# Use of Fuzzy Logic for Predicting Design Cost Overruns on Building Projects

Karla Knight[1] and Aminah Robinson Fayek, A.M.ASCE[2]

**Abstract:** This paper describes a model for use in predicting potential cost overruns on engineering design projects. The output of the model is useful in assessing the amount of possible risk on a project and the likelihood of making a profit on the job. The model is intended for use by engineering consultants, i.e., structural, mechanical, and electrical engineers, in the building design industry. This research uses fuzzy logic to model the relationships between the characteristics of a project and the potential risk events that may occur, and the associated cost overruns caused by combinations of the project characteristics and risk events. This paper discusses the topic of scope creep and scope definition, which are significant causes of cost overruns on design projects; identifies the project characteristics and risk events used in the model; explains the structure of the model and the use of fuzzy logic; and provides recommendations for future research.

**DOI:** 10.1061/(ASCE)0733-9364(2002)128:6(503)

**CE Database keywords:** Cost control; Decision making; Fuzzy sets; Project management; Construction.



## Introduction

The structural design, mechanical design, and electrical design of a building are important phases in a project. The engineers and designers must produce a design within a given time frame that meets the needs of the building users, is straightforward in its constructability, contains no design errors, and is within certain cost limitations both in creating the design and in constructing the facility. The designers carry a large responsibility in properly managing the project design in order to achieve a successful project both by their own standards and those of the client.

A design firm must apply good project management techniques and practices in order to maintain a well-managed and successful project. Poor project management, among other factors, can lead to negative cost impacts for the design consultant and the entire project team. Although the design phase represents a small portion of the overall cost of a construction project, McGeorge (1988) believes that ‘‘constructability and value engineering [in the design phase] yield construction cost savings of ten to twenty times the cost of the extra design input.’’ For example, McGeorge claims that an increase in design input of 50% will lead to a 10% savings in construction costs. A better design and better management of the design process will lead to a more organized and controlled project, and therefore a more successful project. Design deviations account for approximately 60% of construction project deviations, and poor quality design and engineering can cost almost 10% of the total project cost, according to Bubshait et al. (1999). Many factors, then, that affect design costs will affect the entire project, and therefore impact the overall project costs.

Research on construction design is limited by the lack of documented data available from previous projects. The model presented in this paper attempts to predict design cost overruns or cost savings using data that are realistic to obtain. It identifies the main factors that contribute to escalating costs for the design consultant on construction projects. A list of the main project characteristics and risk events that affect design projects has been compiled. The model uses fuzzy logic to relate the existence of these project characteristics and the occurrence of these risk events, to determine the overall cost impact on the design phase of the project.

## Scope Definition

A lack of scope definition at the onset of a project is one of the main causes of cost overruns during design. The Construction Industry Institute’s (CII) publication on scope definition and control (CII 1986) ranks the lack of scope definition during engineering as having the second highest impact on cost overruns. Dysert (1997) claims that ‘‘poor scope definition at the estimate stage and loss of control of project scope’’ are the most frequent contributors to cost overruns. A poorly defined project is subject to changes initiated by the client that will require extra work and effort by the design team to complete.

With a poorly defined scope there is no baseline against which changes can be evaluated and monitored to identify those that are not within the original scope of work. According to Dumont et al. (1997), these changes may result in cost overruns and a greater potential for disputes. Dumont et al. also claim that these changes may ‘‘delay the project schedule, cause rework, disrupt project rhythm, and lower the productivity and morale of the workforce’’; but an increased level of scope definition will ‘‘improve the accuracy of cost and schedule estimates as well as the probability of

[1]Project Manager, Earth Tech (Canada), Inc., 17203-103 Ave., Edmonton, AB, T5S 1J4, Canada.

[2]Associate Professor, Dept. of Civil and Environmental Engineering, Univ. of Alberta, Edmonton, AB, T6G 2G7, Canada.

Note. Discussion open until May 1, 2003. Separate discussions must be submitted for individual papers. To extend the closing date by one month, a written request must be filed with the ASCE Managing Editor. The manuscript for this paper was submitted for review and possible publication on March 28, 2001; approved on August 21, 2001. This paper is part of the ***Journal of Construction Engineering and Management***, Vol. 128, No. 6, December 1, 2002. ©ASCE, ISSN 0733-9364/2002/6-503–512/$8.00+$.50 per page.



J. Constr. Eng. Manage., 2002, 128(6): 503-512

meeting or exceeding project objectives.'' A survey done by Bresnen and Haslam (1991) showed that of projects surveyed that were overbudget, 40% of the time it was due to additional work and/or design variations.

A poorly defined scope can also be caused by internal problems within a design firm. If the project manager does not adequately define to the designers and draftspeople the project and tasks to be done, extra work may be performed due to lack of direction. A project manager must have good communication with the design team and provide guidance and direction to ensure that everyone is working toward the same goal. Misunderstandings and misinterpretations between parties can cause problems such as rework and extra work, and will invariably cause disputes. Good communication, good organization and control, and proper scope definition are therefore important elements of a successful project.

## Previous Studies on Design Cost Overruns

Literature on design cost control and cost overruns is very limited. A study by Cox et al. (1999) reviews the causes and effects of design changes on a construction project, and Glavinich (1995) reports on the issue of constructability and its effect on the construction process. Love et al. (2000) created a model that simulates the causes and effects of design errors. The CII has a publication entitled ''Scope definition and control'' (CII 1986) that discusses the effects of a poorly defined scope. The CII has also prepared a ''project definition rating index'' (PDRI) for industrial projects (CII 1996; Dumont et al. 1997), and a PDRI for commercial building projects (CII 1999). The PDRI for building projects has a checklist of scope definition elements and a score sheet that allows the user to rate each element to determine the level of scope definition on the project. The risks to a consultant due to the client have also been researched and modeled by Kometa et al. (1995, 1996).

Many of the factors that affect design costs have been identified and studied, but no tools are currently in place for predicting design cost overruns. Project management of the design phase is underresearched in comparison to the amount of research that has been done on project tracking, monitoring, and control for the construction phase of a project. Factors for the proposed model were chosen partly through literature that reviewed the issues that are important to design work, and partly through discussion with project engineers and other personnel working with a design firm.

## Fuzzy Logic in Construction

Fuzzy logic is the modeling technique used in the model to predict design cost overruns. Fuzzy logic has become an increasingly popular modeling tool in construction research. Fuzzy logic and fuzzy expert systems are used increasingly in situations where little deterministic data are available. ''Fuzzy set theory was originally devised to model uncertainty associated with human perception or subjective probability judgments'' (Nguyen 1985).

