


## Research Article

# An Effective Approach to Input Variable Selection for Preliminary Cost Estimation of Construction Projects

Meseret Getnet Meharie <sup>1,2</sup> Zachary C. Abiero Gariy,<sup>3</sup>  
Raphael Ngumbau Ndisya Mutuku,<sup>4</sup> and Wubshet Jekale Mengesha<sup>5</sup>

<sup>1</sup>Department of Civil Engineering, Pan African University Institute for Science, Technology and Innovation (PAUISTI), P.O. Box 62000 00200 JKUAT—Juja Campus AICAD—Block C, Room 101 Nairobi, Juja, Kenya

<sup>2</sup>School of Civil Engineering and Architecture, Department of Civil Engineering, Adama Science and Technology University, P.O. Box 1888 Adama, Oromia Region, Adama, Ethiopia

<sup>3</sup>School of Civil, Environmental and Geospatial Engineering, Department of Civil, Construction and Environmental Engineering, Jomo Kenyatta University of Agriculture and Technology (JKUAT), P.O. Box 62000—00200, Nairobi, Juja, Kenya

<sup>4</sup>Faculty of Engineering & Technology, Department of Building and Civil Engineering, Technical University of Mombasa, Tudor, Tom Mboya Street, P.O. Box 90420—80100, Mombasa, Kenya

<sup>5</sup>Construction Technology and Management Competency Unit and Construction Law Chair, Addis Ababa University (Ethiopian Institute of Architecture Building Construction and City Development (EIABC)), P.O. Box 7048, Addis Ababa, Ethiopia

Correspondence should be addressed to Meseret Getnet Meharie; mesget86@gmail.com

Received 18 February 2019; Accepted 12 May 2019; Published 25 June 2019

Academic Editor: Heap-Yih (John) Chong

Copyright © 2019 Meseret Getnet Meharie et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Accurate cost estimates are vital to the effective realisation of construction projects. Extended knowledge, wide-ranging information, substantial expertise, and continuous improvement are required to attain accurate cost estimation. Cost estimation at the preliminary phase of the project is always a challenge as only limited information is available. Hence, rational selection of input variables for preliminary cost estimation could be imperative. A systematic input variable selection approach for preliminary estimating using an integrated methodology of factor analysis and fuzzy AHP is presented in this paper. First, the factor analysis is used to classify and reduce the input variables and their variable coefficients are determined. Second, fuzzy AHP based on the geometric mean method is employed to determine the weights of input variables in a fuzzy environment where the subjectivity and vagueness are handled with natural language expressions parameterized by triangular fuzzy numbers. Then, the input variables are suggested to be selected starting with those having high coefficient and high importance weight. A set of three variables, one from each group, can be added to the estimating model at a time so that the problem of collinearity can vanish and good accuracy of the estimate can be ensured. The proposed approach enables cost estimators to better understand the complete input variable selection process at the early stage of project development and provide a more accurate, rational, and systematic decision support tool.

## 1. Introduction

Construction project success can be attained through the good performance of indicators of the project such as time, cost, quality, and customer satisfaction [1]. In construction industry, one of the critical factors of a feasibility study in the early planning and design phase of a project is the early

understanding of the cost of the construction project [2–4]. Several studies have also stated that the failure or success of a certain construction project largely depends on the accuracy of various estimations prepared throughout the project lifecycle [2, 3, 5–9]. However, establishing a more realistic and accurate cost estimate for a certain project at its preliminary phase is one of the most difficult tasks. Because

preparation of preliminary cost estimation (PCE) must be done before the detail design of the project is completed [3] as there exist lack of detail information, high complexity, and uncertainty [10]. Therefore, it is required to have a systematic way of exploring the cost-estimate accuracy factors which can be used for preliminary cost estimation of construction projects. In this regard, there are several studies that explored the techniques for identification of input cost factors for the PCE model in highway construction industry [7, 11–15]. These literature studies identified quantitative and qualitative cost data which can be classified as numerical or categorical data, respectively. However, there are some limitations with existing techniques or methods that affect the quality of cost estimation.

The first deficiency of the published paper is the absence of dealing with the correlation between input variables (cost factors) that will result in multicollinearity problems as multicollinearity is an issue for developing a model [16]. Multicollinearity among cost data will give the same effect on the estimation outcomes because of trade-offs and repetitive errors [17]. Second, during the factor selection process, in most cases, the large number of input variables was considered for estimation purpose, which requires high computational requirements. This also makes the developed cost model more complex and time-consuming as far as the formulation of the historical database is concerned. Finally, there are few fuzzy logic approaches, in the literature, aimed at evaluating the relative importance and influence of cost factors though the use of fuzzy set theory is the hot research topic. The expert evaluation of qualitative cost attributes is always subjective and imprecise and vague [18]. In such a situation, fuzzy set theory is a suitable and powerful tool in dealing with the uncertain environment with vagueness, imprecision, and ambiguity. Thus, this study integrated a factor analysis and a fuzzy analytical hierarchy process (fuzzy AHP) approach and provides a rational and systematic input variable selection process for cost estimators to estimate the more realistic preliminary cost of highway construction projects while filling the knowledge gaps mentioned above.

The rest of this paper is organized as follows. Section 2 provides the cost-estimation accuracy factors or input variables and their selection process. Section 3 presents how the integrated methodology of factor analysis and fuzzy AHP can be adopted. Section 4 shows numerical analysis and results of factor analysis and fuzzy AHP and presents how the input variables can be selected. Section 5 presents discussions, contributions, and managerial implications. Finally, general conclusions and remarks are then drawn in Section 6.

