Al-Azhar University Civil Engineering Research Magazine (CERM)
Vol. (41) No. (3) July, 2019

Determination of the Contingency Values for the Construction of Water Treatment Plant Projects in Egypt

Madonna Nabil Roshdy[1], Ibrahim Abdel Rashid [2], Mohamed A. El Mikawi [3], Mohamed Tantawy[4]

[1]M.Sc. Student Department of Structural Engineering, Ain shams University, Cairo, Egypt
[2] Professor of Construction Management, Structural Engineering Department, Faculty of Engineering, Ain Shams University, Egypt
[3]Associate Professor of Construction Management, Structural Engineering Department, Faculty of Engineering, Ain Shams University, Egypt
[4] Assistant Professor of Construction Management, Civil Engineering Department, Faculty of Engineering, Helwan University

ملخص البحث:

تغيرت صناعة الإنشاءات كثيرا على مدى السنوات الماضية، حيث تواجه الشركات مزيد من المخاطر عن ذي قبل. نتيجة لذلك، شهدت جميع أنواع مشاريع البناء، بما في ذلك مشاريع محطات معالجة المياه، زيادات كبيرة في التكاليف. نظرًا لأن مشاريع محطات معالجة المياه تنطوي على عدد كبير من المخاطر، ولتجنب المخاطر المحتملة ولتقليل عواقب نتائج الأحداث غير الآمنة، صار تقدير احتياطي التكاليف المخصصة للطوارئ ضروري لضمان نجاح المشروع. إن احتياطي التكاليف المخصصة للطوارئ هو رصيد للتعامل مع تلك التهديدات. لذلك فإن تقدير التكاليف المخصصة للطوارئ وكفايتها النهائية ذات أهمية قصوى في مشروعات محطات معالجة المياه، لتقليل إحتمال حدوث تجاوز في التكلفة، أو على الأقل، لتقليل مقدار التجاوز نفسه في التكلفة. تحدد الورقة متغيرات المدخلات الفعالة التي تؤثر على مشاريع محطات معالجة المياه. تمت مراجعة الأدبيات وأجريت مقابلات مع خبراء ومسؤولين في محطات معالجة المياه لاستكشاف جميع المتغيرات التي تؤثرعلي احتياطي التكاليف المخصصة للطوارئ. تم تجميع البيانات لأكثر من 80 محطة لمعالجة المياه في مصرو بناء نموذج على أساس البيانات المتجمعة باستخدام أداة الذكاء الاصطناعي للتنبؤ باحتياطي الطوارئ لتشييد مشروعات محطات معالجة المياه في مصر. تم إجراء تحاليل مكملة لدراسة المتغير الأساسي في النموذج وتأثيره على قيم الطوارئ. ثم تم إجراء تحليل النتائج للتحقق من صحة النموذج وإظهار فعاليته حيث يمكن للمقاولين استخدام النموذج المشيد في المستقبل، قبل مرحلة التقدم بالعطاءات، لتحديد احتياطي المبالغ الطارئة لتشييد مشروعات محطة معالجة المياه المشابهة.

Abstract

Construction projects involve many uncertainties and risks in all phases. As a result, all types of construction projects, including Water Treatment Plant projects, have historically experienced significant cost increases. As Water Treatment Plant Projects involve large number of risks, project contingencies are important items in the cost estimate for compensating unforeseen risks. Contingency serves as well for project costs underestimations and budget overruns. The contingency cost of a Water Treatment project is a fundamental input for decision making process set by contractor during the tendering stage. The paper identifies the effective input variables affecting the Water Treatment Plant projects. Literature has been reviewed and interviews were conducted with experts and officials in Water Treatment Plants to explore all variables that influence the contingency cost. Datasets that consist of more than 80 Water Treatment Plant Projects in Egypt were collected. The effective variable is used to construct an Artificial Neural Network (ANN) Model, using MATLAB Neural Network Toolbox, to predict the contractor's contingency. Analysis of results was performed to validate the model and demonstrate its effectiveness. The constructed model can be simulated to assess contractors in contingency estimation for new data.

Keywords: Artificial Neural Network, Backpropagation Neural Networks, Coefficient of Correlation, Contingency, Deterministic methods, Probabilistic methods, Processing elements and Neurons.