Due to the imprecise nature of many factors that affect construction projects, and a general lack of data for proper quantification of factors, fuzzy logic lends itself well to many construction applications. The measurement of construction factors is often subjective and uncertain, where actual data are not available or when the data come from the experience and judgment of those in the industry. For this reason, fuzzy logic is being used more and more to model construction issues where the process was previously only available in the mind of an experienced construction practitioner. Fuzzy logic supports the use of linguistic variables such as ''high experience'' or ''bad weather'' and allows for ranking or subjective rating of factors.

## Model for Predicting Design Cost Overruns

The goal of the model presented in this paper is to accurately predict cost overruns and underruns, as a percentage of the contract fee, based on the degree of existence of certain project characteristics and the degree of occurrence of certain risk/opportunity events. The model uses 13 project characteristics and eight risk events that are rated by the user in reference to a specific project. Standard strengths within the model represent the relationships between the project characteristics and risk events, and the potential cost overrun/underrun that combinations of these factors may cause. Based on the user input ratings of the project characteristics and risk events, the model calculates the predicted cost overrun or underrun above or below the consultant's fixed engineering fee. A cost overrun, as defined in the model, occurs when the actual cost of the work (including profit and overheads) to the design firm is above the fee paid to the design firm by the client to complete the work. It is assumed that a project fee is estimated at the onset of the project and this fee is set as the maximum upset price for the design work.

The data used to build the model are based on previous projects and information obtained from project managers within a local leading design firm. The model is applicable to the work of structural, mechanical, and electrical design engineers performing work in the commercial building construction sector. The engineering design firm should be either the consultant (i.e., a consultant to the prime consultant; typically the design firm) or the prime consultant (i.e., a consultant directly to the client; typically the architect), and the contract a guaranteed maximum price contract where the fee is estimated by the project manager or determined as a percentage of the total project cost. The project size can vary, but the model is most applicable to projects with an engineering fee in the \$100,000–\$500,000 range.

The model is intended to be used during the proposal stage of the project (i.e., as the consultant is preparing the fee estimate) or at the onset of the consultant's work on the project. The model is intended for use in conducting a risk analysis or for project awareness, and not as a guide for estimating. The consultant's estimate of the fee should be complete at this point, and the model will identify the sensitivity of the project to certain characteristics and risk events.

## Framework of Model

The structure of the model is shown in Fig. 1, and each of its components is discussed next.

### *Project Characteristics*

There are many project characteristics that can potentially affect the costs of a project, both in a negative way and in a positive way. The project characteristics that are most recognizable in their effect on project costs, and the project characteristics that potentially have the greatest impact on project costs have been identified in consultation with the collaborating design firm. The project characteristics ($P_i$) used in this model are the following:



J. Constr. Eng. Manage., 2002, 128(6): 503-512

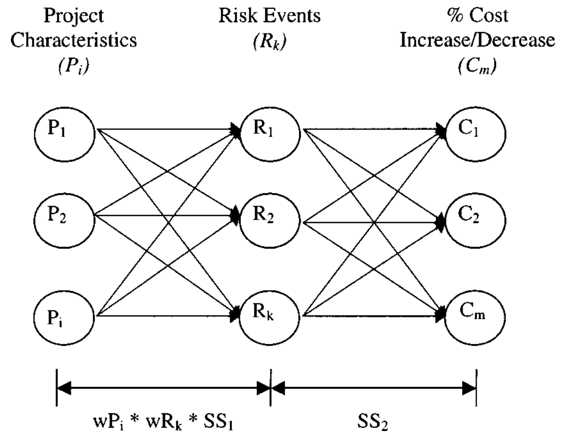


**Fig. 1.** Structure of model for predicting design cost overruns

1. Willingness of the prime consultant to approach the client for extra fees,
2. Time taken by the client/prime consultant/architect/engineer to make decisions,
3. Knowledge base of the client,
4. Level of project scope definition between the consultant and the client/prime consultant at the proposal stage,
5. Definition of scope duties passed on by the consultant's project manager to the design team,
6. Experience of the consultant's project manager,
7. Experience of the prime consultant's project lead or project manager,
8. Skill set of the consultant's design team,
9. Skill set of the prime consultant's design team,
10. Experience of the project team with similar projects,
11. Project complexity,
12. Timeline for design and construction, and
13. Project location.

Each of these project characteristics is given a rating between 0 and 10 by the user. This rating, $wP_i$, represents the degree of existence of the project characteristic, $P_i$, on a specific project. A rating of 0 indicates *poor*, and a rating of 10 is an *excellent* rating. A rating of 5 indicates that the project characteristic has a *neutral* effect on the project, or that the characteristic exists in a state that is typical or expected on a project.

### Risk Events

Risk events are undesirable or fortuitous events that may occur during the design and construction phases of a project, potentially leading to a cost overrun or savings, respectively. Risk should always be taken into account on a project, but due to the competitive nature of the industry not every possible risk event can be accounted for in the project fee. This puts the onus on the consultant and the contractor to deal with, and possibly cover the cost of, both negative and positive risk events. Eight risk events have been selected in consultation with the design firm as being the most common or having the largest impacts on projects. The chosen risk events ($R_k$) are

1. Accuracy of site investigation,
2. Design errors or omissions,
3. Design/scope changes by the client, consultants, or architect,
4. Communication among the project team,
5. Overengineering,
6. Constructability issues,

**Table 1.** Ranges of Cost Increases and Decreases for Large and Small Projects

| Parameter | Cost range, large project (%) | Cost range, small project (%) | Linguistic descriptor |
|---|---|---|---|
| $C_1$ | $-50$ to $-20$ | $-100$ to $-80$ | Very high |
| $C_2$ | $-20$ to $-10$ | $-80$ to $-50$ | High |
| $C_3$ | $-10$ to $-5$ | $-50$ to $-20$ | Medium |
| $C_4$ | $-5$–0 | $-20$–0 | Low |
| $C_5$ | 0 | 0 | Zero |
| $C_6$ | 0 | 0 | Zero |
| $C_7$ | 0–5 | 0–20 | Low |
| $C_8$ | 5–10 | 20–50 | Medium |
| $C_9$ | 10–20 | 50–80 | High |
| $C_{10}$ | 20–100 | 80–100 | Very high |

7. Inadequate design team resources, and
8. Adequacy of the general contractor and subcontractors.

As with the project characteristics, each of these risk events is given a rating by the user between 0 and 10. This rating, $wR_k$, represents the extent to which the risk event, $R_k$, is likely to occur and the extent of its impact on the specific project. A rating of 0 indicates the worst case, and a rating of 10 is the best case. A rating of 5 indicates that the risk event is likely to have a neutral effect on project costs.