## 2. Literature Review

Preliminary or conceptual cost estimation is commonly used to predict the cost at an earlier stage in the project development [19]. It is a greatly experience-based process and involves the assessment of several multifaceted relationships of cost-influencing factors [20]. The first step in estimating the preliminary cost of the project is selection of the appropriate

input factors, and it is vital for attaining good cost-estimation performance or improving the prediction capability of the model [12, 21, 22]. Accordingly, in the realm of estimation, many studies have applied various factor selection techniques, including both statistical and nonstatistical, in order to identify and select the most significant cost factors that are required for estimating the preliminary cost of projects in the highway construction industry [12, 13, 21, 23–25]. In the following paragraphs, the most cost-influencing factors and the way they are identified and selected for preliminary or top-down cost-estimation purpose are described.

Swei et al. [15] presented an alternative approach for preliminary cost estimates through the combination of a maximum likelihood estimator to search for an optimal data transformation combined with least angle regression for input variable selection/dimensional reduction. Gardner et al. [23] quantified the effort expended to carry out conceptual estimates using highway agency data and concluded that input variables require a low amount of effort and that have a high influence on the final predicted cost are desired in data-driven conceptual cost-estimating models for the highway agency studied. Elfaki et al. [26] conducted an extensive review of literature and finally explored the most significant cost-estimation factors such as type of project, type of client, ground conditions, material costs, size of the project, likely design and scope changes, duration, contract type, tendering method, and estimator-specific factors as a benchmark to compare between the cost-estimation proposals. The input data set that contains 12 years cost of different construction items were analysed through graph plotting, which shows the relationship between the year and cost for the purpose of forecasting the future cost of construction projects [19]. In the study conducted by Ma et al. [27], the proficiency of estimators on specifications for construction cost estimation was mentioned as a significant factor that affects the accuracy and efficiency of cost estimation. Hyari et al. [28] identified the factors that could influence engineering services' bids for construction projects. They used five factors including project type, engineering services category, project location, construction costs, and project scope for the development of conceptual cost-estimation model for engineering services.

Kim and Shin [22] identified the types of critical factors required for estimating the cost of a building project using mean value and standard deviation after conducting a thorough literature review and expert interview. Yu and Skibniewski [29] utilised foundation type, structural type, floor area, number of base floors, and number of total floors as an influential attribute in estimating the cost of residential building construction projects in China. Arabzadeh et al. [30] considered variables consisting of thickness, tank diameter, and length of the weld to forecast the construction cost of spherical storage tank projects. Bayram et al. [31] used the approximate cost, total construction area, number of floors, building height, and a contract value as an input parameter for estimate the cost of building construction projects. Shutian et al. [24] identified 10 parameters through the Pearson correlation coefficient for predicting the cost of construction projects.

Zhu et al. [25] obtained the key market drivers which affect cost items based on their priorities through parallel Monte Carlo simulation and Likert scale analysis for the purpose of estimating the cost of the chemical engineering construction project. This study identified four key factors, which include technology and specification standard, vender list, target market and economy, and the condition of the project site. Agyekum—Mensah [32] examined the degree of accuracy and factors that influence the uncertainty of cost estimates in the construction sector. This author identified the main determinant factors for project cost estimates including terms and conditions of payments, type of client, experience, availability of cost information, a guarantee of the job, and repeated work. Zhang and Minchim [33] selected factors such as number of bids, contract days, location, and contract type as predictor variables to develop cost-estimation models for resurfacing projects. Pesko et al. [34] analysed the major group of works in urban road construction and stated that the works on roadway construction and landscaping are cost-significant items according to “Pareto” distribution, i.e., distribution of 20/80. In their study, 12 inputs (work items) were identified for the purpose of estimating the cost of urban road construction.

Gransberg et al. [21] identified 13 input variables on the bases of their influence on cost and effort required to collect data for the purpose of conceptual cost estimation of highway projects. These input variables include project location (rural or urban), site topography, project size, design AADT, typical section, design speed(s), intersection signalization and signage, traffic control, curb and gutter and sidewalk, contract time, letting date, geotechnical-subsurface and slope recommendations, and extent of utility relocations and costs.

### 3. Methodology

The overall framework and steps of the newly proposed input variable selection approach for PCE based on the integration of factor analysis and fuzzy AHP are shown in Figure 1.

**3.1. Factor Analysis Technique.** Factor analysis is a statistical technique for identifying groups and clusters of variables that can be used to characterize the relationships among sets of many interrelated variables [35, 36]. This technique has three key uses: (i) to understand the structure of a set of variables; (ii) to construct a questionnaire to measure an underlying variable; and (iii) to reduce a data set to a more manageable size without much loss of the original information [13, 17, 35]. The present study applied factors analysis to combine, categorize, and reduce the collected factors through a questionnaire survey. To apply the factor analysis based on principal component analysis, the following main steps were carried out. The SPSS version 24 was employed to perform the analysis: (i) testing for the appropriateness of using factor analysis; (ii) extraction of principal components or factors with eigenvalues more than

1, known as Kaiser’s criteria [35]; and (iii) factor rotation, principal components rotate about the original variable’s axis. In doing so, varimax with the Kaiser normalization method of orthogonal rotation is applied to preserve the principal components due to its simplicity in the interpretation of the factors [37] and also a new transformation matrix is formed. This method minimizes the number of variables that have high loadings on each factor. The method of maximum variance, which is most commonly used, causes multiple rotations, and with each rotation, a new variable or factor is created. The newly created factors or principal components are uncorrelated with each other. Thus, by the use of factor analysis based on principal component analysis, the dimensionality of the data is reduced and the multicollinearity is eliminated [17]. So far, very few studies have been carried out on the application of factor analysis (and principal components analysis) to eliminate multicollinearity and select the appropriate input variables (factors) for the purpose of preliminary cost estimation at the early stages of project development in the highway construction industry.

