simulation.

the independence of input parameters should always be investigated.

In practice, we may face a generalized problem in which we have both types of uncertainty, fuzzy and probabilistic. Here, we need to determine the output Y of a function (M) that has $R_1, R_2, \ldots, R_n$ being random variables and represented by probabilistic distributions and $F_1, F_2, \ldots, F_m$ being fuzzy sets (Figure 3).

To estimate the output of this generalized model, most researchers attempt to eliminate or transform one type of uncertainty to another before performing a simulation. For example, Wonneberger et al. (1995) performed a possibility to probability transformation for a problem with both types of uncertainty to change the problem to a purely probabilistic simulation (see Figure 4). However, fuzzy logic and probability theory are orthogonal (and complementary) concepts and capture different types of uncertainty. Also, there is no fully accepted way of transforming one to another (Pedrycz and Gomide, 1998).

Guyonnet et al. (2003) proposed a "hybrid approach" for solving a model that has both fuzzy and random types of uncertainty without transforming one type to another. The essence of their approach is summarized in Figure 5 for a model M that has random variables as probabilistic distributions $R_1, R_2, \ldots, R_n$ and fuzzy sets $F_1, F_2, \ldots, F_m$ for the inputs. To determine the output Y of this model, as indicated in Figure 5, a number of sample sets (w) are generated from the probability distributions. After assigning each sample set $r_1, r_2, \ldots, r_{ni}$ to the random variables of the model, the α-cuts of the fuzzy inputs are calculated for different levels of α. Let us recall that the α-cut of a fuzzy set F at the level of $\alpha \in (0, 1]$ is a set F_α, whose members have a membership degree equal or greater than α. Therefore, the α-cut of each fuzzy input represents a set of values. Guyonnet et al. (2003) calculated the Infimum (Inf) and Supremum (Sup) values of the model M considering all the values that are located within the α-cuts of the input fuzzy sets. In this way, for each sample set (i) and each α-level (α_j) two output values are calculated: $Y_{i\alpha j,inf}$ and $Y_{i\alpha j,Sup}$ (Figure 5). Guyonnet et al. (2003) suggested that minimization and maximization algorithm can be used for finding Inf and Sup values of a general model. However, in their application, the model was a simple monotonic function, and the Inf and Sup values were identified directly without using minimization or maximization algorithms.

For decision making based on the hybrid approach, Guyonnet et al. (2003) developed the histograms of the Inf and Sup values of the α-cuts at each α-level and calculated the final Inf and Sup of the output α-cut based

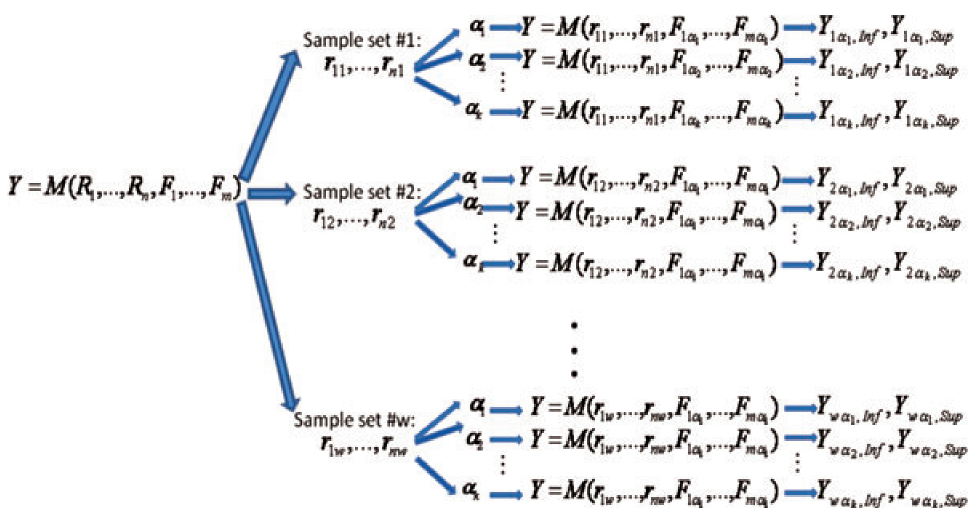


Fig. 5. Guyonnet et al.'s (2003) "hybrid approach" for fuzzy Monte Carlo simulation.

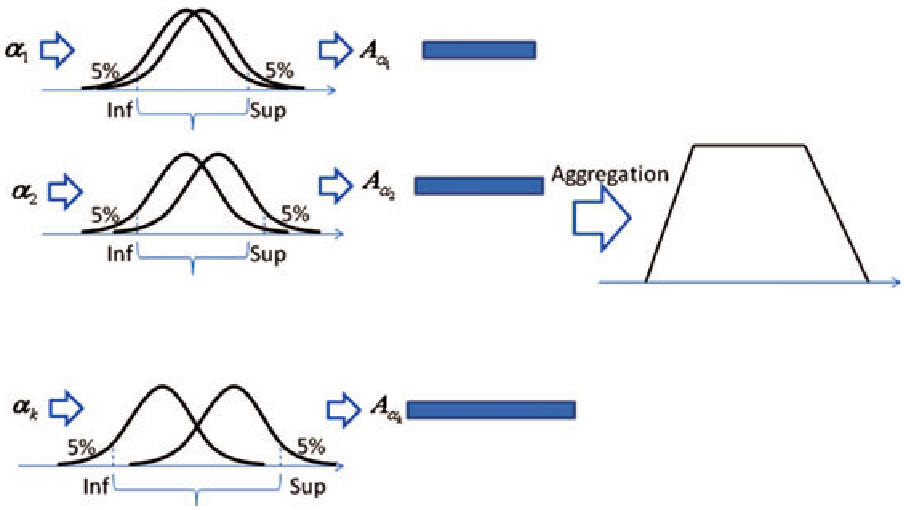


Fig. 6. Calculating the output α-cut based on the histogram of the Inf and Sup values of the α-cuts at the level of α.

on a 5% probability of getting values lower than Inf and higher than Sup (Figure 6). The α-cuts are then aggregated to produce the output as a fuzzy set. This aggregation is done based on the representation theorem, which states that each fuzzy set can be reconstructed from its α-cuts according to Equation (1).

$$A(x) = \text{Sup}_{\alpha\in[0,1]}[\alpha A_\alpha(x)] \quad (1)$$

In their application, Guyonnet et al. (2003) were interested in finding the possibility that the output of the model is less than a specific threshold. For the fuzzy set F with membership function μ_F and a threshold T, the possibility P is calculated based on Equation (2).

$$P(F > T) = \underset{u>T}{\text{Sup}}\,\mu_F(u) \quad (2)$$

The approach of Guyonnet et al. (2003) is unique in considering fuzzy and probabilistic inputs simultaneously in Monte Carlo simulation; however, it is not free from shortcomings. A careful analysis reveals some aspects of the approach that require further evaluation/refinement.