The model proposed here does not take into account the probability of the risk event occurring given that the project characteristics exist to a certain degree, but rather the extent of the combined impact of the existence of project characteristics and the occurrence of risk events. The project characteristics and risk events are considered independent factors in the model, with each combination of project characteristic and risk event having a potential impact on design costs.

### Percent Cost Increase/Decrease

Each combination of project characteristic and risk event results in a potential cost increase or decrease ($C_m$) above or below the contract fee, expressed as a percentage of the contract fee. This percentage is the expected cost overrun, above the contract fee, given that the project characteristic exists to the worst degree possible and the risk event occurs to the worst degree. Or, for a best-case scenario, the percentage is the expected cost underrun, below the contract fee, given that the project characteristic exists to the best degree possible and the risk event occurs to the best degree. Each cost increase/decrease, or $C_m$, is a percentage range that has been identified through expert opinion, as shown in Table 1, and is hard-coded in the model. The ranges each have a linguistic descriptor ranging from very high negative to very high positive. The expert opinions were solicited by means of a panel discussion with five project managers within the collaborating design firm.

### Standard Strengths

There are two sets of standard strengths in the model—both of which were identified by the same panel of experts described earlier. The standard strengths are hard-coded in the model. The first, $SS_1$, relates the project characteristics and risk events; the second, $SS_2$, relates the risk events and cost ranges. The standard strength $SS_1$ represents the sensitivity of the impact of the risk



J. Constr. Eng. Manage., 2002, 128(6): 503-512

event to variations in the project characteristic. For example, if risk event $R$ occurs to the worst degree possible, how strong is its impact on the project cost, based on the degree of existence of project characteristic $P$? The cost impact of the risk event, then, is dependent on the degree to which the project characteristic exists.

The second standard strength, or $SS_2$, represents the strength between each risk event, $R_k$, and cost range, $C_m$. The cost range that was identified by the experts as the most likely cost overrun or underrun given a worst-case or best-case scenario of combinations of each project characteristic and risk event, receives a strength of 1.0. The neighboring cost ranges receive strengths decreasing outwardly in increments of 0.2. These incremental values give the decreasing likelihood of the cost ranges occurring. These standard strengths account for the fact that the chosen cost range may be the most likely to occur, but the other ranges may also occur but with less likelihood.

## Use of Fuzzy Set Theory in Model

The model presented in this paper uses fuzzy binary relations to approximate the relationships between the project characteristics and risk events in order to predict the expected cost overrun or underrun. Binary relations are a fuzzy logic technique that involves approximating the relationship between two data sets given the degree of association between the sets. The advantage of using binary relations is that they do not rely on membership functions, which can require substantial data sets formed on expert opinions. This is unfortunately also the biggest weakness of binary relations, because the user must provide ratings instead, which are often subjective and relative to the user's context. The method used to overcome this weakness was to solicit opinions from a panel of experts, as described, and to base the values used on a consensus of opinion. A fuzzy binary relation can be represented by a matrix, with the values of the matrix representing the values of the membership grades between sets. The fuzzy relation allows for partial membership, as opposed to the crisp binary relation, which only allows for values of 0 or 1, or in other words, either the presence or absence of association (Klir and Yuan 1995).

The first binary relation formed is between the project characteristics and risk events, called the $S(P,R)$ relation. This relation takes into account the user ratings of the project characteristics and the risk events for the specific project under analysis, and the preset standard strengths relating project characteristics and risk events. The $S(P,R)$ relation is calculated using Eq. (1)

$$S(P,R)=wP_i*wR_k*SS_{1ik} \tag{1}$$

where $S(P,R)$=fuzzy binary relation between the project characteristics and risk events, from 0.0 to 1.0; $wP_i$=rating of project characteristic $i$ from 0 to 10, input by the user and scaled by a factor of 10; $wR_k$=rating of risk event $k$ from 0 to 10, input by the user and scaled by a factor of 10; and $SS_{1ik}$=standard strength between project characteristic $i$ and risk event $k$, from 0.0 to 1.0.

The second binary relation, the $F(R,C)$ relation, relates the project characteristics and risk events to percentage cost overruns/underruns, and is equal to values ranging from 0.0 to 1.0 derived based on $SS_2$, the preset standard strengths relating risk events and cost ranges. For each combination of project characteristic and risk event, a most likely range of percentage cost overrun and underrun has been identified through expert opinion. The ranges of cost overruns and underruns used are shown in Table 1. The membership values, or standard strengths, for this matrix are values decreasing outwardly by increments of 0.2, starting with a value of 1.0 for the percentage range identified as the most likely cost overrun for that combination of project characteristics and risk events. For example, if $C_3$ receives a value of 1.0 for $SS_2$, then $C_2$ and $C_4$ each receive a value of 0.8, $C_1$ and $C_5$ each receive a value of 0.6, all the way down to a value of 0, or until the end of the matrix is met. These incremental values give the decreasing likelihood of the cost ranges occurring. They account for the fact that the chosen cost range may be the most likely to occur, but the other ranges may also occur but with less likelihood. This is one of the advantages of using fuzzy logic.

Once these two binary relations are formed, they are combined to produce a third binary relation using the composition operation. This operation, shown in Eq. (2), yields the fuzzy binary relation $Q(P,C)$, which relates the project characteristics to the percent overrun/underrun through their respective relationship to the risk events

$$Q(P,C)=SoF(P,C)=S(P,R)oF(R,C) \tag{2}$$

Two types of composition operations are used and compared—the maximum-minimum (max-min) composition operation and the cumulative-minimum (cum-min) composition operation. The max-min composition operation, shown in Eq. (3), determines the most likely solution based on the strongest indicator or piece of evidence. The cum-min composition operation, shown in Eq. (4), takes into account each factor that points to the output value and increases the strength with which that output value is recommended based on all supporting evidence. The accuracy of each of these two composition operations is compared during the model testing and sensitivity analysis

$$SoF(P,C)=\text{max-min}[S(P,R),F(R,C)] \quad \text{for all } R \tag{3}$$

$$SoF(P,C)=\text{sum-min}[S(P,R),F(R,C)] \quad \text{for all } R \tag{4}$$

After the composition operation is used to combine the $S(P,R)$ and the $F(R,C)$ relations to obtain the $Q(P,C)$ relation, the total strength of each percentage range is calculated using Eq. (5). This equation takes the sum of the strengths for each percentage range, and divides the sum by the total of the project characteristic ratings. This operation for combining the strengths is analogous to the statistical concept of weighted means. The "strengths" are the membership values of the $Q(P,C)$ relation for each percentage range. The highest strength points to the most likely range of percent cost overrun/underrun, based on the standard strengths within the model and the user input ratings for the specific project under analysis

$$Q(P,C_m)=\frac{\Sigma_{n=1}^{i}Q(P_n,C_m)}{\Sigma_{n=1}^{i}wP_n} \tag{5}$$

where $Q(P,C_m)$=total strength with which cost range $C_m$ is expected to occur; $\Sigma\, Q(P_n,C_m)$=sum of the membership values for each $C_m$ element from the $Q(P,C)$ relation; and $\Sigma\, wP_n$=sum of the user input project characteristic ratings.