**3.2. Fuzzy Sets and Fuzzy Numbers.** The theory of fuzzy set was first introduced by Zadeh [38] to tackle the vagueness and uncertainty during the decision-making process in practice. The concept of fuzzy set theory is an extension of the classical notion of set, and he defined it as a “class of objects with a continuum of grades of membership” [38, 39].

*Definition 1.* A fuzzy set in  $\mathbb{R}$  is a function  $\mu: \mathbb{R} \rightarrow [0, 1]$ , with the following properties [39, 40]:

- (1)  $x$  is convex, i.e.,  $\mu(x) \geq \min\{\mu(s), \mu(r)\}$ , for  $s \leq x \leq r$ ;
- (2)  $x$  is normal, i.e.,  $(\exists)t_0 \in \mathbb{R}: \mu(x_0) = 1$ ;
- (3)  $x$  is upper semicontinuous, i.e.,

$$\begin{aligned} &(\forall)t \in \mathbb{R}, (\forall)\alpha \in (0, 1) : \mu(x) < \alpha, \\ &(\exists)\delta > \text{such that } |s - t| < \delta \Rightarrow \mu(s) < \alpha. \end{aligned} \quad (1)$$

*Remark 1.* A fuzzy number  $\tilde{A}$  is a triangular fuzzy number if its membership function  $\mu_{\tilde{A}}: \mathbb{R} \rightarrow [0, 1]$  is equal to following [41]:

$$\mu_{\tilde{A}}(x) = \begin{cases} \frac{x-l}{m-l}, & l \leq x \leq m, \\ \frac{u-x}{u-m}, & m \leq x \leq u, \\ 0, & \text{otherwise,} \end{cases} \quad (2)$$

where the parameters  $l$  and  $u$  represent the lower and upper bounds of fuzzy number  $\tilde{A}$ , and  $m$  is the modal value for  $\tilde{A}$ . The triangular fuzzy numbers can be symbolised by  $\tilde{A} = (l, m, u)$  (see Figure 2).

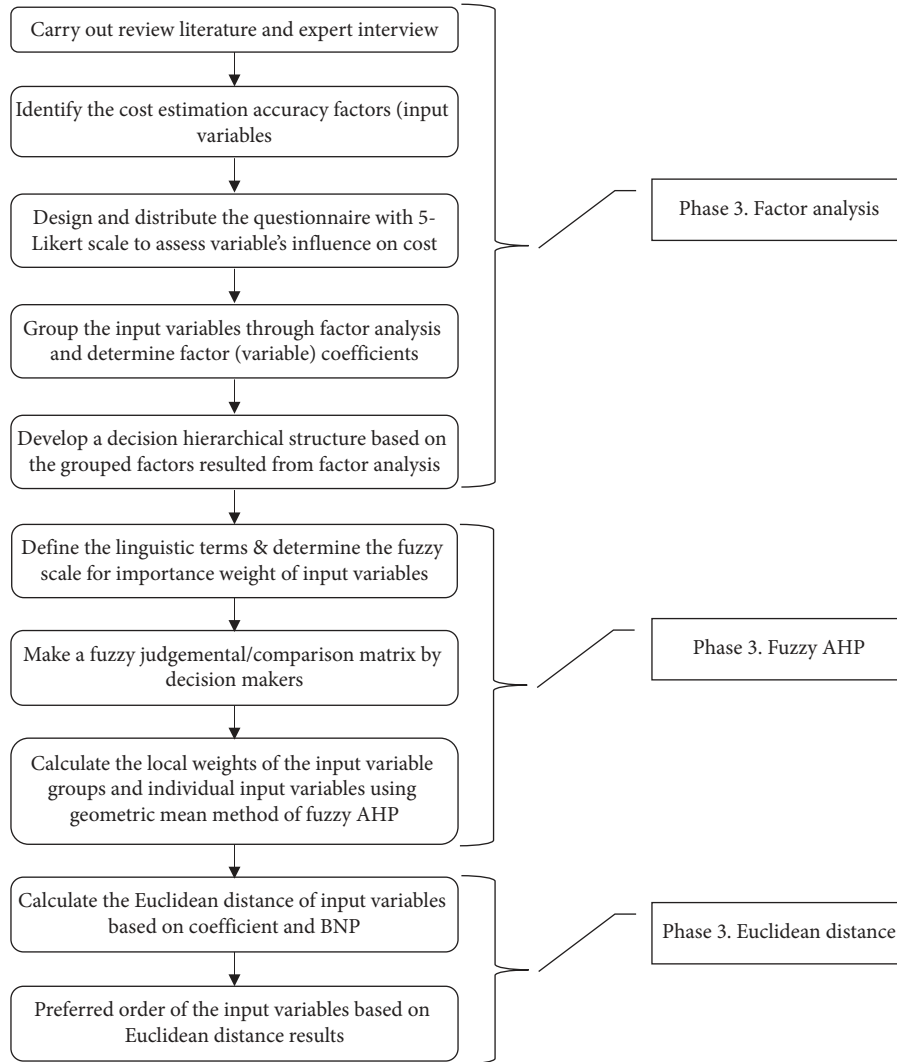


FIGURE 1: The steps for the proposed input variable selection approach.