1. The α-cuts of a fuzzy set cannot always be represented by Inf and Sup values (Figure 7a). Therefore, as indicated in Figure 7b, values that do not actually belong to an α-cut may be considered in the α-cut when representing a fuzzy set only with Inf and Sup values of the α-cuts. Therefore, the specificity of the results may be decreased in this approach.
2. Guyonnet et al. (2003) do not mention why a 5% probability of getting lower and higher values of the histograms of the α-cuts will generate the Inf and Sup of the output α-cut. In this manner, they remove the random type of uncertainty and consider a fuzzy set for the output. Baudrit et al. (2005) indicate this method leads to unrealistic output and overestimation.
3. In addition, if only random inputs are considered as the extreme case for this model, the result will not be similar to the traditional Monte Carlo simulation approach. In this case, the absence of fuzziness results in equal histograms for the Inf and Sup values at all levels of α, because the difference between these histograms result from the fuzzy inputs. Therefore, the method will produce the same α-cuts for all values of α, and the result of their aggregation will be an interval that does not contain enough probabilistic or fuzzy information to help in decision making (Figure 8).

Baudrit et al. (2005) propose an approach for "post-processing" of the hybrid method of Guyonnet et al. (2003) using the theory of evidence (or theory of belief functions; see Shafer, 1976). The final output of their proposed method does not directly represent the fuzziness or randomness, but rather analyzes the output with concepts that are defined in the theory of evidence.

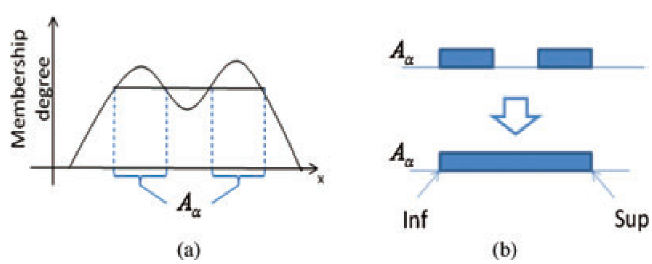


Fig. 7. (a) An α-cut of a non-convex fuzzy set. (b) Considering only Inf and Sup values of the α-cuts and the associated lack of specificity.

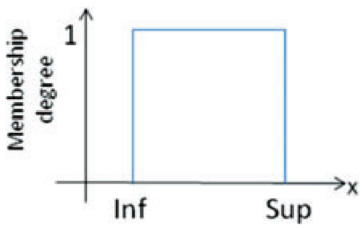


Fig. 8. Output of Guyonnet et al.'s (2003) approach in the absence of fuzziness.

As the sources of uncertainty for different parameters of a project differ, a framework that can handle fuzzy and probabilistic uncertainty simultaneously and represent each type of uncertainty separately is helpful. However, as discussed, existing frameworks do not represent both fuzzy and probabilistic types of uncertainty in the output results. In this article, we propose a new method for handling both fuzzy and probabilistic uncertainty in the inputs of a simulation model.

3 FUZZY MONTE CARLO SIMULATION FRAMEWORK

A FMCS framework is developed for risk analysis of problems that contains both fuzzy and random inputs. Consider a model (M) that has both random variables as probabilistic distributions $R_1, R_2, \ldots, R_n$ and subjective variables as fuzzy sets $F_1, F_2, \ldots, F_m$ for the inputs. In the FMCS framework, sample sets are produced from the probabilistic distributions. After assigning each sample set $r_1, r_2, \ldots, r_{ni}$ to the random variables of the model, the model will contain only fuzzy input variables. We perform fuzzy arithmetic to calculate the output in the form of a fuzzy set (Figure 9).

Based on extension principle (Zadeh, 1975), one can apply fuzzy arithmetic on a function (M) with input fuzzy sets $F_1, F_2, \ldots, F_n$ that are defined on the universes $X_1, X_2, \ldots, X_n$. Assuming that $\mu_1(x_1), \ldots, \mu_k(x_k)$ are membership functions defined on $F_1, F_2, \ldots, F_n$ the membership function of Y can be calculated as in Equation (3) according to Zadeh's extension principle. Similar to Monte Carlo simulation, the independence of input parameters is assumed in fuzzy arithmetic; the output may be overestimated when using fuzzy arithmetic for a function with dependant input parameters (Hanss, 2004).

$$y = \sup_{M(x_{1,\ldots,}x_k)=y} \{\min[\mu_1(x_1), \ldots, \mu_k(x_k)]\} \tag{3}$$

The α-cut method can also be used to perform fuzzy arithmetic on a function. This method is equivalent to the Zadeh's extension principle. However, it is easier to implement because it is based on interval analysis of the α-cuts of the input fuzzy sets. Given a function $Y = M(A_1, A_2, \ldots, A_n)$, Equation (4) shows how the α-cut of Y may be calculated at the level of α using interval analysis (Chang and Hung, 2006). For calculating Equation (4) using a computer program, optimization routines should be carried out for finding the Inf and Sup values of the output α-cut intervals. However, if the model is monotonic (increasing or decreasing) with regard to the input fuzzy sets, we can calculate the output α-cut based on the Inf and Sup values of the input α-cut intervals (Abebe et al., 2000). The output α-cuts are usually calculated for a finite number of α-levels. The α-cuts of Y at different levels of α can be aggregated to produce the fuzzy set for Y according to Equation (1).

$$Y_\alpha = M(A_{1,\alpha}, A_{2,\alpha,\ldots,} A_{n,\alpha}) \tag{4}$$

As explained previously, in FMCS, we perform fuzzy arithmetic for each sample set. Therefore, each output in FMCS framework is in the form of a fuzzy set. The final output is represented as fuzzy sets with random variation that can be modeled with a fuzzy random variable. A fuzzy random variable is a mapping from the probability space to the fuzzy sets (Terán, 2007).

The ultimate goal of any risk analysis model is decision-making support. FMCS is proposed as a general form of Monte Carlo simulation and similar decisions that are made using Monte Carlo simulation can be made based on the FMCS framework. The mean and variance can be calculated to provide an estimate of the output of Monte Carlo simulation. We can benefit from work in measurement theory for calculating the mean

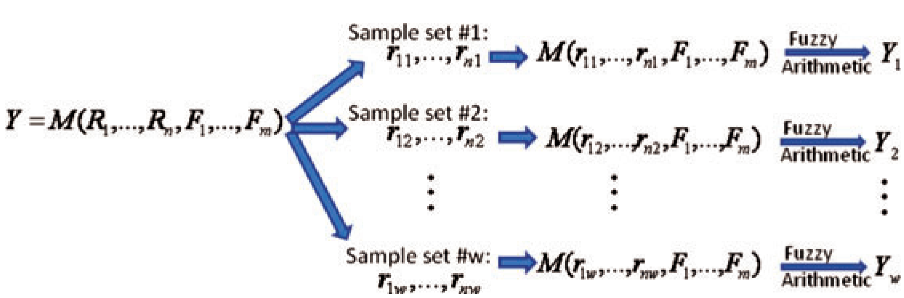


Fig. 9. Fuzzy Monte Carlo Simulation (FMCS) approach.

and variance of the output of the FMCS framework. Terán (2007) used fuzzy random variables to represent the results of measurements. He represented each measurement as a fuzzy set, whereas the variability between fuzzy sets was considered random. This scenario is similar to results obtained from FMCS. Terán (2007) suggested the use of fuzzy arithmetic to perform statistical calculations on the fuzzy samples.