The output of the model can now be defuzzified to obtain a single (crisp) value, $C^*$. Defuzzification methods use the strength for each percentage range to recommend the most likely cost percentage, either by calculating the area under the curve using the center of area (COA) method, or simply by examining the percentage range with the highest strength [the mean of maximum, largest of maximum (LOM), smallest of maximum (SOM), and middle of maximum (MOM) methods]. The center of area method, shown in Eq. (6), takes the centroid, or midpoint, of each range and multiplies it by its membership value, then divides by the sum of the membership values. The values for LOM, SOM,



J. Constr. Eng. Manage., 2002, 128(6): 503-512

and MOM are the largest, smallest, and middle values of the range with the highest membership value. Like with the composition operations, these methods of defuzzification are compared during the model testing and sensitivity analysis

$$C^* = \frac{\sum_{m\ \min}^{m\ \max} m \cdot C(m)}{\sum_{m\ \min}^{m\ \max} C(m)} \tag{6}$$

where $C^*$=crisp defuzzified output; $m$=midpoint value of the numerical range of $C_m$; and $C(m)=Q(P,C_m)$=total strength with which cost range $C_m$ is expected to occur.

### *Accounting for Cost Overruns and Underruns*

Although the primary goal of this research is to predict cost overruns from the project characteristics and risk events on a project, potential cost underruns must also be taken into account. The standard strengths for the model are identified assuming a worst-case scenario for all project characteristics and risk events. In using this worst-case scenario approach, potential cost underruns are not accounted for. Therefore, a best-case scenario with standard strengths is also provided for the model. The calculation for the predicted cost overrun/underrun is actually determined through two calculations; the first is a worst-case scenario, and the second is a best-case scenario.

One scenario is the inverse of the other, except that each has its own standard strengths representing the degree of sensitivity of the impacts of the risk events to variations in the project characteristics, and the expected cost overrun or underrun. Therefore, one model will predict cost overruns and the other will predict cost underruns based on the same user input. The best-case and the worst-case scenarios are joined after the composition operation by summing the membership values in the $Q(P,C)$ relation.

The user input ratings of the factors for the best-case scenario are simply the inverse of the ratings of the worst-case scenario. For example, if the experience of the consultant's project manager has a membership value of 0.2 in *poor*, or the worst-case scenario, then it has a membership value of 0.8 in *excellent*, or the best-case scenario. These calculations are illustrated in the sample project.

## Data Collection and Results

Two interview surveys were used to collect the information needed for developing and calibrating the model. Project managers from the collaborating design firm participated in the interviews, giving the standard strengths that are hard-coded into the model, and the case studies used to test the model. The expert opinions were solicited by means of a panel discussion with five project managers. This discussion provided the standard strengths, $SS_1$ and $SS_2$, and the cost ranges, $C_m$, used in the model. A further survey questionnaire was conducted in the form of an interview with seven project managers; they were each asked to choose two to three completed projects and rate the existence of the project characteristics and the occurrence of the risk events on these projects. They were also asked to provide the fee paid by the client for the work and the actual cost of the work to the company (inclusive of profit and overheads).

An analysis of the data collected from the expert opinions and the project manager interviews provides some insight into the characteristics and risk events on a project that have the most significant impacts and occur most frequently. The project characteristics that were rated as having the greatest potential to affect project costs are the experience level of the consultant's project manager, the skill set of the consultant's design team, and the level of project scope definition between the consultant and the client/prime consultant. The most frequently occurring adverse project characteristics are the time taken by the client/architect/prime consultant/engineer to make decisions, a too short timeline for design and construction, and an unwillingness of the architect/prime consultant to approach the owner for extra fees. The risk events that were identified as having the greatest impacts on costs and that are the most commonly occurring are a poor general contractor and subcontractors; design/scope changes by the client, consultants, or architect; and poor communication among the project team.

## Model Calibration and Validation

The accuracy of the model was tested using case studies obtained from the collaborating design firm, as previously discussed. Eighteen completed projects were obtained and used for testing the model.

The actual cost overrun from the sample projects was calculated by subtracting the final project fee, or billed amount, from the actual project cost to the consultant, and dividing by the billed amount [Eq. (7)]

$$\text{Actual cost overrun/underrun (\%)} = \frac{\text{actual cost (\$)} - \text{billed amount (\$)}}{\text{billed amount (\$)}} \times 100\% \tag{7}$$

The actual cost overrun/underrun was then compared to the model's prediction of the cost overrun/underrun from the defuzzified output. This was the first criterion used in assessing the accuracy of the model—the absolute value of the difference between the actual cost overrun/underrun and the predicted cost overrun/underrun. A difference of 10% above or below the actual value was deemed acceptable. This is an acceptable margin of error, considering the high tolerance for variations in cost overruns and underruns, and the inaccuracy and subjectivity in estimating the contract fee.

The second criterion used for model validation is the accuracy of the associated linguistic variable of the predicted cost overrun/underrun compared to the associated linguistic variable of the actual cost overrun/underrun. Table 1 gives the cost ranges associated with each linguistic descriptor. In the design context, linguistic terms are often used rather than crisp numbers to describe design cost performance, since there tend to be wide margins of variation and therefore tolerance. The use of fuzzy logic in predicting cost overruns and underruns makes linguistic term prediction possible. The accuracy of the model is evaluated by determining how many ranges the predicted value is off from the actual value. Predictions that fall in the same range as the actual value are deemed acceptable.

A need for the model to be calibrated was established based on the first test run of the model. Two problems with the model were found. First, the model was poorly predicting cost underruns; second, the model was underpredicting in certain cases.