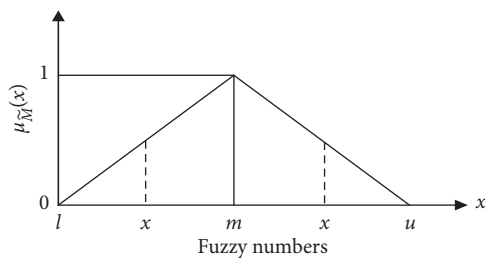


FIGURE 2: The membership functions of the triangular fuzzy numbers.

**3.3. Fuzzy AHP Method.** A fuzzy AHP method is selected as a factor-weighting model because studies have declared its effectiveness in prioritizing and ranking of factors, criteria, alternatives in construction management, and other various disciplines [42–50]. Fuzzy AHP can be applied in several ways. In this study, Buckley’s geometric mean method is used because it is easy to extend to the fuzzy case, of its

computational easiness, and guarantees a unique solution [41, 51].

**3.3.1. Geometric Mean Analysis on Fuzzy AHP.** Buckley [51] introduced the geometric mean method to extend the hierarchical analysis to the situation of using linguistic variables. The geometric mean method is applied to compute the fuzzy weights for each fuzzy comparison matrix.

The steps for the fuzzy AHP analysis based on the geometric mean method are summarized as follows [40, 41, 51].

Step 1. Construct a fuzzy pairwise comparison matrix,  $\tilde{A} = [\tilde{a}_{ij}]$  (see equation (3)) The fuzzy pairwise comparison matrices among all the input variables in the variable groups of the hierarchy system and each decision maker allot linguistic variable represented by triangular fuzzy numbers to the pairwise comparison among all input variables.

$$\tilde{A} = \begin{bmatrix} 1, 1, 1 & \tilde{a}_{12} & \cdots & \tilde{a}_{1n} \\ \tilde{a}_{21} & 1, 1, 1 & \cdots & \tilde{a}_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{a}_{n1} & \tilde{a}_{n2} & \cdots & 1, 1, 1 \end{bmatrix} \quad (3)$$

$$= \begin{bmatrix} 1, 1, 1 & \tilde{a}_{12} & \cdots & \tilde{a}_{1n} \\ \frac{1, 1, 1}{\tilde{a}_{12}} & 1, 1, 1 & \cdots & \tilde{a}_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1, 1, 1}{\tilde{a}_{1n}} & \frac{1, 1, 1}{\tilde{a}_{2n}} & \cdots & 1, 1, 1 \end{bmatrix}$$

The judgment comparison matrix  $\tilde{A}$  is an  $n \times n$  fuzzy matrix containing fuzzy numbers  $\tilde{a}_{ij}$ .

where  $\tilde{a}_{ij} \otimes \tilde{a}_{ji} \approx 1$  and  $\tilde{a}_{ij} \cong w_i/w_j, i, j = 1, 2, \dots, n$ . Meaning that  $\tilde{a}_{ij}$  is the importance of input variable  $V_i$  with respect to subfactor  $V_j$ .

Step 2. Compute the fuzzy geometric mean value. The fuzzy geometric mean value  $\tilde{r}_i$ , for each subfactor  $i$  is computed as follows:

$$\tilde{r}_i = [\tilde{a}_{i1} \otimes \tilde{a}_{i2} \otimes \cdots \otimes \tilde{a}_{in}]^{1/n} \quad (4)$$

Step 3. Calculate the fuzzy weight. The fuzzy weight  $\tilde{w}_i$  for each subfactor  $i$  is calculated as follows:

$$\tilde{w}_i = \tilde{r}_i \otimes (\tilde{r}_1 \oplus \tilde{r}_2 \oplus \cdots \oplus \tilde{r}_n)^{-1} \quad (5)$$

where  $\tilde{r}_k = (l_k, m_k, u_k)$  and  $(\tilde{r}_k)^{-1} = (1/u_k, 1/m_k, 1/l_k)$ .

Step 4. Defuzzify the fuzzy weights. The fuzzy weights  $\tilde{w}_i = (l_i, m_i, u_i)$  can be defuzzified by any defuzzified method. In this study, the Center of Area (CoA) method [52, 53] is applied to compute the best non-fuzzy performance (BNP) value of the fuzzy weights of each input variable and it is calculated as follows:

$$\tilde{w} = [\tilde{w}_1, \tilde{w}_2, \dots, \tilde{w}_n], \quad (6)$$

$$\text{BNP}_{wi} = \frac{[(u_{wi} - l_{wi}) + (m_{wi} - l_{wi})]}{3} + L_{wi}$$

3.3.2. *Fuzzy Comparison Matrix for Input Variables.* So as to take the vagueness of evaluation on the pairwise comparison of input variables into consideration, the triangular fuzzy numbers (1~9) are utilised to characterize the evaluation from “equally important to absolute important.” Table 1 depicts the linguistic terms and triangular fuzzy numbers.

3.4. *Input Variable Selection Approach.* In order to meet dual objectives during PCE, it was proposed that input variables be opted starting with those that have a high coefficient, obtained from factor analysis, but also have a high

importance weight (or BNP value), obtained from fuzzy AHP. Accordingly, the input variables suggested to be preferred and opted are located in the top right-hand quadrant as shown in Figure 3.

