

## Estimation of cost overrun in construction projects using Fuzzy Logic

Neelakash Haloi<sup>1</sup>, Tapas Goyal<sup>2</sup>, Faizan Zahoor<sup>3</sup>, Harsh Jain<sup>4</sup>, Ravikiran S Wali<sup>5</sup>

Student, Department of Civil Engineering<sup>1,2,3,4</sup> R.V College of Engineering, Bangalore, India

Assistant Professor, Department of Civil Engineering<sup>5</sup> R.V College of Engineering, Bangalore,  
India

[haloineelakash@gmail.com](mailto:haloineelakash@gmail.com)<sup>1</sup>, [tapasgoyal.777@gmail.com](mailto:tapasgoyal.777@gmail.com)<sup>2</sup>, [faizanzahoorkanroo@gmail.com](mailto:faizanzahoorkanroo@gmail.com)<sup>3</sup>,  
[hjain6415@gmail.com](mailto:hjain6415@gmail.com)<sup>4</sup>, [ravikiransw@rvce.edu.in](mailto:ravikiransw@rvce.edu.in)<sup>5</sup>

**Abstract:** In recent years, the impact of the construction industry on the economy of India is increasing. However, there are a number of problems that are being faced by the construction companies or contractors or other relevant stakeholders, in the completion of a construction project. One of the major problems being faced is time and cost overrun. Thus, it is important to find ways of mitigating this critical problem. One of the solutions that is being used is the prediction of the amount and time that a project might overrun beforehand i.e. estimating the overrun during the design phase itself. This prediction can be done using a number of different techniques, the important one of which used in this paper is through the use of Fuzzy Logic. It also provides a brief description about the factors that are responsible for causing this cost overrun. The factors are identified by conducting a questionnaire survey and gathering responses from the relevant stakeholders which include contractors, project engineers, architects and consultants. The responses collected are further processed and analysed using SPSS software. The top 5 factors have been selected after the analysis and further used in preparation of the fuzzy logic model in the fuzzy toolbox of MATLAB. Further, the validation of this model has been done with real-time data of projects from various published research papers. Thus, the model developed provided prediction of the percentage of the cost overrun based on the percentage input of the top five factors.

**Keywords:** Cost overrun, fuzzy logic, relative importance.

### 1. INTRODUCTION

The construction industry plays an important role in the economic development of a country, contributing about 3175.14 INR Billion to the GDP of India in the first quarter of 2021 as per the MSPI. The completion of projects on time within estimated budget is important.

The practical implementation of any construction model is preceded by a phase wherein the calculation of the costs is done. However, at the time of the implementation of this model, the actual costs exceed the estimated costs. This phenomenon is referred to as cost overrun. According to the report by MSPI, 437 infrastructure projects have been hit by cost overruns of over Rs 4.37 lakh crore in 2020 in India. Thus, the problem of cost overrun needs to be mitigated to reduce losses and increase project management efficiency.

A number of factors influencing construction time and cost overrun have been identified using various statistical and analytical tools. One of such tools that can be used is Fuzzy Logic. It has the potential to provide legit results when there are a lot of uncertainties involved, as in the case of the construction industry. Fuzzy theory is based on uncertainties where there is an inherent impreciseness and it provides mathematical tools to deal with imprecise and vague data. It enables the assessment of the project

characteristics and risk events collectively and hence helps to analyze the effects of costs involved in construction projects.

Further section of the paper is divided as follows: Section 2 studies in detail the methodology adopted in this paper; section 3 briefs about the stepwise analysis for the development of the model; section 4 briefs about the validation of the developed model; section 5 demonstrates the results obtained; section 6 gives the suggestive framework to mitigate the cost overrun; section 7 provides a conclusion of the paper; section 8 provides an overview of future scope.

## **2. Research Methodology**

The brief methodology that was implemented in this research paper on cost overrun using fuzzy logic has been described as:

- Project characteristics were identified due to which the cost overrun occurs with the help of literature review.
- Questionnaire prepared was distributed to relevant stakeholders of the industry to understand the severity of all considered factors.
- Data filtration and validation using statistical approaches (RII and Cronbach's Alpha).
- Incorporating Fuzzy logic for developing cost overrun assessment model using fuzzy toolbox in MATLAB.
- Defining membership functions and constructing a rule base to control the output variables.
- Performing fuzzy inference to obtain the desired results.