The model was poorly predicting cost underruns because the standard strengths between the project characteristics and the risk events for the best-case model were too weak. The experts reasoned that most fee estimates are prepared assuming minimal problems on the project, which in turn assumes a best-case scenario for the project. The standard strengths for the best-case



J. Constr. Eng. Manage., 2002, 128(6): 503-512

**Table 2.** Results of Model Testing

| Composition operation | Defuzzification method | Average % difference (absolute value) | Linguistic term difference=0 | Linguistic term difference=±1 | Linguistic term difference≥±2 |
|---|---|---|---|---|---|
| (a) Large projects (>$55,000) | | | | | |
| cum-min | COA | 10 | 2 | 2 | 6 |
| max-min | COA | 10 | 2 | 2 | 6 |
| cum-min | MOM | 8 | 3 | 3 | 4 |
| max-min | MOM | 10 | 2 | 3 | 5 |
| cum-min[a] | LOM[a] | 9[a] | 5[a] | 2[a] | 3[a] |
| max-min | LOM | 10 | 4 | 2 | 4 |
| cum-min | SOM | 8 | 5 | 3 | 2 |
| max-min | SOM | 9 | 4 | 2 | 4 |
| (b) Small projects (≤$55,000) | | | | | |
| cum-min | COA | 14 | 3 | 1 | 4 |
| max-min | COA | 15 | 2 | 2 | 4 |
| cum-min | MOM | 9 | 4 | 3 | 1 |
| max-min | MOM | 10 | 3 | 4 | 1 |
| cum-min[a] | LOM[a] | 9[a] | 7[a] | 1[a] | 0[a] |
| max-min | LOM | 8 | 7 | 1 | 0 |
| cum-min | SOM | 13 | 4 | 4 | 0 |
| max-min | SOM | 13 | 4 | 4 | 0 |

Note: COA=center of area, MOM=middle of maximum, LOM=largest of maximum, and SOM=smallest of maximum.

[a]Best prediction results.

scenario were therefore quite low; all except two of them were below 5, limiting the influence of the underruns on the project output. The first calibration to the model was to normalize the standard strengths of the best-case model. Two of the existing standard strengths were 8, which became 10, and the remaining strengths at 5 and below were doubled. This calibration increased the influence of the best-case model and enabled the model to predict cost underruns without changing the relationships between the factors in the best-case model.

The next problem with the model was that it was underpredicting certain overruns and underruns. It was observed that the majority of the projects with the high cost overruns and underruns were smaller projects. This is logical, since the percent overrun/underrun is relative to the size of the project. For example, on a small project of $10,000, if the project is $2,000 overbudget this represents a 20% cost overrun. However, on a larger project of $100,000, this would only be a 2% overrun, which is considered quite low. Therefore, it was necessary to separate the large and the small projects, and adjust the ranges used on the small projects. This calibration addressed the problem of underprediction in the majority of cases.

Upon examination of the data, $55,000 (design fee) was chosen as the cutoff between large and small projects. This put eight projects into the small project category, and 10 into the large project category. The cost ranges for the small projects were then amended to attempt to more accurately predict their cost overruns and underruns. The new ranges chosen, based on the given data, are shown in Table 1. These values were chosen to best fit the existing data, and were not chosen by experts, as with the original cost ranges.

### *Validation Results and Sensitivity Analysis*

Two composition operations and four defuzzification methods were tested during the model validation. The composition operations tested were the cum-min operation and the max-min operation; the defuzzification methods tested were the COA, MOM, LOM, and SOM, as discussed earlier. Table 2 shows the results of model testing for each combination of composition operation and defuzzification method, after calibration of the model. The cum-min composition operation and the LOM defuzzification method provide the best prediction results (footnoted items in Table 2). For large projects, the average absolute difference between the actual and the predicted cost overrun/underrun was 9%; for five of these projects, the model predicted the same linguistic descriptor for the magnitude of the overrun/underrun as the actual term; for two of the remaining five projects, the linguistic term prediction was only one term off from the actual. For small projects, the average absolute numerical difference was 9%; seven of the project overruns/underruns were accurately predicted in linguistic terms, and one project was one term off.

## Sample Project

The following sample project illustrates the calculations performed in the model and the use of the model. The use of the model is illustrated by screens captured from the software implementation of the model using Microsoft Visual Basic 6.0 (Microsoft Corp., Redmond, Wash.); Fig. 2 illustrates the required user input, and Fig. 3 illustrates the output of the model. For the sample project calculations, the full range of factors is not used in order to simplify the calculations and reduce the size of the matrices. The output of the sample calculations therefore does not match the output of the model shown in Fig. 3. The sample project is the construction of a clubhouse for the local cross-country ski group in a small remote town in northern Alberta, Canada, with a design fee of $84,300. It is therefore analyzed using the model for large projects. Three project characteristics, time taken to make decisions ($P_1$), project location ($P_2$), and experience of the consultant's project manager ($P_3$), and three risk events, accuracy of the site investigation ($R_A$), adequacy of