#### 4. Case Study: Numerical Analysis

4.1. *Initial Identification of Input Variables.* At this stage, the factors affecting the accuracy of the cost estimate, input variables, are identified and defined through intensive literature review and expert interview for the required numerical analysis. The study on various literature studies relevant to the cost estimation of highway construction project explored the most influential cost-estimation accuracy factors or input variables [7, 12, 23, 25, 33, 54–56]. It is not necessarily true that increasing the number of input variables in an early estimate may seem to improve the accuracy of the estimate [23]. Moreover, increasing the number of factors results in several questions when making pairwise judgmental or comparison matrices and lowers the efficiency of the analytical hierarchy process (AHP) model [57]; hence, using interviews with highway engineers along with review of literature studies, the final set of 12 input variables were identified that have an influence on the cost of highway construction projects in the case of Ethiopian Highway Construction Industry; these are shown in Table 2. To select these input variables, variable’s significant influence on cost based on Engineer’s opinion, variable’s frequency in the literature, and variable whose information is available at an earlier stage are considered as the criteria.

4.2. *Factor Analysis Results.* Based on the determined factors, the first questionnaire was developed to collect data for factor analysis. The survey was distributed to 105 highway professionals and academicians, who have had sufficient work experience in the highway construction industry in Ethiopia, via e-mail and physical presence. Respondents of the survey were asked to answer “how influential/impact do you perceive this factor/variable is on the construction cost of a project?” on each of the 12 variables identified under a five-point Likert scale, i.e., 5—very high, 4—high, 3—moderate, 2—low, and 1—not at all. A total of 70 responses were received with three incomplete responses. Results from the survey indicated that the majority of respondents (87%) who responded for the survey were highway engineers and practitioners from various construction organizations and Ethiopian Road Authority who had had significant participation in the planning, designing, and implementing of various types of highway construction projects, and the rest 13% of respondents were academicians from higher institutions.

To test the reliability of the data set, Cronbach’s  $\alpha$  coefficient method was used. Cronbach’s  $\alpha$  coefficient for the data set is more than 0.7, is a recommended threshold value [35] i.e. 0.716. Therefore, the data set is considered reliable. In addition, the Kaiser–Meyer–Olkin (KMO) and Bartlett’s test were conducted as a preliminary analysis to check the suitability the data set collected through the questionnaire

TABLE 1: Linguistic terms and triangular fuzzy numbers.

Linguistic terms	Fuzzy number	Triangular fuzzy numbers
Equal important (EI)	$\bar{1}$	(1, 1, 1)
Equal to moderate important (intermediate) (E-MI)	$\bar{2}$	(1, 2, 3)
Moderate important (MI)	$\bar{3}$	(2, 3, 4)
Moderate to strong important (intermediate) (M-SI)	$\bar{4}$	(3, 4, 5)
Strong important (SI)	$\bar{5}$	(4, 5, 6)
Strong to very strong important (intermediate) (S-VSI)	$\bar{6}$	(5, 6, 7)
Very strong important (VSI)	$\bar{7}$	(6, 7, 8)
Very strong to absolute important (intermediate) (VS-AI)	$\bar{8}$	(7, 8, 9)
Absolute important (AI)	$\bar{9}$	(8, 9, 10)

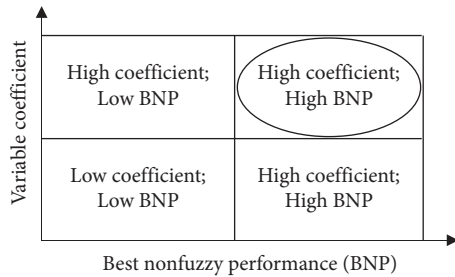


FIGURE 3: Input variable selection to meet the dual objectives of PCE.

TABLE 2: Identified input variables.

Input variables	References
Project type	[12, 54]
Project complexity	[7, 10]
Project location	[10–12, 21, 28, 33, 54, 58]
Project scope	[7, 11, 12, 26, 54]
Project size	[8, 11, 12, 21, 26, 54]
Site topography	[12, 21, 25, 56, 59]
Bridge type	[21, 23]
Number of bridges	[12, 21, 23]
Existence of groundwater	[12]
Soil type	[12, 26, 56, 60]
Inflation rate	[11, 12, 54]
Project duration	[7, 10, 12, 21, 26]

surveys for running the factor analysis [13]. Consequently, the KMO measure verified the sampling adequacy for the analysis,  $KMO = 0.603$ , which is well above the acceptable limit of 0.5. According to Field [35], Bartlett's test of sphericity should be significant (the value of Sig. should be less than 0.05). For these data, Bartlett's test is highly significant since the significance ( $p$ ) is 0.001, indicating that there exist some relationships between the variables. Therefore, the results of KMO and Bartlett's test revealed that factor analysis is appropriate. An initial analysis was run to find eigenvalues for each variable group in the data set. Three variable groups or principal components have had

eigenvalues 1, which is Kaiser's criterion and all together explained 58.568% of the variance.

Factor coefficients are then computed after orthogonal rotation (varimax) to determine the contribution of each input variable with respect to the variable group or principal component it belongs to. In doing so, variable group loadings less than  $\pm 0.40$  should be removed due to their insignificance for variable group interpretation [37]. In other words, input variables within a variable group with loading values above  $\pm 0.40$  are considered as substantive values and involved in the calculations of the variable coefficients. The variable loadings are used to compute the variable coefficients. These coefficients are computed by dividing a loading value of each input variable in the variable group by the sum of the variable group loadings of all input variables within the same variable group [37]. It can be noticed that an input variable with the largest variable loading values in a certain variable group would have the greatest effect on the variable group value. Table 3 depicts the three variable groups and the related input variable coefficients. These coefficients are then considered for as one element for the input variable selection process.