1. Introduction

The issue of setting appropriate contingency is one that often poses difficulties for most Water Treatment Plant Projects as a vital component of a project budget is "Cost Contingency". Cost contingency for projects has been part of the project management process for more than 50 years. Regardless of the omnipresence of project cost contingency in the theory of project cost management, there has been little empirical research into project practitioners understanding of the concept, and its methods of estimating or management. A major impact on the cost of a project is the accuracy of an estimated price (Baccarini, 2005).

The purpose of cost contingency therefore is to generate a reserve fund that is adequate to cover the ingrained risks within the project total budget and within the project completion duration. Its establishment eliminates the adverse impact of unforeseen event.

One of the key success criterion for project owners is the cost performance for all the different types of the Construction Projects. Cost contingency is usually included within the estimated budget that will be presented to the project sponsor. Contractors usually never accurately calculate the cost contingency; they subjectively estimate it as 5–10% from the total cost estimated; their basis for estimation depends solely on their experience in past similar projects. However, this method does not have a sound basis and is difficult to justify or defend. Therefore, cost contingency estimation is of a great importance to projects.

This paper is structured as follows; Section 2 briefly reviews the literature on ANN to predict the contingency of construction projects. Section 3 presents a detailed overview of the methodology. Section 4 discuses variables identification, implementation and development of the ANN as well as the model's results and analysis. Finally, the paper concluded remarks and results are discussed in Section 5.

2. Literature Review

Traditionally, contractors determine cost contingency simply by adding; say 10% contingency onto the estimated cost of a project (Mills, 2001). However, this conventional method is arbitrary and difficult to justify or defend (Baccarini D. , 2006).Several methods to determine cost contingency have been introduced, such as Monte Carlo Simulation (Al-Bahar, 1988), Artificial Neural Network (ANN) (Chen, 2000), Belief Network (Khalafallah, 2002)and linear Regression (Sonmez R. E., 2007). However, most of these methods are based on statistical analyses which rely on historical data or sometimes require statistical and mathematical knowledge from the user. This becomes one of its limitations since most of the construction's project managers or construction personnel do not have any knowledge of the formalized technique to estimate cost contingency (Smith G. R., 1999).

According to (Dikmen, 2007) and (Kangari R. &., 1989) in the construction field, often the use of expert knowledge, experience, intuitive judgment and rule of thumbs that are usually ill-defined and vague cannot be avoided due to limited statistical data. The impreciseness and vagueness are usually characterized by using linguistic terms such as low, medium and high. Fuzzy set is a mathematical tool that can accommodate the use of such linguistic terms which is based on subjective judgment (Kasabov, 1996). Previous researchers have presented the use of fuzzy sets in estimating project cost contingency (Paek, 1993); (Tah, 1993). However, these methods still have a limitation, particularly the difficulty in its application (Sonmez R. E., 2007). Fuzzy expert system, which is developed based on fuzzy logic concept, is a technique that provides an easy method in dealing with the fuzzy set. Fuzzy logic could accommodate the human approximate reasoning, which is usually represented using IF-THEN rule. This method is easy to be developed, understood and applied (Kasabov, 1996) and has been widely used in many areas such as control, decision making, and management (Negnevitsky, 2004).

Research Methodology

This section presents the research approach and methodology that was followed in this study. The methodology proposed in this research is depicted in Fig. 1. To achieve the above-mentioned study objectives, the following procedure was carried out:

1. Explore initial list of variables that deem important and influence contingency cost estimation of water treatment plant construction projects through, literature review and experts' interviews. Also, determine the form of the output variables as the application objectives of the artificial neural network (ANN) model.
2. Collect Project Database for Contingency.
3. Use Artificial Neural Networks (ANN) through MATLAB, Neural Network Toolbox to Predict Contingency for Water Treatment plants.
4. Evaluate and Validate the Developed ANN Model with additional project contingency data.
5. Future simulation for the model to predict the contractor contingency for contractors.