Similarly, we can apply fuzzy arithmetic to the fuzzy outputs of FMCS to find the mean and variance. This approach provides an analysis tool that behaves reasonably well. When we have no fuzziness, the results are the same as those obtained with a classical statistical analysis approach. Also, when we have no randomness, the results are compatible with the fuzzy set theory analysis approach. To compare, the approach proposed by Guyonnet et al. (2003) does not allow for the calculation of the mean or variance.

However, the mean and variance is not enough for the risk analysis of construction projects. In construction management, a decision maker is usually interested in two other important statistics: (1) an arbitrary quantile and (2) the probability of exceeding (or not exceeding) a specific threshold. For example, one may want to estimate the completion time of a project with 95% confidence. This value is referred to as the 95th quantile of the output. In the context of the simulation process, this means that 95% of the conducted simulation results are less than the completion time. Decision makers are also interested in finding the probability that a project will exceed a certain value of cost or time (Ahuja et al., 1994).

The CDF is typically used for finding the probability of not exceeding a given threshold. Equation (5) defines the CDF function of a random variable X (Ahuja et al., 1994). The inverse of the CDF is used for finding the arbitrary quantile.

$$F_x(x) = \Pr\{X < x\} \tag{5}$$

Considering a finite number of random samples resulting from a Monte Carlo simulation, the CDF function can be estimated from Equation (6).

$$F_x(t) = \frac{\text{Number of samples that are less than } t}{\text{Total number of samples}} \tag{6}$$

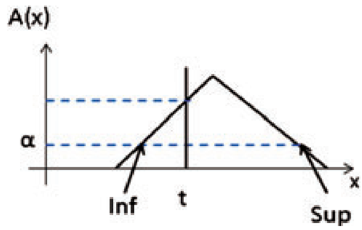


Fig. 10. The threshold is in the sample fuzzy set.

However, the samples of the output of a FMCS are fuzzy sets, and when the threshold is a member of a sample fuzzy set, there is an uncertainty in considering the sample fuzzy set as less than or greater than the threshold t (Figure 10).

We solve this problem by incorporating fuzziness into the CDF and generating a fuzzy CDF. We consider α-cuts of the samples and calculate two CDFs at each α-level: $F_{\alpha,\max}$ is calculated based on the Inf values of the α-cuts of the samples at the level of α (Equation (7)), and $F_{\alpha,\min}$ is calculated by considering the Sup values of the α-cuts of the samples at the level of α (Equation (8)). When we compare the Inf values of the α-cuts with the threshold, more α-cuts are less than the threshold t, which results in the maximum CDF, $F_{\alpha,\max}$. Similarly, Sup values of the α-cuts will generate the minimum CDF function, $F_{\alpha,\min}$. For example, in Figure 10, the fuzzy set A is considered less than t for calculating $F_{\alpha,\max}(t)$, and is considered greater than t for calculating $F_{\alpha,\min}(t)$.

$$F_{\alpha,\max}(t) = \frac{\text{Number of Inf of } \alpha\text{-cuts of the samples that are less than } t}{\text{Total number of samples}} \tag{7}$$

$$F_{\alpha,\min}(t) = \frac{\text{Number of Sup of } \alpha\text{-cuts of the samples that are less than } t}{\text{Total number of samples}} \tag{8}$$

$F_{\alpha,\min}$ and $F_{\alpha,\max}$ will generate a CDF bound $F_\alpha(x)$ at each α-level. The final fuzzy CDF, $F(x)$, can be determined based on its corresponding α-cuts, $F_\alpha(x)$, according to the representation theorem (Equation (1)). The graphical representation of aggregating CDF bounds to produce the final fuzzy CDF is shown in Figure 11. The fuzzy CDF is a fuzzy function, meaning that if we provide a numeric input, the function produces a fuzzy set as output.

By using the fuzzy CDF method in the FMCS framework, we remove the second and third shortcomings associated with Guyonnet et al.'s (2003) hybrid approach. The output of FMCS captures both fuzzy and probabilistic uncertainty, and therefore we do not have the overestimation problem that exists in Guyonnet et al.'s (2003) method. Furthermore, when we have no

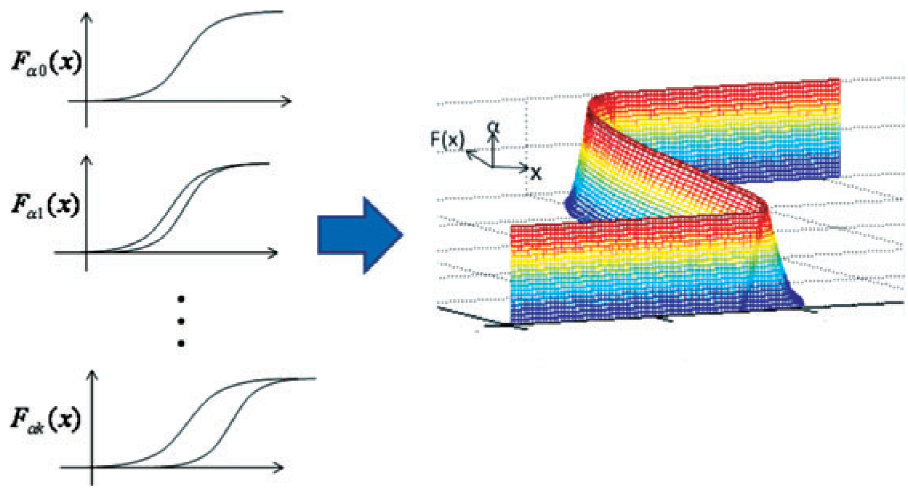


Fig. 11. Aggregating the CDF bounds at different levels of α to generate the fuzzy CDF.

fuzziness, each sample value is a real number instead of a fuzzy set. In this case, $F_{\alpha,\max}(x)$ equals to $F_{\alpha,\min}(x)$, and all of the α levels will have equal CDFs. Therefore, the results are exactly equal to those obtained by traditional Monte Carlo simulation. Finally, having developed the fuzzy CDF, we can perform any decision analysis that can be performed using CDF, as discussed in the section below. However, the fuzzy CDF analysis method still retains the first shortcoming of the hybrid approach, in that only the Inf and Sup values of the α-cuts are considered for decision making. The ultimate goal of any risk analysis model is decision making. The estimator can estimate the probability that the output is less than a threshold t. The answer is in the form of a fuzzy set that is obtained by intersecting the fuzzy CDF graph at the desired threshold. A method for making decisions using fuzzy sets is based on the confidence level. The estimator is able to decide on a confidence level between 0 and 1 to get a range of values for the final output. This range is calculated by finding the α-cut at the value of 1 minus the confidence level (Mauris et al., 2001). In this way, the estimator can choose from a range of values instead of a crisp output. An arbitrary quantile can also be estimated using the inverse of the fuzzy CDF. In Section 5, we explain decision making using fuzzy CDF using an example.