J. Constr. Eng. Manage., 2002, 128(6): 503-512

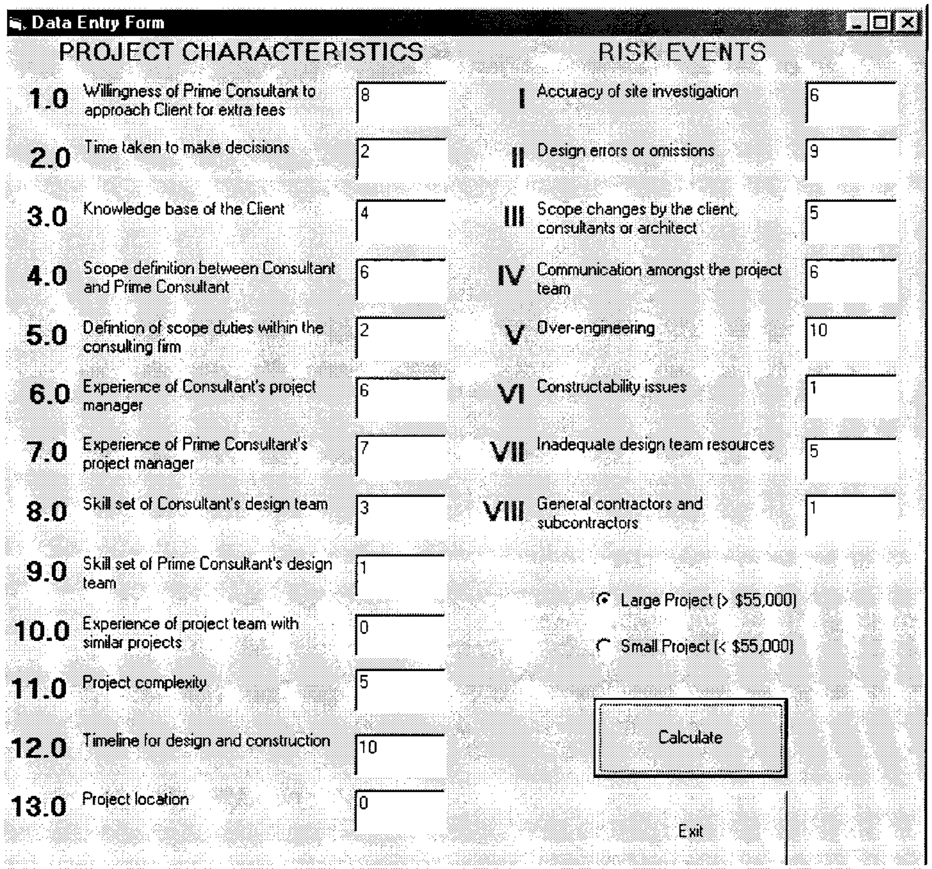


**Fig. 2.** User input form

the general contractor and subcontractors ($R_B$), and inadequate design team resources ($R_C$), are used for illustration purposes. Sample calculations are provided at each step.

### Step 1

Identify the user input ratings for the project characteristics and risk events. The user rates each project characteristic and risk event on a scale from 0 (worst case) to 10 (best case). The model scales the input ratings by a factor of 10, giving numbers between 0 and 1.0 for the best-case scenario, and then calculates the complement of these numbers for the worst-case scenario.

For the sample project, $P_1$, the time taken to make decisions, receives a rating of 2 because the owners are a volunteer organization with 12 members on the project committee and are very slow at making decisions. As explained before, the model converts this to a degree of existence ($wP_1$) for the best-case scenario of 0.2, and 0.8 for the worst-case scenario. $P_2$, project location, receives a rating of 0 because the project is in a small town 6 h north of the consultant's office ($wP_2=0$ for the best-case scenario, and 1.0 for the worst case). $P_3$, the experience of the consultant's project manager, receives a rating of 6, as the engineer is an intermediate engineer with two years of experience as a project manager ($wP_3=0.6$ for the best-case scenario, and 0.4 for the worst case).

Next, the ratings for the risk events are determined. $R_A$, the accuracy of the site investigation, receives a rating of 6 because typically soils in this area have been slightly unpredictable. The degree of occurrence for the best-case scenario, $wR_A$, is 0.6, and 0.4 for the worst-case scenario. $R_B$, the adequacy of the general contractor and subcontractors, is given a rating of 1 because the contractors in this remote location are unknown and all contractors are presently busy with other work, so it will be difficult to get a good contractor for this project ($wR_B=0.1$ for the best-case scenario, and 0.9 for the worst-case scenario). The final risk event, $R_C$, inadequate design team resources, receives a rating of 5, as the design firm is also presently busy, so its designers' commitment to the project may be low at times, although they have enough resources to complete the project ($wR_C=0.5$ for both the best-case scenario and the worst-case scenario).

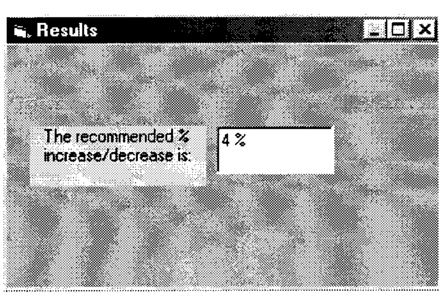


**Fig. 3.** Output form

### Step 2

Identify the standard strengths between the project characteristics and risk events, and between the risk events and the cost ranges



J. Constr. Eng. Manage., 2002, 128(6): 503-512

**Table 3.** Standard Strengths for Sample Project

| $P_nR_k$ | $SS_1$ | Cost range |
|---|---|---|
| | (a) Worst case | |
| 1A | 1.0 | $C_9$ |
| 1B | 0.8 | $C_9$ |
| 1C | 0.9 | $C_8$ |
| 2A | 0.4 | $C_6$ |
| 2B | 0.5 | $C_7$ |
| 2C | 0.9 | $C_{10}$ |
| 3A | 0.3 | $C_6$ |
| 3B | 0.6 | $C_8$ |
| 3C | 0.7 | $C_7$ |
| | (b) Best case | |
| 1A | 0.6 | $C_2$ |
| 1B | 0.5 | $C_3$ |
| 1C | 0.5 | $C_3$ |
| 2A | 0.8 | $C_1$ |
| 2B | 0.3 | $C_4$ |
| 2C | 0.4 | $C_2$ |
| 3A | 0.2 | $C_3$ |
| 3B | 0.1 | $C_4$ |
| 3C | 0.1 | $C_5$ |

(Table 3). These values are hard-coded into the model. The cost ranges used are identified in Table 1 as the cost range for large projects.

### Step 3

Calculate the values of the $S(P,R)$ relation (Table 4).

**Sample Calculation**

$$S(P_i,R_k)=wP_i\times wR_k\times SS_{1ik}$$

$$S(P_1,R_A)=wP_1\times wR_A\times SS_{11A}=0.8\times 0.4\times 1.0=0.32$$

### Step 4

Calculate the values of the $F(R,C)$ relation (the most likely cost overrun/underrun receives a value of 1.0, and the surrounding costs receive values decreasing outwardly by increments of 0.2) (Table 5).

**Table 4.** Membership Values of $S(P,R)$ Relation for Sample Project

| $S(P,R)$ | A | B | C |
|---|---|---|---|
| | (a) Worst case | | |
| 1 | 0.32 | 0.58 | 0.36 |
| 2 | 0.16 | 0.45 | 0.45 |
| 3 | 0.05 | 0.22 | 0.14 |
| | (b) Best case | | |
| 1 | 0.07 | 0.01 | 0.05 |
| 2 | 0.00 | 0.00 | 0.00 |
| 3 | 0.07 | 0.01 | 0.03 |