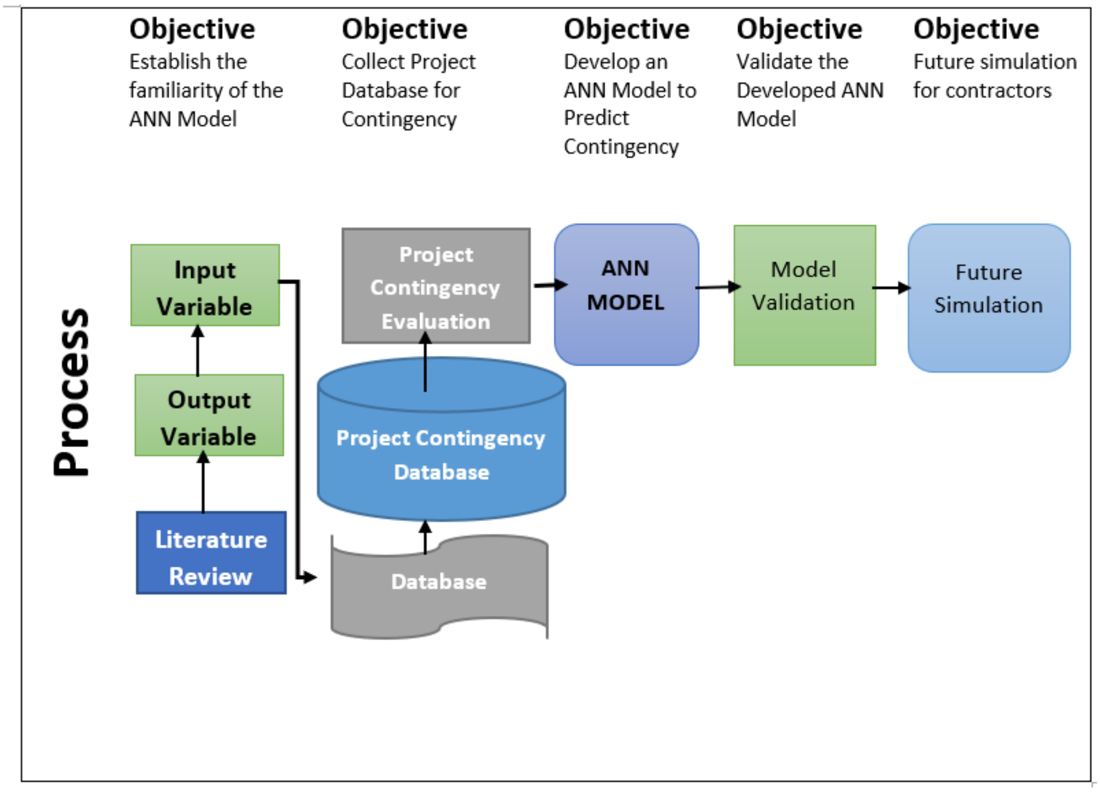

Figure 1: Flowchart for Proposed Methodology

3. Artificial Neural Network Model

The main modeling tool for this study was to use artificial neural network (ANN) technique which is an exciting form of Artificial Intelligence (AI). ANN is inspired by the way the biological nervous systems, such as the brain works - neural networks learn by example. An ANN is formed from single units, (artificial neurons or processing elements - PE), connected with coefficients (weights), which constitute the neural structure and are organized in layers. The power of neural computations comes from connecting neurons in a network. Each PE has weighted inputs, transfer function and one output. The behavior of a neural network is determined by the transfer functions of its neurons, by the learning rule, and by the architecture itself. The weights are the adjustable parameters. The weighed sum of the inputs constitutes the activation of the neuron.

Before constructing the ANN model, it was important to finish the data gathering process and define the input and output variables for the Model. The data gathering process included several interviews with sector heads at the National Organization for Potable Water and Sewage Drainage (NOPAWSD). From previous literature and based on those interviews, the most significant input parameter affecting the contingency cost of Water Treatment Plants was the plant's design capacity and the target parameter for the model was the contingency percentage. The neural network model was created in MATLAB. The ANN model consisted of three types of samples listed hereinbelow;

Training Sample: These are presented to the network during training, and the network is adjusted according to its error. The number of Water Treatment plants for the training sample is 48 samples which accounts to 60% of the total gathered data.

Validation Sample: These are used to measure network generalization, and to halt training when generalization stops improving. This data set is used to minimize overfitting. It is not used for adjusting the weights of the network, it is just for verifying that any increase in accuracy over the training data set yields an increase in accuracy over a data set that has not been shown to the network before, or at least the network hasn't trained on it (i.e. validation data set). If the accuracy over the training data set increases, but the accuracy over the validation data set stays the same or decreases, then you're overfitting your neural network and you should stop training. The number of Water Treatment plants for the Validation sample is 12 samples which accounts to 20% of the total gathered data.

Testing Sample: This data set is used only for testing the final solution in order to confirm the actual predictive power of the network. These has no effect on training and so provide an independent measure of the network performance during and after training. The number of Water Treatment plants for the testing sample is 12 samples which accounts to 20% of the total gathered data.