### **2.1 Identification of factors**

Based on the literature referred, 45 factors were determined to have an impact on cost overrun in construction projects. These factors were categorised into 6 broad groups, which are:

- 1) Time related factors include delay in decision making process, delay in site handover etc.
- 2) Human Resource and Communication related factors consist of lack of managerial capability of the company, lack of managerial capability of the company etc.
- 3) Design related factors include inadequate project planning and scheduling, lack of technical expertise.
- 4) Equipment related factors consist of frequent equipment breakdown, inadequate modern equipment etc.
- 5) Labour related factors include labour injuries at site, personal conflicts between labourers.
- 6) Contractor related factors rise due to problems like lack of experience, poor communication between project parties etc.

### **2.2 Questionnaire Design**

A questionnaire survey comprising 44 questions was prepared. The questionnaire consisted mainly of 3 parts.

- First part was focused on obtaining the personal information of the respondents viz. (name ,age, profession, experience in the construction industry.)
- Second Part aimed to get the opinions of the factors affecting cost overrun wherein the respondents were asked to mark their level of agreement to each factor (Strongly agree-5, Somewhat agree- 4, Neutral -3, Somewhat disagree- 2, Strongly disagree-1).
- Third part was focused on getting the respondent's perspective on the ways that can be implemented to avoid or prevent cost overrun.

## 2.3 Questionnaire Analysis

The questionnaires were distributed randomly among 150 personnel working in the construction industry across India and 70 sets of questionnaires were reverted successfully from which 88.2 % respondents agreed to facing cost overrun in construction projects and 8.8 % opposed cost overrun. The response rate observed is as follows:

1. 67.6% are Project Engineers.
2. 11.8% are contractors.
3. 17.6% are consultants.
4. 2.9% are Architects.

## 2.4 Data Analysis Methods

### 2.4.1 Reliability Analysis

Reliability analysis was done to check the internal consistency of the data having multiple scales. The statistical package IBM SPSS 22 was used for the test. The most popular and common technique to measure the scale of reliability is Cronbach's alpha ( $\alpha$ ), Field [21]. Cronbach Alpha is a coefficient of reliability (or consistency) that can be written as a function of the number of test items and the average inter-correlation among the items. Given below is the formula for Cronbach alpha .

$$\alpha = \frac{N.c}{v + (N-1)c} \quad (1)$$

where  $\alpha$  is Cronbach alpha, N is number of items, c is average inter-item covariance among the items, and v is average variance. Reliability of the data is considered as high level when Cronbach's alpha is 0.7 which implies that data is suitable for further analysis[4]. The computed value of Cronbach's alpha was 0.879 which was greater than 0.7. Hence the data obtained was reliable and used for further analysis.

### 2.4.2 Ranking Analysis

Relative importance index can be calculated as per the following formula -

$$RII = \frac{\sum_{i=1}^5 W_i \cdot X_i}{A \times N} \quad (2)$$

where RII = Relative importance index, w = weighting given to each factor by respondents and it ranges from 1 - 5. x = frequency of response given for each cause. A= highest weight that is 5 in our case, N= total no. of responses. RII was calculated for all the 29 isolated factors and the top 5 factors have been tabulated in descending order of RII in the table given below.