**4.3. Fuzzy AHP Results.** The variable groups and input variables resulting from the factor analysis (see Table 3) are converted to the hierarchical structure to transform these clustered input variables as the schematic structure depicted (see Figure 4). The hierarchical structure of the decision problem is shown in Figure 4, and its ultimate goal is to select the input variables for preliminary cost estimation based on the relative importance weight of the input variables, which is placed in the first level in the hierarchical structure based on the pairwise comparison. The three variable groups and twelve input variables are located in the second and third levels, respectively.

At this phase, to define and compute the importance weights of variable groups and individual input variables under each variable group, the fuzzy AHP method based on geometric mean analysis was applied. Succeeding the hierarchical structure, the experts or decision makers are required to fill the evaluation matrix. The overall

TABLE 3: Variable groups and coefficients.

Variable group	Eigenvalue	Variance (%)	Input variables	Variable coefficients
Variable group ( $V_1$ )	4.0420	25.8079	Project size	0.1525
			Project type	0.1499
			Project complexity	0.1417
			Project scope	0.1274
			Project duration	0.1253
			Project location	0.1044
			Bridge type*	0.0923
Variable group ( $V_2$ )	1.8707	17.5530	Existence of groundwater	0.3191
			Bridge type	0.2804
			Number of bridges	0.2436
			Soil type*	0.1570
Variable group ( $V_3$ )	1.1155	15.2073	Project location*	0.1739
			Inflation rate	0.3409
			Site topography	0.2657
			Soil type	0.2195

\*Input variables which are clustered more than one variable group.

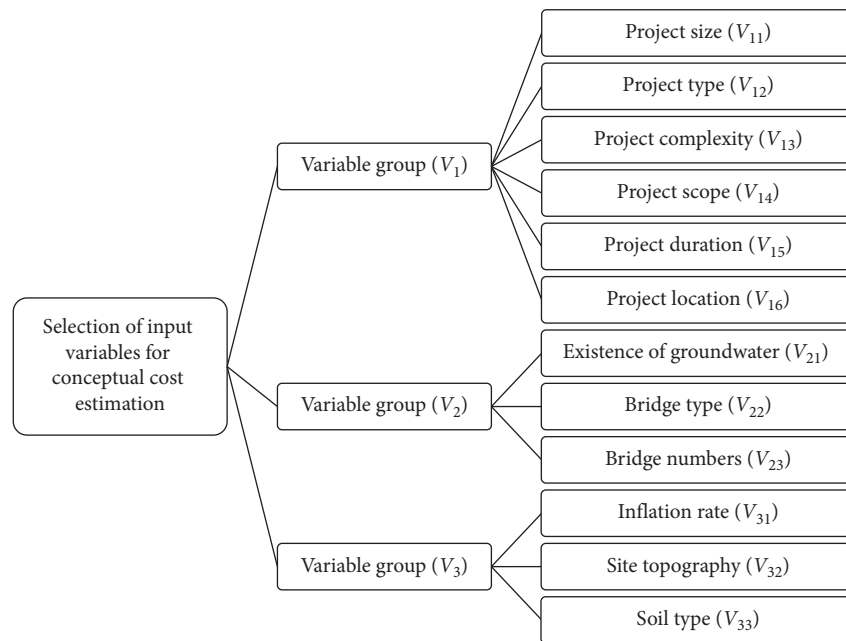


FIGURE 4: A hierarchical structure for input variable importance weight determination.

computational procedures to determine the weights of variable groups placed at the second level and input variables located at the third level in the hierarchy are demonstrated as follows:

- (a) First, the decision-making committee consisted of thirteen highway experts use the triangular fuzzy numbers tabulated in Table 1 to compare the importance or preference of variable group or input variable over another with the help of the questionnaire. Linguistic preferences of the decision

makers or experts were transformed into triangular fuzzy numbers.

- (b) Fuzzy-integrated pairwise comparison weights for the variable group in the hierarchy were computed through combining the collected data from all experts by using a geometric mean method, that is,  $\tilde{a}_{ij} = (\tilde{a}_{ij}^1 \otimes \tilde{a}_{ij}^2 \otimes \dots \otimes \tilde{a}_{ij}^{13})$ .

As a sample calculation, the fuzzy-integrated pairwise comparison value for the variable group ( $V$ ) are shown in the following matrix  $\tilde{A}$ .

		$V_1$	$V_2$	$V_3$
$\tilde{A} =$	$V_1$	(1, 1, 1)	(3.2169, 4.0179, 5.0104)	(3.5137, 4.6915, 5.7851)
	$V_2$	(0.1996, 0.2489, 0.3109)	(1, 1, 1)	(1.0173, 1.4336, 2.0138)
	$V_3$	(0.1729, 0.2132, 0.2846)	(0.4966, 0.6976, 0.9830)	(1, 1, 1)

(c) Calculating the consistency index and consistency ratio. Consistencies of the pairwise comparison results were measured using the approach adopted by Chan et al. [61] to specify whether or not the targets can be arranged in a suitable order of ranking and how consistent the pairwise judgmental or comparison matrices are. The consistency index (C.I.) and the consistency ration (C.R.) for a comparison matrix are calculated using the following equations.

$$C.I. = \frac{(\lambda_{\max} - n)}{(n - 1)}, \tag{7}$$

$$C.R. = \left( \frac{C.I.}{RI_{(n)}} \right) 100\%, \tag{8}$$

where, in this case,  $n$  is the number of variable groups to be compared,  $\lambda_{\max}$  is the largest eigenvalue of the comparison or judgmental matrix, and  $RI_{(n)}$  is the random index based on  $n$ . Following this computation, it can be decided that if the C.R. of a certain comparison matrix is less than 0.1 (10%), the consistency of the judgment is considered to be acceptable [61].