4 PRACTICAL ASPECTS OF USING FMCS IN CONSTRUCTION RISK ASSESSMENT

A simulation based approach for risk analysis of a problem in construction management can be summarized in the following steps: (1) identifying the structure of the problem, (2) quantifying uncertainty in different parameters of the model, (3) performing a simulation, (4) analyzing the results, and (5) making a consensus decision (Walls III and Smith, 1998). Traditionally, all the input uncertainties are modeled based on probability theory, and simulation is performed to find the output results. FMCS framework extends the practical use of Monte Carlo simulation by providing the capability of choosing between fuzzy sets and probability distributions for quantifying the input uncertainties of a Monte Carlo simulation model. However, different simulation methods may be applied on a model depending on the structure of the problem. For example, discrete event simulation can be used to analyze the sensitivity of dynamic schedule and resource constraints to unexpected construction scenarios, whereas Monte Carlo simulation is applied to a model that does not depend on the time. Future research can be conducted to provide the capability of considering both fuzzy and probabilistic uncertainty in other simulation approaches such as discrete event simulation.

One can find many practical examples in Monte Carlo simulation models of construction projects in which some of the input variables are estimated based on experts' judgment and some are derived from historical data. Life-cycle analysis of a pavement design by Walls III and Smith (1998) is a good practical example. They have used recent bid records to find probability distributions of the costs of construction and rehabilitation of a project, whereas experts' judgment has been used for estimating the service life of the pavement. Monte Carlo simulation is used as the method of risk analysis in this study. Although explaining the details of this project is beyond the scope of this article, we suggest this work can be used as an actual case study for future research.

Range estimating is another example in which FMCS can be applied. Range estimating using Monte Carlo simulation is a common process for risk analysis and decision making regarding the budget and schedule of construction projects. The approach is based on considering the work breakdown structure (WBS) of a project and estimating the cost or duration of each work package in the form of a PDF. The Monte Carlo simulation method is used to aggregate the work packages and to estimate the range and degree of uncertainty for the overall project cost or duration (Shaheen et al., 2007; Ahuja et al., 1994).

Accurate cost estimation plays a major role in the success of a construction project. Different methods have been suggested in the literature for estimating the cost of construction projects (e.g. Adeli and Wu, 1998) and for determining a contingency value (Touran, 2003). Contingency is the anticipated cost for unknowns that may increase the total cost of a project (Ahuja et al., 1994). Monte Carlo simulation is a common approach that is performed for estimating the cost and contingency. The process of Monte Carlo simulation for cost range estimating can be summarized as follows:

1. Provide the WBS and remove work packages that do not have major effects on the total cost of the project. Ahuja et al. (1994) suggest that those work packages that affect the total cost of the project with at least 0.5% should be considered major.
2. Provide the quantity and unit cost related to each work package. Use a PDF to represent the uncertainty associated with different values of the quantity and unit cost of each work package.
3. Use Monte Carlo simulation to determine the uncertainty associated with the total cost of the project.

Although expert judgment is usually used in range estimating of construction projects (Ahuja et al., 1994), only the random type of uncertainty can be considered using this approach. Experts' judgment is especially useful in the preliminary stages of a project when not enough data are available for many factors. For example, before performing geotechnical tests, experts may estimate the geotechnical parameters to calculate the cost of the project. In the later stages of the project, we may still have some fuzzy parameters due to the unique aspects of the project, lack of data, or subjectivity. Shaheen et al. (2007) suggested an alternative method of using fuzzy set theory for modeling uncertainties in range estimating problems. The researchers proposed a range-estimating model that uses fuzzy arithmetic to estimate the cost or duration of a project with purely fuzzy inputs. However, as the source of information about various parameters of a project differs, we may have probabilistic uncertainty for some of the input variables and fuzzy uncertainty for others. Therefore, we need a range-estimating model that is capable of handling both types of fuzzy and probabilistic inputs. The FMCS framework can be used to solve this problem by using the PDF to represent random uncertainty and fuzzy sets for representing subjective or linguistically expressed values in the WBS. We have developed a cost range estimating template based on FMCS framework. This template illustrates how the FMCS framework can be implemented for practical use in construction management.

A special purpose simulation (SPS) template has been developed by connecting the Simphony.NET© 2005 platform (The University of Alberta, Edmonton, Alberta) and MATLAB (The MathWorks, Inc., Natick, Massachusetts) for range estimating based on the proposed FMCS. Simphony.NET is a simulation software application for construction processes that is capable of developing different SPS templates. The SPS template provides a tool for an expert, who is not necessarily knowledgeable in simulation, to develop a simulation model in the area of his/her expertise (Hajjar and AbouRizk, 1999). Our developed cost range estimating SPS template allows the user to represent the WBS of a project by dragging and dropping the elements on a computer screen and connecting them according to the structure of the WBS. Using the fuzzy Monte Carlo cost range estimating template, input values can be entered as the properties of each element in the form of fuzzy sets or probabilistic distributions.

We have used the α-cut method to perform fuzzy arithmetic on the fuzzy sets in FMCS. As the calculations for cost range estimating are limited to addition and multiplication, which are monotonically increasing, finding the Inf and Sup values of the output α-cut intervals is straightforward by using the Inf and Sup of the input α-cut intervals. The elements of the developed template are listed in Table 1. The root element is responsible for calculating the value of α and deciding whether a minimum or maximum value of the α-cut should be calculated in each run of the simulation. Other elements identify their appropriate actions based on the status of the root element in each run. The cost of each child work package is calculated by multiplying its unit cost and quantity. A number of child work packages or parent work packages may be defined under a parent work package. Therefore, it is possible to have any number of levels in the WBS. The cost for the parent work package is the sum of the costs of its lower level work packages multiplied by the quantity of the parent work package. The analysis element collects the output results and sends them to a MATLAB routine. In this routine fuzzy CDF graph is created for decision making.

Table 1
Elements of fuzzy Monte Carlo cost range estimating template

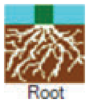
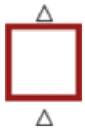
Element name	*Graphical representation*	*Description*
Root		This element defines the status of the simulation and controls the actions of all the other elements.
Parent Work Package		Parent Work Package represents a group of work packages that will be defined under this element.
Child Work Package		Child Work Package represents the lowest level of the WBS, the unit cost and quantity can be defined for this element.
Analysis Element		This element collects the outputs and calculates the statistics.