**Table 1. Top 5 Factors in Descending Order in Accordance with RII**

Rank	Factor
1	Inflation in price of raw materials
2	Additional work not included in the contract

3	Poor management and supervision
4	Delay in decision making process
5	Inexperienced technical staff deployed at site

## 2.5 Correlation Test

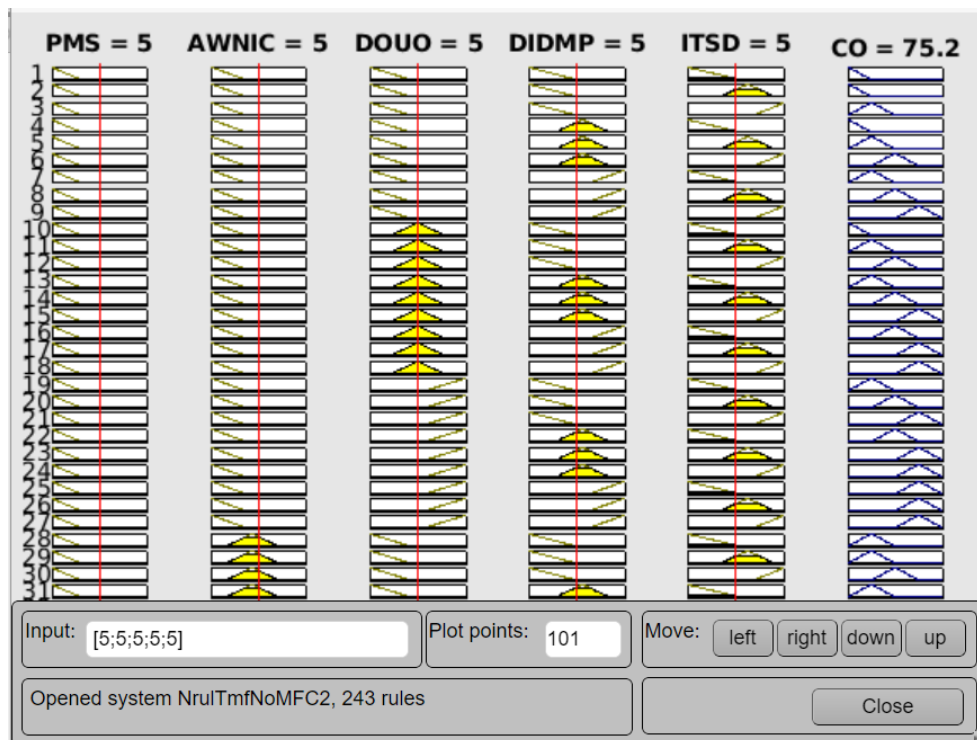
Spearman's correlation test has been performed to examine the relationship between the factors having an impact on the cost overrun. The purpose of this test is to examine and evaluate the relationships involving the factors contributing to cost overrun.

It was observed that factors delay in material availability and delay in works carried out by sub contractors were highly correlated with a correlation coefficient value of 0.624 having a significance factor of 0.002. On the other hand, delay in decision making process and additional work not included in the contract was observed to be moderately correlated with a correlation coefficient value of 0.443 and a significance factor of 0.044.

## 3. Fuzzy Inference System Modelling

A cost overrun estimation model was developed considering the top five factors using the Fuzzy Logic Designer tool that is available in MATLAB. The analysis steps for the development of the model are:

- The first step of the fuzzy inference system is defining the input and output variables for the model development. In the FIS editor of MATLAB, a five-input and one-output system was developed. Input variables were chosen as the top-5 most important factors obtained via relative importance index.
- In the second step, the input variables were fuzzified with the help of membership functions. Membership functions for each input and output variable were aligned in the Triangular form of function wherein each input variable were given 3 linguistic divisions - Low, Medium and High while the output variable consisted of 5 linguistic divisions - Very Low, Low, Medium, High, Very High. Ranges of these divisions were defined according to the findings from literature review and suggestions from experienced personnel in the industry.
- In the third step fuzzy rules that connect input variables to output variables were defined to perform the inference. In this study, 243 rules have been constructed in the form of IF- then.
  - The relative percentage based on the Relative Importance Index (RII) of cost overrun was found and assigned as weightage to the fuzzy rules to develop the assessment model.
  - Since the obtained RI values for the top 5 factors were found to be less distinct, the range of input and output linguistic variables were defined in the light of expert's opinion. This method has been useful in achieving more accurate fuzzy rule weights resulting in each IF-Then rule to differ indicating the relative importance of fuzzy rules in the model making.
- In the fourth step, the model performed defuzzification of the combined fuzzy output to generate crisp output value. The percentage of cost overrun was determined in the interval of [0 100]. The entire roadmap of the process can be seen in the rule viewer editor given below in Figure(1).