For instance, taking a fuzzy pairwise comparison matrix of the variable group, the associated crisp matrix was obtained using a defuzzified method followed by Chan et al. [61], that is,  $\tilde{M}_{\text{crisp}} = (l + 4m + u)/6$ .

where 1,  $m$ , and  $u$  denote triangular fuzzy numbers.

		$V_1$	$V_2$	$V_3$
$\tilde{M}_{\text{crisp}} =$	$V_1$	1	4.0498	4.6775
	$V_2$	0.251	1	1.4609
	$V_3$	0.2183	0.7116	1

$$\lambda_{\max} = 3.0309, n = 3, \text{ and } RI_{(n=3)} = 0.58 \text{ [43].}$$

Therefore, the C.I., using equation (7), and the C.R., using equation (8), of the pairwise judgmental matrix can be calculated as follows:  $C.I. = (3.0309 - 3)/(3 - 1) = 0.0155$   $C.R. = 0.0155/0.58 = 0.0267 < 0.1$

Thus, the judgmental matrix is acceptable.

The consistency ratios of all other matrices are shown in Tables 4–6. It was found that all ratios are less than 10%. Thus, all the judgments are consistent.

(d) In order to obtain the fuzzy weights of variable groups, the fuzzy geometric mean value  $\tilde{r}_i$ , for each variable group  $i$  is first computed using equation (4) from the matrix  $\tilde{A}$ .

$$\begin{aligned} \tilde{r}_1 &= [\tilde{\alpha}_{11} \otimes \tilde{\alpha}_{12} \otimes \tilde{\alpha}_{13}]^{1/3} \\ &= (1 \times 3.2169 \times 3.5137)^{1/3} \\ &\quad (1 \times 4.0179 \times 4.6915)^{1/3} \\ &\quad (1 \times 5.0104 \times 5.7851)^{1/3} \\ &= (2.2442, 2.6614, 3.0718). \end{aligned} \tag{9}$$

Likewise, the remaining  $\tilde{r}_i$  values are obtained as follows:  $\tilde{r}_2 = (0.5878, 0.7093, 0.8554)$ .  $\tilde{r}_3 = (0.4411, 0.5298, 0.6540)$

To determine the fuzzy weight of each variable group, equation (5) is applied.



TABLE 4: The weight of the input variables with respect to the variable group ( $V_1$ ).

Input variable ( $V_{1j}$ )	Fuzzy weight	Crisp weight (BNP)
Project size	(0.1659, 0.2662, 0.4180)	0.2834
Project type	(0.0735, 0.1182, 0.1965)	0.1294
Project complexity	(0.1453, 0.2396, 0.3910)	0.2586
Project scope	(0.1068, 0.1717, 0.2772)	0.1853
Project duration	(0.0708, 0.1127, 0.1811)	0.1216
Project location	(0.0573, 0.0916, 0.1496)	0.0995

TABLE 5: The weight of input variables with respect to the variable group ( $V_2$ ).

Input variable ( $V_{2j}$ )	Fuzzy weight	Crisp weight (BNP)
Existence of groundwater	(0.2100, 0.3240, 0.4839)	0.3393
Bridge type	(0.1802, 0.2626, 0.3974)	0.2801
Number of bridges	(0.2753, 0.4134, 0.6211)	0.4366

TABLE 6: The weight of input variables with respect to the variable group ( $V_3$ ).

Input variable ( $V_{3j}$ )	Fuzzy weight	Crisp weight (BNP)
Inflation rate	(0.2234, 0.3438, 0.5199)	0.1293
Site topography	(0.2951, 0.4322, 0.6333)	0.1442
Soil type	(0.153, 0.2241, 0.3347)	0.0798

$$\begin{aligned} \tilde{w}_1 &= \tilde{r}_1 \otimes (\tilde{r}_1 \oplus \tilde{r}_2 \oplus \tilde{r}_3)^{-1} \\ &= (2.2442, 2.6614, 3.0718) \otimes \begin{pmatrix} \frac{1}{(3.0718 + 0.8554 + 0.6540)} \\ \frac{1}{(2.6614 + 0.7093 + 0.5298)} \\ \frac{1}{(2.2442 + 0.5878 + 0.4411)} \end{pmatrix} \\ &= (0.4999, 0.6823, 0.9385). \end{aligned} \tag{10}$$

Similarly, the remaining fuzzy weights  $\tilde{w}_i$  are computed as

$$\begin{aligned} \tilde{w}_2 &= (0.1283, 0.1818, 0.2614), \\ \tilde{w}_3 &= (0.0963, 0.1358, 0.1998). \end{aligned} \tag{11}$$

- (e) Computing the BNP value of the fuzzy weights of each variable dimension using COA method. For instance, to find the BPN, the value of the weight of the variable group ( $V_1$ ) is calculated using equation (6) as follows:

$$\begin{aligned} \text{BNP}_{w1} &= \frac{[(u_{w1} - l_{w1}) + (m_{w1} - l_{w1})]}{3} + l_{w1} \\ &= \frac{[(0.9385 - 0.4999) + (0.6825 - 0.4999)]}{3} + 0.4999 \\ &= 0.1837. \end{aligned} \tag{12}$$

In a similar fashion, the BNP value of the other two variable groups can be found, and the results are shown in Table 7. Table 7 also shows the relative fuzzy weight (score), as well as of the total integral crisp values of each variable group obtained by fuzzy AHP based on geometric mean value analysis.