The Network properties are as follows:

- Network input: Plant capacity (m^3/day)
- Network output: Contingency percentage
- Network type: Feed-Forward Back-Propagation.
- Training function: TRAINLM.
- Adaption learning function: LEARNGDM.
- Performance function: MSE.
- Number of hidden layers: 2

Validation for the ANN Model

For evaluations and validations of neural network models, the Mean Squared Error (MSE) error metric is used. The definition for MSE is given in Equation 1.

$$MSE = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m}(Oij - Tij)^2}{n} \qquad \text{Equation 4-1}$$

Where:
n: The number of patterns in the test set
m: The number of components in the output vector
O: The output of a single neuron j
T: The target for the single neuron j, and i is each input pattern

Figure 2 presents the performance curve of the developed model. The performance curve shows the plot of the training errors, validation errors, and test errors. This is used to validate the network performance. The best validation performance is 0.096865 and occurred at epoch 134. The results of the model are very good due to the following;

- The final mean-square error is small.
- The test set error and the validation set error have similar characteristics.
- No significant overfitting has occurred by iteration 134 (where the best validation performance occurs).

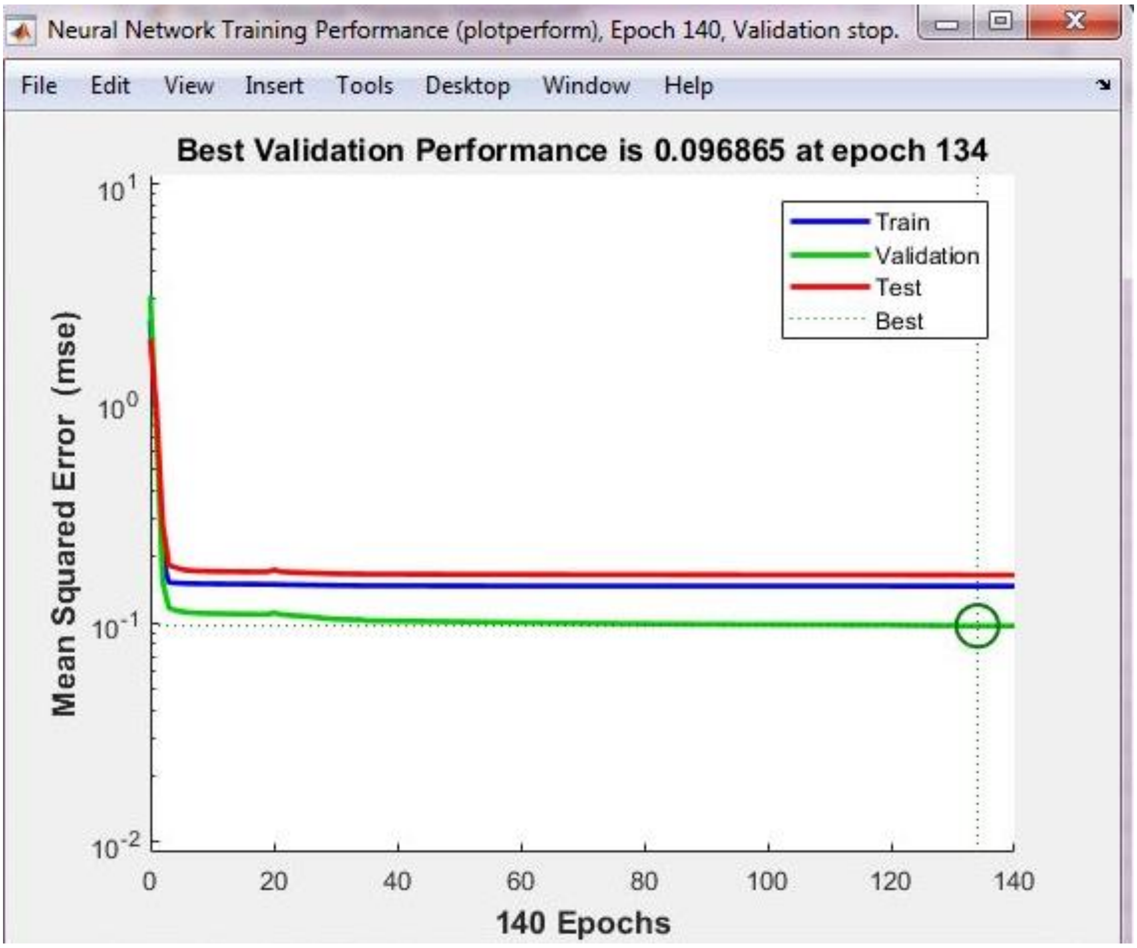

Figure 2: Performance Curve- Mean square Error vs. Epochs

In addition to the performance curve, the Regression plots are used to validate the network performance. The regression plots presented in Figure 3 plots the linear regression of targets relative to outputs. It displays the network outputs with respect to targets for training, validation, and test sets, as well as a plot for the three sets combined.