**Figure 1. Fuzzy Rule Viewer**

Thus, the model forecasts cost overruns due to the worst activity of the top five factors. The results have shown that the increase in expected project cost (cost overrun) in percentage may be calculated for different percentage losses of each of the input variables, i.e. elements responsible for cost overrun.

#### 4. Model Validation

Model validation was carried out using quantitative data available in the published papers (Gaytri Thampi; 2016 and Shreenath.A et al; 2015) as focus groups and survey interviews could not be adopted to verify the model due to data confidentiality issues. The papers included the detailed case study analysis of 5 building projects and have validated the survey findings on the cost affecting factors.

**Table 2. Details of Projects considered for Validation**

Project Details	Type of Project	Location	Estimated Cost	Final cost
Project 1	Residential Building	Kollam	2,44,83,000	2,67,75,358
Project 2	Residential Building	Kollam	2,73,52,000	2,98,4,308
Project 3	Residential Building	Chennai	8,50,00,000	10,28,50,000
Project 4	Residential Building	Chennai	1,00,00,000	1,21,50,000
Project 5	Residential Building	Kollam	2,24,88,000	2,48,68,800

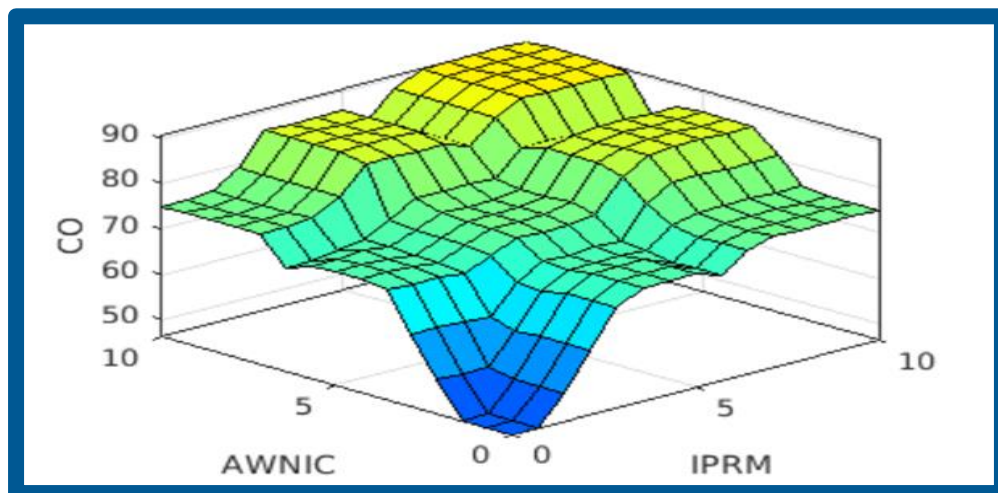
## 5. Results

- Results obtained from implementing the developed model are tabulated in Table(3).