Table 2
Major work packages of WBS for a highway overpass project and their associated cost and quantity (adapted from Ahuja et al. 1994)

Work package	*Quantity*	*Unit cost*
1. Excavation (m^3)	Uniform (2,200, 2,500)	Triangular (10, 11, 13)
2. Backfill (m^3)	Uniform (1,700, 2,200)	Triangular (9, 10, 13)
3. Pilings and Bells		
Piling (300 dia) (m)	Constant (160)	Constant (29)
Piling (750 dia) (m)	Constant (510)	Triangular (175, 183, 190)
Bells (1,500 dia) (ea)	Constant (42)	Triangular (370, 390, 420)
Bells (1,200 dia) (ea)	Constant (16)	Constant (340)
4. Cast in place concrete		
Pier footing (m^3)	Constant (73)	Triangular (320, 330, 350)
Pier column (m^3)	Constant (55)	Triangular (600, 650, 700)
Abutments (m^3)	Constant (635)	Triangular (200, 235, 290)
Approach slabs (m^3)	Constant (55)	Triangular (220, 230, 400)
Bridge girder (m^3)	Constant (1,310)	Triangular (370, 390, 450)
Parapets incl. finish (m)	Constant (171)	Triangular (150, 160, 175)
Concrete median (m)	Constant (67)	Constant (124)
5. Concrete slope protection (m^3)	Uniform (1,000, 1,100)	Triangular (42, 45, 50)
6. Hot mix asphaltic concrete paving (m^2)	Constant (1,900)	Triangular (17, 18, 19)
7. Deck water proofing	Uniform (1,800, 2,000)	Constant (5.7)
8. Class 5 finish (NIC parapets) (m^2)	Constant (565)	Constant (6)

5 AN ILLUSTRATIVE EXAMPLE TO COMPARE MONTE CARLO SIMULATION AND FMCS

In this section, we analyze the behavior of FMCS framework in comparison with Monte Carlo simulation using a cost range estimating example. Consider a sample application by Ahuja et al. (1994) of a cost range estimating problem for a highway overpass project. The unit cost and quantity for major work packages of this project are shown in Table 2, and probabilistic distributions are used to express the uncertainty regarding those variables. These uncertainties may result from uncertainty regarding the accuracy of take-off values or different scenarios that may happen in the field

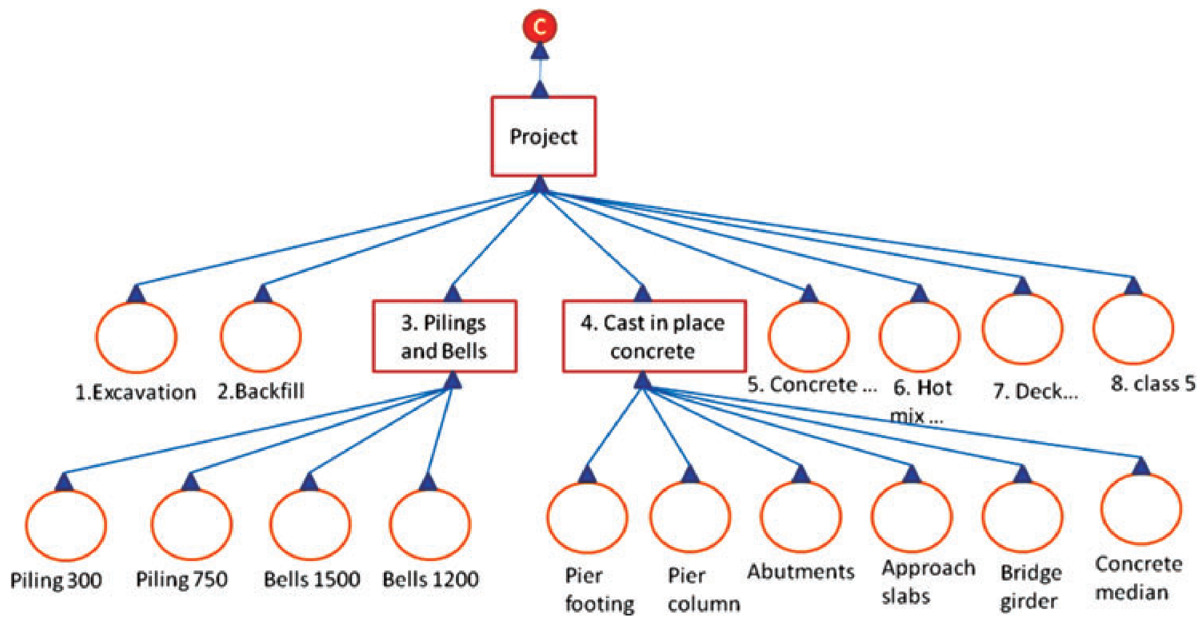


Fig. 12. Developed model for the highway overpass example using fuzzy Monte Carlo cost range estimating SPS template.

during construction. For example, uncertainty in the unit cost may be a result of uncertainty associated with the productivity of workers or variability in weather conditions. Ahuja et al. (1994) used subjective judgment to derive the given probabilities, and it is beyond the scope of this article to verify these distributions, which are also specific to this example and its assumptions. We assume that the model is developed correctly and their suggested probability values are appropriate. We should note that this assumption does not bring any limitations to our analysis, becuse the model is used with the sole goal of performing a sensitivity analysis on the FMCS framework and comparing the results with the probabilistic approach. Figure 12 illustrates the model developed for this example using the fuzzy Monte Carlo cost range estimating template.

To experiment with the FMCS approach using a combination of fuzzy and probabilistic inputs, some of the probability distributions of Table 2 are transformed into fuzzy sets using the probability-possibility transformation method of Dubois et al. (2004). For example, Figure 13 represents the transformation of the triangular distribution for the unit cost of the pier footing process in Table 2 to a fuzzy membership function based on Dubois et al. (2004). In this method, the confidence level of the intervals is estimated using the probability of that interval. This probability is equal to the area under the PDF that is bounded within that interval. Among different intervals of the same confidence level, Dubois et al. (2004) proved that the most informative interval is the one with minimal length, and this interval should be considered as the α-cut of the final fuzzy set. This approach will produce a nested family of intervals that are considered as the α-cuts of the final fuzzy set.