**Table 5.** Values for $F(R,C)$ Relation for Sample Project

| | Best Case | | | | | Worst Case | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $F(R,C)$ | $C_1$ | $C_2$ | $C_3$ | $C_4$ | $C_5$ | $C_6$ | $C_7$ | $C_8$ | $C_9$ | $C_{10}$ |
| 1A | 0.8 | 1.0 | 0.8 | 0.6 | 0.4 | 0.4 | 0.6 | 0.8 | 1.0 | 0.8 |
| 1B | 0.6 | 0.8 | 1.0 | 0.8 | 0.6 | 0.4 | 0.6 | 0.8 | 1.0 | 0.8 |
| 1C | 0.6 | 0.8 | 1.0 | 0.8 | 0.6 | 0.6 | 0.8 | 1.0 | 0.8 | 0.6 |
| 2A | 1.0 | 0.8 | 0.6 | 0.4 | 0.2 | 1.0 | 0.8 | 0.6 | 0.4 | 0.2 |
| 2B | 0.4 | 0.6 | 0.8 | 1.0 | 0.8 | 0.8 | 1.0 | 0.8 | 0.6 | 0.4 |
| 2C | 0.8 | 1.0 | 0.8 | 0.6 | 0.4 | 0.2 | 0.4 | 0.6 | 0.8 | 1.0 |
| 3A | 0.6 | 0.8 | 1.0 | 0.8 | 0.6 | 1.0 | 0.8 | 0.6 | 0.4 | 0.2 |
| 3B | 0.4 | 0.6 | 0.8 | 1.0 | 0.8 | 0.6 | 0.8 | 1.0 | 0.8 | 0.6 |
| 3C | 0.2 | 0.4 | 0.6 | 0.8 | 1.0 | 0.8 | 1.0 | 0.8 | 0.6 | 0.4 |

### Step 5

Apply the maximum-minimum composition operation and the cumulative-minimum composition operation to obtain the $Q(P,C)$ relation (Table 6).

**Sample Calculation Using max-min**

$$Q(P_i,C_m)=\text{max-min}[S(P_i,R_k),F(R_k,C_m)]$$

$$Q(P_1,C_6)=\text{max-min}[(0.32,0.40),(0.58,0.40),(0.36,0.60)]$$

$$=\max(0.32,0.40,0.36)=0.40$$

Calculate the total strength of each cost range

$$\text{sum}\Big/\sum wP=\sum C_m\Big/\sum wP_i$$

For $C_6$ the total strength $=(0.40+0.45+0.22)/2.2=0.49$.

**Sample Calculation Using cum-min**

$$Q(P_i,C_m)=\text{sum-min}[S(P_i,R_k),F(R_k,C_m)]$$

$$Q(P_1,C_6)=\text{sum-min}[(0.32,0.40),(0.58,0.40),(0.36,0.60)]$$

$$=\text{sum}(0.32,0.40,0.36)=1.08$$

### Step 6

Combine the best-case and the worst-case models by combining the strengths of each cost range calculated in the $Q(P,C)$ relation

**Table 6.** Membership Values of $Q(P,C)$ Relation for Sample Project

| | Best Case | | | | | Worst Case | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $Q(P,C)$ | $C_1$ | $C_2$ | $C_3$ | $C_4$ | $C_5$ | $C_6$ | $C_7$ | $C_8$ | $C_9$ | $C_{10}$ |
| | | | | | (a) max-min | | | | | |
| 1 | 0.07 | 0.07 | 0.07 | 0.07 | 0.07 | 0.40 | 0.58 | 0.58 | 0.58 | 0.58 |
| 2 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.45 | 0.45 | 0.45 | 0.45 | 0.45 |
| 3 | 0.07 | 0.07 | 0.07 | 0.07 | 0.07 | 0.22 | 0.22 | 0.22 | 0.22 | 0.22 |
| Sum/$\Sigma$ $wP$ | 0.18 | 0.18 | 0.18 | 0.18 | 0.18 | 0.49 | 0.57 | 0.57 | 0.57 | 0.57 |
| | | | | | (b) cum-min | | | | | |
| 1 | 0.13 | 0.13 | 0.13 | 0.13 | 0.13 | 1.08 | 1.26 | 1.26 | 1.26 | 1.26 |
| 2 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.81 | 1.01 | 1.06 | 1.06 | 1.01 |
| 3 | 0.11 | 0.11 | 0.11 | 0.11 | 0.11 | 0.40 | 0.40 | 0.40 | 0.40 | 0.40 |
| Sum/$\Sigma$ $wP$ | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 | 1.04 | 1.21 | 1.24 | 1.24 | 1.21 |



J. Constr. Eng. Manage., 2002, 128(6): 503-512

**Table 7.** Normalized Strengths of Each Cost Range Combining Worst-case and Best-case Scenarios for Sample Project

| $C_1$ | $C_2$ | $C_3$ | $C_4$ | $C_5$ | $C_6$ | $C_7$ | $C_8$ | $C_9$ | $C_{10}$ |
|---|---|---|---|---|---|---|---|---|---|
| (a) max-min composition operation | | | | | | | | | |
| 0.18 | 0.18 | 0.18 | 0.18 | 0.18 | 0.49 | 0.57 | 0.57 | 0.57 | 0.57 |
| 0.32[a] | 0.32[a] | 0.32[a] | 0.32[a] | 0.32[a] | 0.86[a] | 1.00[a] | 1.00[a] | 1.00[a] | 1.00[a] |
| (b) cum-min composition operation | | | | | | | | | |
| 0.30 | 0.30 | 0.30 | 0.30 | 0.30 | 1.04 | 1.21 | 1.24 | 1.24 | 1.21 |
| 0.24[a] | 0.24[a] | 0.24[a] | 0.24[a] | 0.24[a] | 0.84[a] | 0.98[a] | 1.00[a] | 1.00[a] | 0.98[a] |

[a]Normalized.



(for both the max-min and the cum-min composition operations), and normalize the values of the matrix. The values are normalized to prevent having recommendations greater than 1.0 and to ensure that at least one recommendation equals 1.0 (Table 7).

### *Step 7*

Apply the center of area method of defuzzification.

**Sample Calculation from max-min Composition Operation [Eq. (6)]**

$$C* = \sum m*C_m \Big/ \sum C_m = [(-35\%*0.32)+(-15\%*0.32)$$
$$+(-7.5\%*0.32)+(-2.5\%*0.32)+(0\%*0.32)$$
$$+(0\%*0.86)+(2.5\%*1.00)+(7.5\%*1.00)+(15\%*1.00)$$
$$+(60\%*1.00)]/(0.32+0.32+0.32+0.32+0.32+0.86$$
$$+1.00+1.00+1.00+1.00)=10.2\%$$

The output of this sample project predicts that the conditions of this project will cause a 10.2% cost overrun above the estimated contract fee, using the max-min composition operation, or an 11.6% cost overrun above the estimated contract fee using the cum-min composition operation.