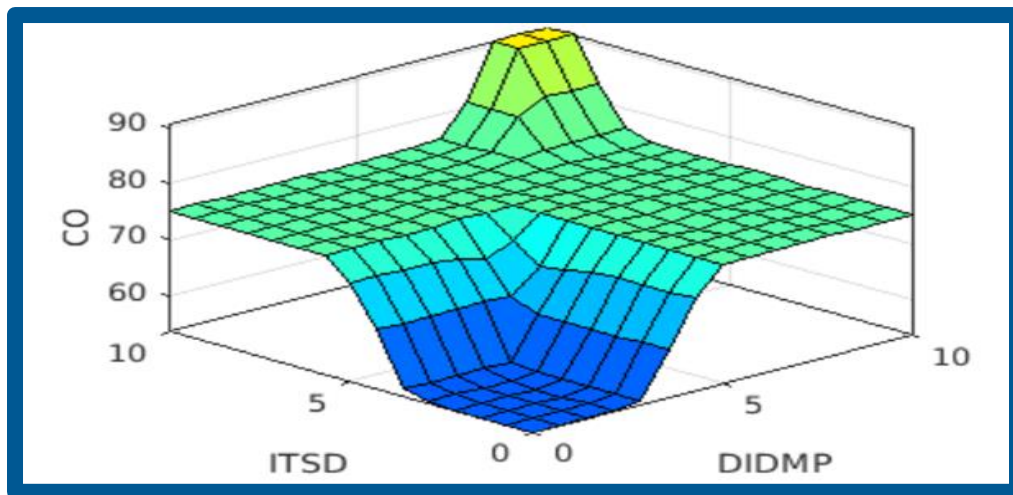
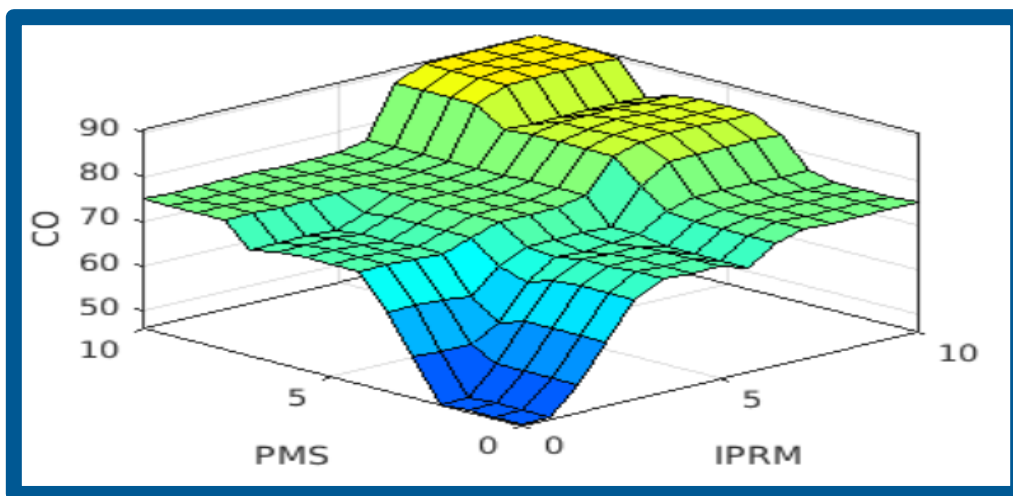
**Table 3. Cost Overrun for Different Percentage Loss of Factors**

Factors considered in the study	Escalation of material prices (% loss)	Additional work not included in the contract (%loss)	Poor management and Supervision (% loss)	Delay in decision making process (% loss)	Inexperienced technical staff deployed at site (% loss)	Cost overrun estimate actual (in the paper) %	Cost overrun estimate %	%error= $\{(actual - estimate)/actual\} * 100$
1.	25	15	15	0	10	24.8	24.5	1.20
2	35	20	15	0	10	32.9	30.9	10.03
3	22	10	15	12	5	21	23.7	12.85
4	20	10	10	10	5	21.5	22.9	6.51
5	26	6	10	10	10	26	24.7	5

- Considering the combinations of the different percentage loss of the factors, the percentage error between cost overrun and actual cost overrun was calculated as 7.11% average by the model. Hence the proposed model holds good for predicting the percentage of cost overrun for the given factors.
- Variation in result of cost overrun among the factors considered for analysis with respect to the Fuzzy rules can be understood by surface diagrams presented below. Each diagram considers involvement of 2 factors in contribution towards cost overrun. Surface diagrams show the result area which is analogous with each set of input variables. The x-axis and y-axis represent the input variables while the z-axis represents cost overrun. Surface diagrams for top three factors:



**Figure 2. Surface Diagram for Additional Work Not Included in Contract and**

**Inflation in Price of Raw Materials****Figure 3. Surface Diagram for Inexperienced Technical Staff Deployed at Site and Delay in Decision Making Process****Figure 4. Surface Diagram for Poor Management and Supervision and Inflation in Price of Raw Materials**

## 6. Suggestive Framework

The respondents have voluntarily provided possible solutions to abate the impact of cost overrun on construction projects.