We are not recommending that such transformations from probabilistic to fuzzy sets be done in practice but

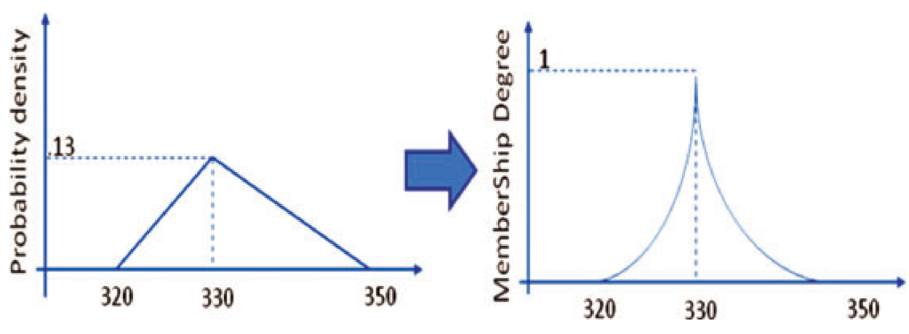


Fig. 13. Transformation of the triangular distribution for the unit cost of the pier footing process to a fuzzy membership function.

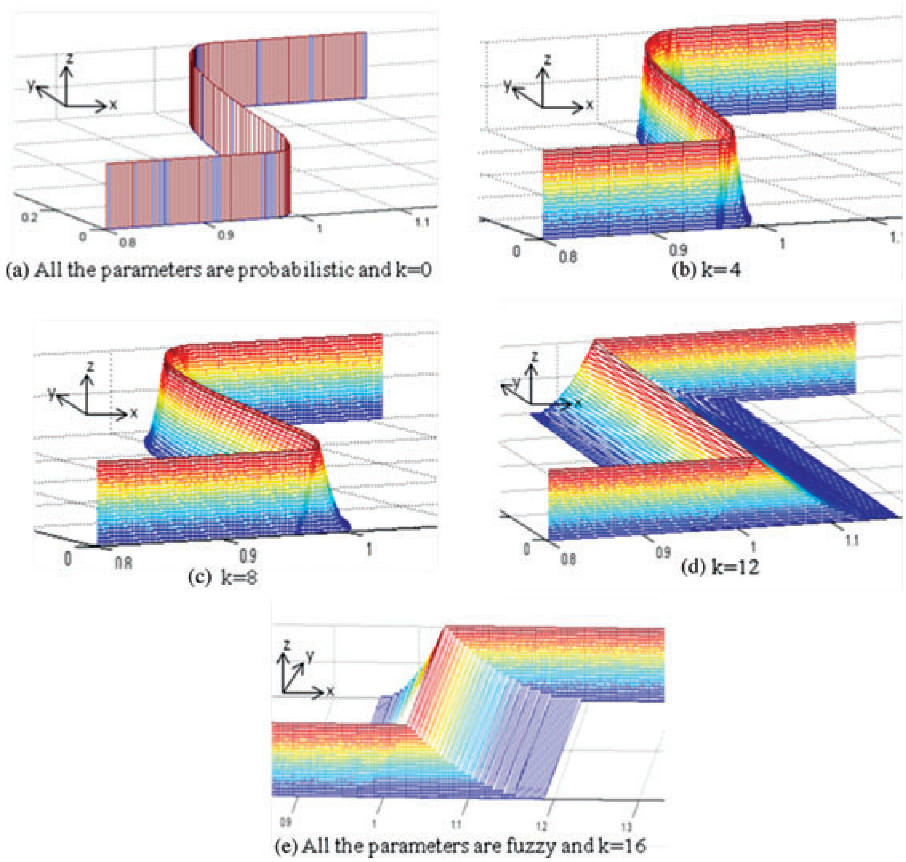


Fig. 14. Three-dimensional view of fuzzy CDF resulting from the output of FMCS for the highway overpass project; k indicates the number of fuzzy sets in each experiment.

rather that the fuzzy sets be derived directly (e.g., from expert judgment). However, these transformations are performed in this study only to be able to compare the FMCS framework with traditional Monte Carlo simulation.

In Section 5.1, we perform a sensitivity analysis to investigate the effect of different numbers of fuzzy sets on the output of the FMCS framework. Section 5.2 discusses how similar decisions that can be made using Monte Carlo simulation can be made based on the FMCS framework.

5.1 Sensitivity analysis of the FMCS framework

For experimenting with differing numbers of fuzzy sets as inputs to the FMCS framework, we select the first k uncertain variables from Table 2, which are not constant, and transform them into fuzzy sets, although keeping the rest of the inputs as probabilistic distributions. For example, when k equals 4, the unit cost and quantity of the excavation and backfill processes are transformed into fuzzy numbers, because these parameters comprise the first four uncertain parameters in Table 2. We gradually increase the value of k in each experiment. The total number of uncertain variables in Table 2 is 16; therefore, when k equals 16, all of the uncertainty is in the form of fuzzy numbers, and we have no randomness in the model. Other variables are constant and are considered as crisp values in the model.

Figure 14 illustrates the three-dimensional graphs of the fuzzy CDFs that are generated by MATLAB for several experiments performed using the FMCS framework. The x-axis indicates the total cost of the model in millions of dollars, the y-axis is the probability, and the z-axis is the α value associated with each output. Therefore, these graphs illustrate both probabilistic and fuzzy uncertainty. We can see how the fuzziness of the output increases when the number of fuzzy inputs (k) increases, illustrating the intuitively appealing behavior of the method.

The x–y view of the fuzzy CDF is also represented in Figure 15. These figures illustrate the CDF bounds of

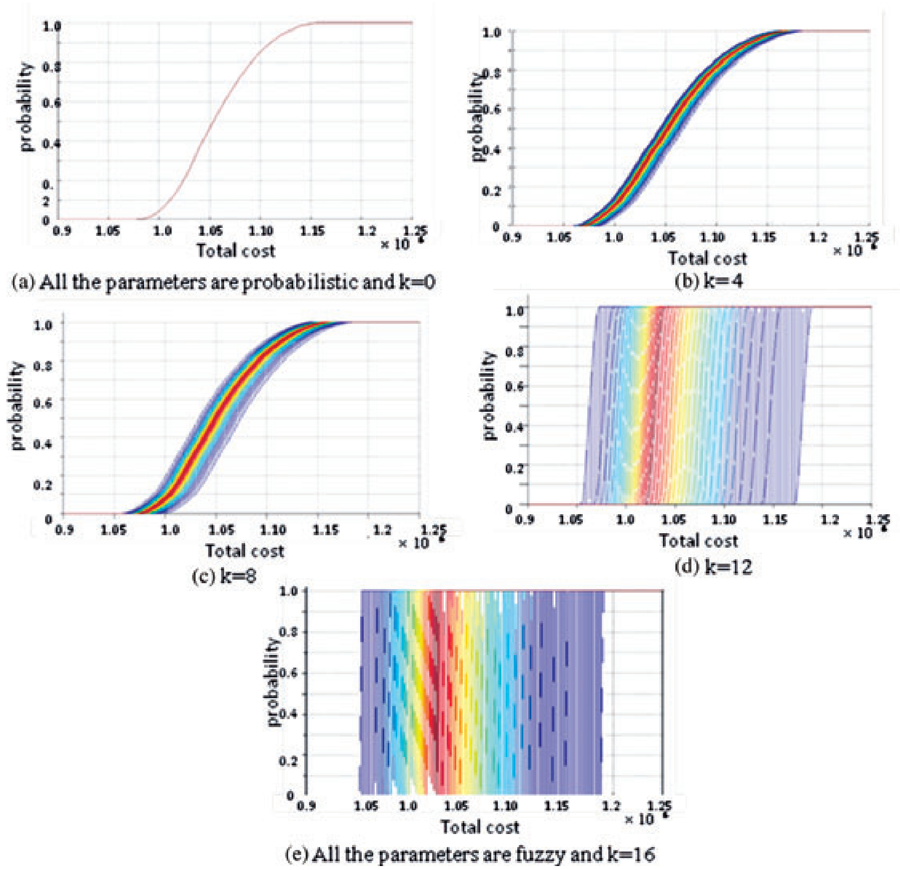


Fig. 15. *x*–*y* view of output results of Figure 14; *k* indicates the number of fuzzy sets in each experiment.