## Conclusions and Future Development

The use of fuzzy logic to predict cost overruns/underruns yields a more realistic model in the design context. It enables the assessment of the project characteristics and risk events to be made subjectively, which is usually the case in practice. The output of the model is presented both numerically and linguistically, providing the decision maker with a useful and realistic guide to the likely cost overrun or underrun for the project. This feedback is useful to the user in setting an appropriate fee for the project based on its likely performance. The user may also choose to modify the existing conditions surrounding the project that he or she can control in order to change the likely outcome of the project. The modifications to the project characteristics and/or risk events can then be rerun through the model to assess the impact on cost due to changes in project conditions.

The model can also be used as a risk assessment tool. The user may decide that a project is too risky to pursue or find ways of transferring the risk back to the client or to other parties. By rating each of the factors, it also makes the user more aware of potential problems on the project that he or she should address prior to project commencement, or be prepared to handle later on in the project. For example, the exercise of assessing the level of scope definition will aid the user in further defining the scope of work if necessary, or suggest that the owner be approached for clarification on the project and the duties of the consultant.

There are several refinements and improvements that could be made to the model to improve its accuracy and usefulness. First of all, the accuracy of the model may be improved if it is further tailored to one specific type of project. This model is intended for use on commercial building projects, but it could be refined to suit either new buildings or renovations, for example. The size of the projects applicable to the model should be further defined, and possibly the engineering discipline (e.g., mechanical, electrical, structural, and so on). An alternative to predicting a percent cost overrun/underrun would be to predict an actual dollar value overrun or underrun.

The model would benefit from future research by expanding it to include more project characteristics and risk events, and by refining the standard strengths used in the model. A more elaborate survey may provide more accurate standard strengths that are based on real projects instead of a consensus of opinions. There are also other factors that could be taken into account, such as the bidding conditions, whether the owner of the project was public or private, and the project delivery method—all of which could be used to further classify projects.

A key factor in the accuracy of the model is the accuracy of the estimated contract fee that is used to calculate the actual percent cost overrun/underrun. Further research should examine methods of making the fee estimates more rigorous and accurate. Another solution would be to include a factor within the model that identifies that amount of time and preparation that was put into the estimate of the fee, the project manager's confidence in his or her estimate, or a rating of the project manager based on the accuracy of his or her previous estimates, to assess the likely accuracy of the estimated fee. The estimates from previous projects could be compared to the actual costs of the projects to assess the project manager's historical estimating accuracy.

One of the possible inaccuracies in the data used in the model is that the project ratings are not always evaluated on the same scale, because it is subjective rather than objective. Increased uniformity and objectivity of the project ratings could reduce this inaccuracy. Categories could be developed for the rating of each factor, and the user need only choose the appropriate category into which his or her project factors fall. For example, the experience of the consultant's project manager and the prime consultant's project lead could be categorized into the number of years of experience in the industry and the number of projects managed. The skill set of the design teams could be categorized by the number of junior and senior designers, the number of years of experience, and a rating of the designers' abilities (e.g., poor, average, or excellent).

The design industry could greatly benefit from future research in the area of controlling and managing project costs during the design phase. This research was a first attempt at addressing the effects of combinations of factors on design project costs. It has provided a realistic preliminary model for forecasting potential design cost variations. The results of this research have helped to identify some of the issues that need to be addressed in conducting future research in managing design projects.

## Acknowledgments

The writers would like to acknowledge the support of the participating design firm in conducting this research. This research was



J. Constr. Eng. Manage., 2002, 128(6): 503-512

partially funded through a Natural Sciences and Engineering Research Council of Canada Individual Research Grant held by Dr. Robinson Fayek.

## References

Bresnen, M. J., and Haslam, C. O. (1991). ‘‘Construction industry clients: A survey of their attributes and project management practices.’’ *Constr. Manage. Econom.*, 9, 327–342.

Bubshait, A. A., Farooq, G., Jannadi, M. O., and Assaf, S. A. (1999). ‘‘Quality practices in design organizations.’’ *Constr. Manage. Econom.*, 17, 799–809.

Construction Industry Institute (CII). (1986). ‘‘Scope definition and control.’’ *Publication 6-2*, Cost/Schedule Controls Task Force, Austin, Tex.

Construction Industry Institute (CII). (1996). ‘‘PDRI: Project definition rating index—Industrial projects.’’ *Implementation Resource, 113-2*, Front End Planning Research Team, Austin, Tex.

Construction Industry Institute (CII). (1999). ‘‘PDRI: Project definition rating index—Building projects.’’ *Implementation Resource, 155-2*, PDRI for Building Projects Research Team, Austin, Tex.

Cox, I. D., Morris, J. P., Rogerson, J. H., and Jared, G. E. (1999). ‘‘A quantitative study of post contract award design changes in construction.’’ *Constr. Manage. Econom.*, 17, 427–439.

Dumont, P. R., Gibson, Jr., G. E., and Fish, J. R. (1997). ‘‘Scope management using project definition rating index.’’ *J. Manage. Eng.*, 13(5), 54–60.

Dysert, L. R. (1997). ‘‘Scope development problems in estimating.’’ *Transactions, AACE Int.*, Morgantown, W. Va., 128–131.

Glavinich, T. E. (1995). ‘‘Improving constructability during design phase.’’ *J. Archit. Eng.*, 1(2), 73–76.

Klir, G. J., and Yuan, B. (1995). *Fuzzy sets and fuzzy logic: Theory and applications*, Prentice-Hall, Englewood Cliffs, N.J.

Kometa, S. T., Olomolaiye, P. O., and Harris, F. C. (1995). ‘‘Quantifying client-generated risk by project consultants.’’ *Constr. Manage. Econom.*, 13, 137–147.

Kometa, S. T., Olomolaiye, P. O., and Harris, F. C. (1996). ‘‘Validation of the model for evaluating client-generated risk by project consultants.’’ *Constr. Manage. Econom.*, 14, 131–145.

Love, P. E. D., Mandal, P., Smith, J., and Li, H. (2000). ‘‘Modelling the dynamics of design error induced rework in construction.’’ *Constr. Manage. Econom.*, 18, 567–574.

McGeorge, J. F. (1988). ‘‘Design productivity: Quality problem.’’ *J. Manage. Eng.*, 4(4), 350–362.

Nguyen, V. U. (1985). ‘‘Tender evaluation by fuzzy sets.’’ *J. Constr. Eng. Manage.*, 111(3), 231–243.





J. Constr. Eng. Manage., 2002, 128(6): 503-512