1. More efforts need to be made in procuring locally available raw materials.
2. Social liabilities included in the project during the construction phase must be tackled skilfully.
3. Multilateral authorities must be avoided to mitigate the delay in the decision making process.
4. Transparency between clients and contractors should exist.
5. Avoiding rework by ensuring the work is rightly done despite a small delay.
6. Frequent interaction with clients is required to completely understand the perception of the client.

## 7. Conclusion

Cost overrun is a frequently occurring major issue in construction projects as agreed by 90% of the respondents and to avoid cost overrun there is need to develop a cost overrun/prediction model as a decision-support tool for project managers and cost estimators prior to the bidding stage. Only few projects get completed within stipulated cost and time. Almost all the estimates of the construction projects are calculated by adding contingency charges which leads to ignoring the analysis of the parameters that leads to cost overrun.

This project focuses on constructing a cost overrun forecasting model by identifying, classifying and ranking of the factors responsible for the cost overrun in construction projects. The main causes for cost overrun wherein 45 factors were identified and grouped into six broad categories in the light of literature review and expert suggestions. The correlation between the several factors affecting cost of construction projects was established using the Spearman correlation test. A fuzzy logic model was developed to predict the percentage of the cost overrun based on the percentage input of the top five factors. A suggestive framework for the five significant factors was also provided in this paper.

## 8. Future Scope

As established, cost overrun is a major problem faced by the construction industry, it involves vast aspects to be covered to obtain a true estimation on the topic. Fuzzy modelling involved in this project is based upon expert's opinions and the understanding gathered from literature review which implies a range of variations in this model can be seen based on the perspective of other experts. In addition, Fuzzy Logic provides a rough idea towards any project. Different approaches can be used to develop models for cost overrun estimation which might offer better accuracy. This project even involved a small sample size which brings in the possibility of error. The model may vary according to the structure and quantity of members in the fuzzy logic tool. Moreover, the focus remained in building projects for analysis, other aspects of the construction industry like roads and bridges could even be incorporated. Quantitative approach can also be used to get marginally better results. More to that, in the field of construction management research, hybrid fuzzy approaches such as the neurofuzzy and fluid neural networks have been increasingly used and may also be more wide-spread since some construction issues are better solved, which fuzzy setting/fuzzy logic alone may not match well. In modelling patterns and automated learning, for example, neural networks are robust and pattern identification are powerful. Their combination can assist in certain ways to build models with minimised uncertainty.

## REFERENCES

### Journal Articles

- [1] Issa, U. H., Mosaad, S. A. A., & Salah Hassan, M., "Evaluation and selection of construction projects based on risk analysis". Structures, ISSN: 2352-0124, Vol: 27, Page: 361-370 (2020), doi:10.1016/j.istruc.2020.05.049.
- [2] Edyta Plebankiewicz, Wieczorek, D., "Prediction of Cost Overrun Risk in Construction Projects". Sustainability 2020, 12(22), 9341, <https://doi.org/10.3390/su12229341>.
- [3] Elmousalami, H. H." Intelligent methodology for project conceptual cost prediction". Heliyon, ISSN: 2405-8440, Vol: 5, Issue: 5, Page: e01625, (2019), doi:10.1016/j.heliyon.2019.e01625.
- [4] Gaurav Dobariya, Swapnil P. Wanjari, "Identifying factors causing cost overrun of the construction projects in India", Sādhanā volume 41, pages 679–693,(2016), <https://doi.org/10.1007/s12046-016-0498-3>.
- [5] Larsen, J. K., Shen, G. Q., Lindhard, S. M., & Brunoe, T. D., "Factors Affecting Schedule Delay, Cost Overrun, and Quality Level in Public Construction Projects". Journal of Management in Engineering Vol. 32, Issue 1, (2016), doi:10.1061/(asce)me.1943-5479.0000391.
- [6] Love P. E. D., Wang X., Sing C., & Tiong, R. L. K., "Determining the Probability of Project Cost Overruns". Journal of Construction Engineering and Management, 139(3), 321–330, (2013), doi:10.1061/(asce)co.1943-7862.0000575.
- [7] Kumar, V. S. S., & Reddy, G. C. S., "Fuzzy Logic Approach to Forecast Project Duration in Construction Projects". Construction Research Congress 2005. doi:10.1061/40754(183)136.
- [8] Touran, A., "Probabilistic Model for Cost Contingency". Journal of Construction Engineering and Management, 129(3), 280–284, (2003), doi:10.1061/(asce)0733-9364(2003)129:3(280).
- [9] Knight, K., & Robinson Fayek, A., "Use of Fuzzy Logic for Predicting Design Cost Overruns on Building Projects". Journal of Construction Engineering and Management, 128(6), 503–512, (2002), doi:10.1061/(asce)0733-9364(2002)128:6(503).