The plot shows how much the estimated output is deviating from the actual target. The ideal case is both output and target being exactly equal where the data should fall along the 45-degree dotted line. In other words, if the networks had learned all training patterns perfectly, then all points would fall on the 45° line shown. This represents a perfect fit where the network outputs are equal to the targets. In general, the target and output are never the same; hence, the output will always deviate from the target. This deviation is shown by blue, green, red and black lines. The regression plots in this study indicates a strong correlation between the independent variable and the dependent variable as the fit is almost perfect for all data sets, with R values in each case of 0.95 or above.

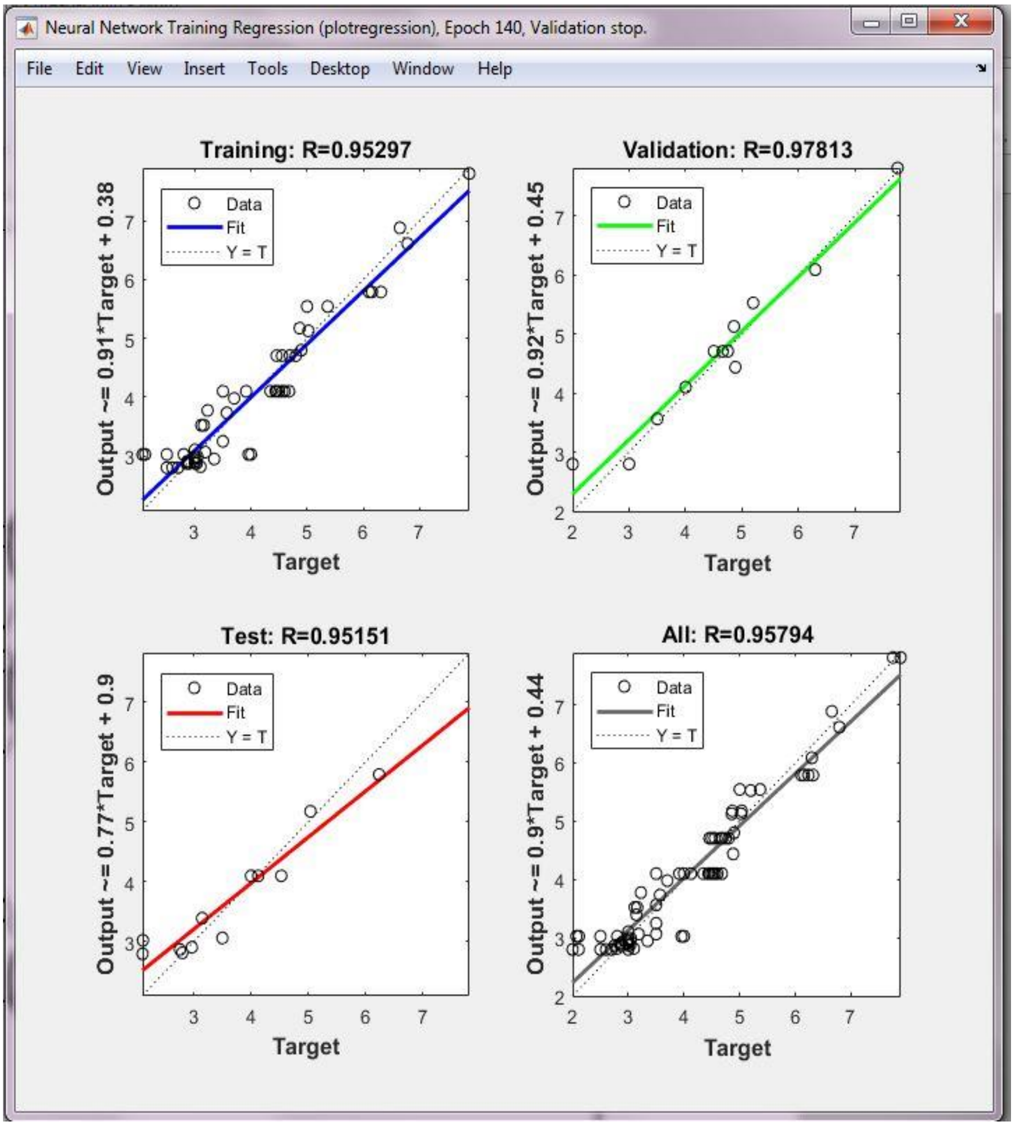

Figure 3: Regression Plots for training, validation, test sets and the three sets combined