the output of the experiments for different values of *k*. As expected, for smaller values of *k*, the CDF function has less fuzziness, and the CDF bound is narrower.

The reasonable behavior of the FMCS in the absence of fuzziness or randomness is also illustrated by these experiments. If a traditional Monte Carlo cost range estimating model is developed using the inputs of Table 2, the output will be exactly equal to the one shown in Figure 15a. Therefore, in the absence of fuzziness, the results of the proposed methodology in Figure 15a will be exactly equal to the traditional CDF derived from the purely probabilistic Monte Carlo simulation method. Also, the results in Figure 15e indicate that when we have no randomness in the model, the CDF bound does not contain any probabilistic information. However, the fuzzy information can be viewed in the *x*–*z* view of the output in Figure 16. This figure is exactly equal to the output of the same model solved using the purely fuzzy cost range estimating method suggested by Shaheen et al. (2007). This example illustrates the reasonable behavior of the proposed methodology in the sense that the output is similar to a purely fuzzy model in the absence of randomness.

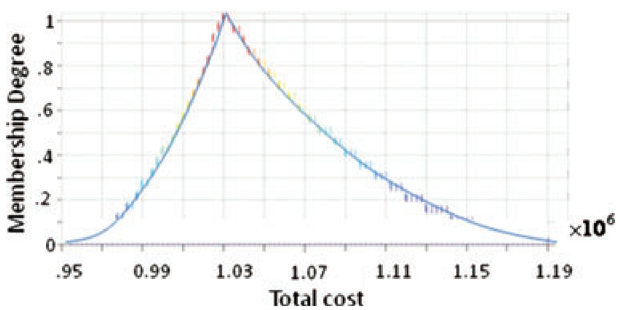


Fig. 16. The fuzzy information of the output in the absence of randomness ($k = 16$).

5.2 Decision making based on fuzzy CDF

Similar to the CDF function resulting from Monte Carlo simulation, an estimator can use the fuzzy CDF of the total cost to estimate the probability of finishing the project within a certain budget. For example, using the CDF function for *k* equals 4 in which we transform the first 4 uncertain variables of the example by Ahuja et al. (1994) to fuzzy sets, Figure 17a indicates how the

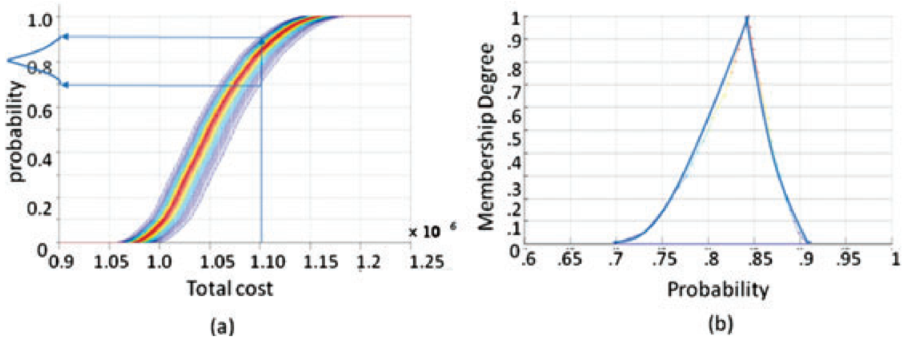


Fig. 17. (a) Intersecting the fuzzy CDF to find the probability of not exceeding a specific threshold. (b) The fuzzy set representing this probability.

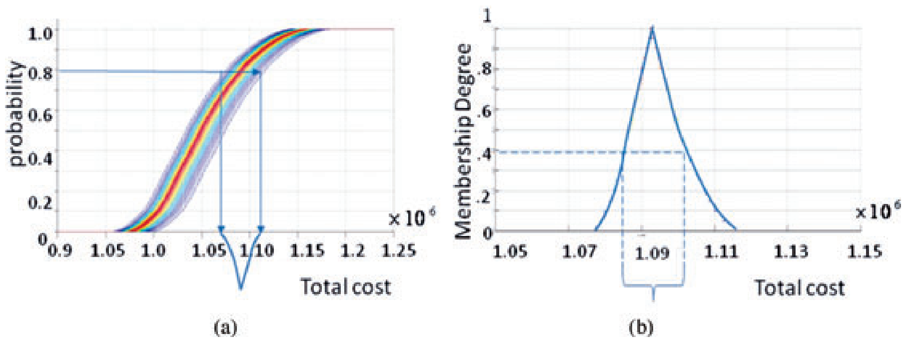


Fig. 18. (a) Intersecting the fuzzy CDF to find an arbitrary quantile. (b) The fuzzy set representing this arbitrary quantile.

probability that the total cost of the project will be less than $1,100,000 is calculated by assigning $1,100,000 to the *x*-axis of the fuzzy CDF. This probability is in the form of a fuzzy set, as shown in Figure 17b. We can defuzzify a fuzzy set to get a crisp value and use that value for decision making. The centroid method is one of the most common methods for defuzzification, in which the defuzzified value is calculated by finding the center of the area under the membership function. By defuzzifying the fuzzy output of Figure 17b using the centroid method of defuzzification, we can state that the probability of finishing this project with $1,100,000 is about 0.82.

An arbitrary quantile can be used to find an appropriate contingency value for a project. Traditionally, this decision is made by considering a quantile value and using the CDF to find the output. In a fuzzy CDF, the arbitrary quantile is in the form of a fuzzy set. Figure 18a illustrates how the 80th quantile of the total cost of the project is calculated by intersecting the *y*-axis of the fuzzy CDF at y is equal to 0.80. The result indicates that, with 80% probability, a budget of around $1,095,000 is enough for recovering the total cost of the project (Figure 18b).