- [10] Gul Polat, Ferzan Okay, Ekin Eray, "Factors affecting cost overrun in micro-scaled construction companies". *Procedia Engineering* 85, 428-435,(2014) doi: 10.1016/j.proeng.2014.10.569.
- [11] Albert P. C. Chan, Daniel W. M. Chan, M. and John F. Y. Yeung., "Overview of the Application of "Fuzzy Techniques" in Construction Management Research". *JOURNAL OF CONSTRUCTION ENGINEERING AND MANAGEMENT*,(2009), DOI: 10.1061/ASCECO.1943-7862.0000099.
- [12] Ji Hyun Jung, Dae and Young Kim and Hak Ki Lee., "The Computer-based Contingency Estimation Through Analysis Cost Overrun Risk of Public Construction Project". *KSCE Journal of Civil Engineering* volume 20, pages 1119–1130,(2016), DOI 10.1007/s12205-015-0184-8.
- [13] O. Dehzangi, M. J. Zolghadri, S. Taheri, S.M. Fakhrahmad., "Efficient fuzzy rule generation: A new approach using data mining principles and rule weighting". *Fourth International Conference on Fuzzy Systems and Knowledge Discovery*,(2007), DOI:10.1109/FSKD.2007.267.
- [14] Edyta Plebankiewicz, Damian Wiczorek., "Adaptation of a Cost Overrun Risk Prediction Model to the Type of Construction Facility". *Symmetry* 2020, 12, 1739. doi:10.3390/sym12101739.
- [15] Edyta Plebankiewicz., "Model of Predicting Cost Overrun in Construction Projects". *MDPI Sustainability* 2018, 10, 4387, doi:10.3390/su10124387.
- [16] Hai Pham, Truong-Van Luu, Soo-Yong Kim, Duc-Thinh Vien, "Assessing the Impact of Cost Overrun Causes in Transmission Lines Construction Projects". *KSCE Journal of Civil Engineering* (2020) 24(4):1029-1036, DOI 10.1007/s12205-020-1391-5.
- [17] Haytham H. Elmousalami, Comparison of Artificial Intelligence Techniques for Project Conceptual Cost Prediction". *IEEE Transactions on Engineering Management* ( Volume: 68, Issue: 1, Feb. 2021), DOI: 10.1109/TEM.2020.2972078.
- [18] Arazi Idrus, Muhd Fadhil Nuruddin, M. Arif Rohman, "Development of project cost contingency estimation model using risk analysis and fuzzy expert system". *Expert Systems with Applications* 38 (2011) 1501–1508, doi:10.1016/j.eswa.2010.07.061.
- [19] Burcu Aklnci, Martin Flischer, "Factors Affecting Contractors' Risk Of Cost Overburden". *Journal of Management in Engineering* Vol. 14, Issue 1 (January 1998), [https://doi.org/10.1061/\(ASCE\)0742-597X\(1998\)14:1\(67\)](https://doi.org/10.1061/(ASCE)0742-597X(1998)14:1(67)).
- [20] Jong-Ho Ock, Seung-Heon Han, "Measuring Risk-associated Activity's Duration: A Fuzzy Set Theory Application". *KSCE Journal of Civil Engineering* (2010) 14(5):663-671, DOI 10.1007/s12205-010-1003-x.
- [21] A. Field, *Discovering Statistics Using SPSS*, Sage, London (2005).