4. Conclusion

Traditionally, contingencies are often calculated as an across-the-board percentage addition on the base estimate, typically derived from intuition, past experience and historical data. The literature review highlighted several serious flaws with this estimating method. This judgmental and arbitrary method of contingency calculation is difficult for the estimator to justify or defend. A percentage addition results in a single-figure prediction of estimated cost, which implies a degree of certainty that is simply not justified. It does not encourage creativity in estimating practice, promoting a routine and mundane administrative approach requiring little investigation and decision making.

This paper briefly presented a robust approach to estimating project cost contingency for Water Treatment Plant Projects in Egypt. A model was developed using ANN that can predict the contingency cost of Water Treatment Plants for the Tendering stage. The first objective of this research was to gather data for the contingency values for the construction of Water Treatment Plant projects in Egypt. The data gathering process was such a tedious task due to the confidentiality of such data and the unavailability of online databases by the government. At first, many input factors were included in the submitted form to professionals. Several interviews were conducted with sector heads at the National Organization for Potable Water and Sewage Drainage (NOPAWSD) and the most significant input parameter was employed for this study namely the plant's design capacity.

Data for 80 Water Treatment Plants was collected and randomly divided into three sets: 60% for training, 20% for validating the performance and 20% as a completely independent test of network generalization. One input parameter, namely the Contingency percentage was employed. The input was entered in the ANN architecture and simulated in MATLAB by using the Neural Network Toolbox, to predict the contractor's contingency. The feedforward backpropagation technique was used to generate the contingency percentage. The ANN architecture consisted of one input layer, two hidden layers and one output layer. The resulting ANN model reasonably predicted the contingency percentage with favorable training and testing phase outcomes. For evaluations and validations of neural network models, the Mean Squared Error (MSE) error metric was used. In the future, the model can be further simulated to assess contractors in contingency estimation for novice data and provide them with more realistic results than the mere deterministic method. It is worth mentioning that ANN uses non-linear mathematics and therefore can be used to model highly complex and non-linear functions.

References

1. Al-Bahar, J. F. (1988). Risk management in construction project: A systematic analytical approach for contractors. *Civil Engineering Department, Doctor of Philosophy: University of California at Berkeley*.

2. Baccarini, D. (2005). Estimating project cost contingency - beyond the 10% syndrome. *Australian Institute of Project Management National Conference.* Australia.

3. Baccarini, D. (2006). The maturing concept of estimating project cost contingency. Western Australia, Australia: Department of Construction Management, Curtin University of Technology.

4. Chen, D. a. (2000). A neural network approach to risk assessment and contingency allocation. *AACE Transactions*, Risk.07.01-6.

5. Dikmen, I. B. (2007). Using fuzzy risk assessment to rate cost overrun risk in international construction project. *International Journal of Project Management, 25*, 494–505.

6. Kangari, R. &. (1989). Construction risk assessment by linguistics. *IEEE Transaction on Engineering Management, 36*, 126–130.

7. Kasabov, N. K. (1996). *Foundation of neural networks, fuzzy system and knowledge engineering.* MIT Press.

8. Khalafallah, A. M.-S. (2002). Estimating residential project cost contingencies using belief network. Cairo University.

9. Mills, A. (2001). A systematic approach to risk management for construction. *Structural Survey, 19*, 245–252.

10. Negnevitsky, M. (2004). *Artificial intelligent:, A guide to intelligent system (2nd ed.).* London, UK: Addison-Wesley.

11. Paek, J. H. (1993). Pricing construction risk: Fuzzy set application. *Journal of Construction Engineering and Management, 119*, 743–756.

12. Smith, G. R. (1999). Small to medium contractor contingency and assumption of risk. *Journal of Construction Engineering and Management, 125*, 101–108.

13. Sonmez, R. E. (2007). Quantitative methodology for determining of cost contingency in international project. *Journal of Management in Engineering, 23*, 35–39.

14. Tah, J. H. (1993). Contractor project risk contingency allocation using linguistic approximation. *Computing Systems in Engineering, 4*, 281–293.