We should note that the real intent of FMCS framework is not to defuzzify the output results, but rather to indicate the fuzziness that exists in the output and to allow the estimator to use his/her subjective judgment in deciding on a final value. After obtaining the output fuzzy sets using the above-mentioned methods, the estimator has to decide on a confidence level between 0 (no confidence) and 1 (full confidence) to get a range of values. This range is calculated by finding the α-cut at the value of 1 minus the confidence level (Mauris et al., 2001). The final value should be selected from this range based on the optimistic or pessimistic view of the decision maker. For example, a manager may wish to estimate a final bid price based on the fuzzy set obtained from the 80th quantile of the project. If the decision maker chooses 0.6 as the confidence level, the α-cut at the level of 0.4 (1 – 0.6) represents the range of outputs [1,085,000, 1,100,000] (Figure 18b). Finally, the decision maker can choose the bid price from this range. For example, a conservative decision maker may go for the Sup value of this range, which is $1,100,000.

6 CONCLUSIONS

This article proposes a FMCS framework as a generalized form of Monte Carlo simulation for modeling construction projects. This framework is capable of considering both fuzzy and probabilistic uncertainty in a problem. In FMCS, the output is modeled using fuzzy random variables. We have introduced the fuzzy CDF as a generalized form of CDF. Fuzzy CDF has the unique feature of representing both fuzzy and probabilistic uncertainty in a single figure. The proposed FMCS framework is capable of considering imprecise information in the form of fuzzy sets without assuming probabilistic information that is not actually available in a simulation model. Therefore, the decision maker is presented with the uncertainty in the output in the form of fuzziness and probabilistic uncertainty, and he/she can use subjective judgment and experience to make the final decision. Practical examples are suggested for applying the FMCS framework on real construction projects. However, actual testing on real projects by industry personnel should be conducted to better justify the benefits that FMCS framework brings to the construction industry.

Examples are provided indicating that the FMCS framework is very effective for providing decision support for risk assessment of construction projects. We have applied FMCS to develop a cost range estimating template for construction projects. The template is used for sensitivity analysis of the FMCS framework based on a highway overpass example. The results illustrate the reasonable behavior of the FMCS framework.

Finally, although the fuzzy CDF is developed as part of the proposed FMCS framework, the fuzzy CDF approach is a general method based on fuzzy random variables and may be used for risk analysis in any application, in which both fuzzy and probabilistic uncertainty are involved. For example, fuzzy CDF can be used in measurement theory to analyze the uncertainty of the data resulting from measurements in cases in which there is both probabilistic and fuzzy uncertainty (for example, Terán, 2007).

REFERENCES

Abebe, A. J., Guinot, V. & Solomatine, D. P. (2000), Fuzzy alpha-cut vs. Monte Carlo techniques in assessing uncertainty in model parameters, in *Proceeding of the 4th International Conference of Hydroinformatics*. Iowa, USA.

Adeli, H. & Jiang, X. (2003), Neuro-fuzzy logic model for freeway work zone capacity estimation, *Journal of Transportation Engineering*, **129**(5), 484–93.

Adeli, H. & Wu, M. (1998), Regularization neural network for construction cost estimation, *Journal of Construction Engineering and Management*, **124**(1), 18–24.

Ahuja, H., Dozzi, S. & Abourisk, S. (1994), *Project Management, Techniques in Planning and Controlling Construction Projects*, John Wiley and Sons, New York.

Baudrit, C., Guyonnet, D. & Dubois, D. (2005), Postprocessing the hybrid method for addressing uncertainty in risk assessments, *Journal of Environmental Engineering*, **131**(12), 1750–54.

Carden, E. P. & Brownjohn, J. M. W. (2008), Fuzzy clustering of stability diagrams for vibration-based structural health monitoring, *Computer-Aided Civil and Infrastructure Engineering*, **23**(5), 360–72.

Chang, P. & Hung, K. (2006), α-cut fuzzy arithmetic, simplifying rules and a fuzzy function optimization with a decision variable, *IEEE Transactions on Fuzzy Systems*, **14**(4), 496–510.

Chao, L. C. (2007), Fuzzy logic model for determining minimum bid markup, *Computer-Aided Civil and Infrastructure Engineering*, **22**(6), 449–60.

Dubois, D., Foulloy, L., Mauris, G. & Prade, H. (2004), Probability-possibility transformation, triangular fuzzy sets and probabilistic inequalities, *Reliable Computing*, **10**(4), 273–97.

Fayek, A. R. & Oduba, A. (2005), Predicting industrial construction labour productivity using fuzzy expert systems, *Journal of Construction Engineering and Management*, **131**(8), 938–41.

Garthwaite, P., Kadane, J. & O'Hagan, A. (2005), Statistical methods for eliciting probability distributions, *Journal of the American Statistical Association*, **100**, 680–701.

Goldstein, M. (2006), Subjective Bayesian analysis: principles and practice, *International Society for Bayesian Analysis*, **1**(3), 403–20.

Guyonnet, D., Bourgigne, B., Dubois, D., Fargier, H., Côme, B. & Chilès, J. (2003), Hybrid approach for addressing uncertainty in risk assessments, *Journal of Environmental Engineering*, **129**(1), 68–78.

Hajjar, D. & AbouRizk, S. (1999), Simphony: an environment for building special purpose construction simulation tools, in *Proceedings of the 1999 Winter Simulation Conference*, Phoenix.

Hanss, M. (2004), *Applied Fuzzy Arithmetic: An Introduction with Engineering Applications*, Springer, Berlin.

Jiang, X. & Adeli, H. (2008), Dynamic fuzzy wavelet neuroemulator for nonlinear control of irregular highrise building structures, *International Journal for Numerical Methods in Engineering*, **74**(7), 1045–66.

Jin, X. L. & Doloi, H. (2009), Modelling risk allocation decision-making in ppp projects using fuzzy logic, *Computer-Aided Civil and Infrastructure Engineering*, **24**(7), 509–24.

Lorterapong, P. & Moselhi, O. (1996), Project-network analysis using fuzzy sets theory, *Journal of Construction Engineering and Management*, **122**(4), 308–18.

Mauris, G., Lasserre, V. & Foulloy, L. (2001), A fuzzy approach for the expression of uncertainty in measurement, *Measurement*, **29**(3), 165–77.

Nomura, Y., Furuta, H. & Hiorkane, M. (2007), An integrated fuzzy control system for structural vibrations, *Computer-Aided Civil and Infrastructure Engineering*, **22**(4), 306–16.

Pedrycz, W. (1998), *An Introduction to Fuzzy Sets, Analysis, and Design*, MIT Press, Cambridge, MA.

Pedrycz, W. & Gomide, F. (1998), *An Introduction to Fuzzy Sets: Analysis and Design*, MIT Press, Cambridge, MA.

Pedrycz, W. & Gomide, F. (2007), *Fuzzy Systems Engineering, Toward Human-Centric Computing*, John Wiley and Sons, New York.

Sarma, K. & Adeli, H. (2000a), Fuzzy genetic algorithm for optimization of steel structures, *Journal of Structural Engineering*, **126**(5), 596–604.