According to Table 7, the first variable group ( $V_1$ ) exhibits the highest importance in comparison with the other two variable groups ( $V_2$  and  $V_3$ ). Certainly,  $V_1$  ranks almost 3.08 times more important than  $V_2$  and nearly 7.01 times more important than  $V_3$ .

Following a similar computational procedure, the relative weights of the input variables placed at the third level in the hierarchy structure can be obtained (see Tables 4–6).

**4.4. Input Variable Selection Process.** Input variables are suggested to be used in the PCE model, one at a time, starting with the input variable closest to the most preferred to the least preferred input variables as shown in Figure 5. So, to give preference for input variables, Euclidean distance was adopted to calculate each distance [23, 62]. The formula to find each distance to ideal input (points) is shown as follows:

$$\text{distance to ideal input variable} = \sqrt{(C_i - X)^2 + (W_i - Y)^2}, \tag{13}$$

where  $C_i$  = the variable coefficient obtained from factor analysis,  $X = 1$ , the ideal maximum variable coefficient based on the coefficient analysis and the ideal value as shown in Figure 5,  $W_i$  = the relative importance weight (BNP value) obtained from fuzzy AHP,  $Y = 1$ , the ideal maximum weight (BNP value) based on fuzzy AHP analysis and the ideal value as shown in Figure 5, and  $i$  = the input variables being evaluated.

Applying equation (11), the proposed input variable preference order is found based on the distance to the ideal input variable and shown in Table 8.

## 5. Discussions and Contributions

The selection of the appropriate input variables can improve the accuracy of the cost estimate, particularly at the early stages of the project development. In this study, a set of 12 input variables or attributes were considered. This study explored the appropriate set of input variables for the purpose of PCE in highway construction projects in the case of the Ethiopian construction industry. The hybrid

TABLE 7: The weight of variable groups with respect to the goal.

Variable group ( $V_i$ )	Fuzzy weight	Crisp weight (BNP)
Variable group ( $V_1$ )	(0.4999, 0.6823, 0.9385)	0.1837
Variable group ( $V_2$ )	(0.1283, 0.1818, 0.2614)	0.0597
Variable group ( $V_3$ )	(0.0963, 0.1358, 0.1998)	0.0462

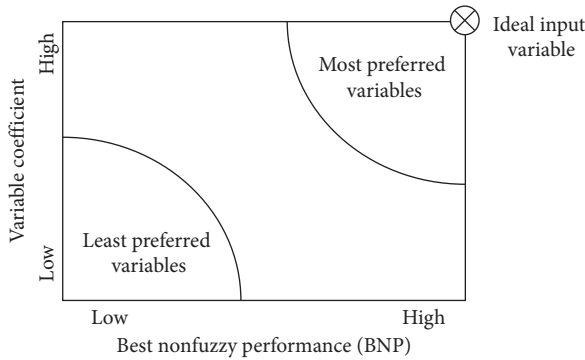


FIGURE 5: Preference for selecting input variables.

methodology of factor analysis and fuzzy AHP are used to supplement the rational input variable selection process. The input variable coefficient and weight (BNP value) are shown in Table 8 along with the computed distance to the ideal input variable shown in Figure 5. The PCE model can use the input variables by selecting them in the order beginning with the shortest distance from the ideal input to the largest distance (applying this for each variable group). Hence, each time, a new set of three input variables, one from each variable group, can be selected and added to the model. In order to verify the practicability on the proposed input variable selection approach, the process can be repeated in the reverse order, meaning starting with the input variable having the largest distance from the ideal input variable. According to the results of the distance computed shown in Table 8, the PCE model development can start with a set of three input variables, one from each variable group, such as project size, number of bridges, and inflation rate. These three input variables exhibit a short distance from the ideal input variable (point). This implies that these input variables are preferred first because of their high influence on construction cost as well as high relative importance when compared to other variables for PCE.

This proved that this study has validity in preliminary or conceptual cost-estimation practice as “project size” is found to be the most significant and important input variable (from the first variable group ( $V_1$ )), in the case of highway construction projects. In regard to project size or length, it can be generalized that, the larger the project, the more expensive it will be. This finding can be supported by other researchers [23, 63]. Elfaki et al. [26] also proved that the project size and the number of labours have a strong correlation. Looking into the second variable group, it can be straightforwardly found out that “number of bridges”

has been evaluated as the first preferred input variable. The bridge construction becomes the major part of construction projects, particularly in highway projects, where waterbody exists. The number of bridges in the project scope greatly affects the cost of construction and has a direct relationship with the construction cost as it enlarges the overall scope of the project. In the third variable group, inflation was also the most preferred input variable for PCE of highway construction projects. So, this input variable should be a serious concern while estimating and managing the cost of highway projects in the case organization. This finding can be justified as it is in line with the findings of previous studies [60]. It is pretty much factual as inflation increases the original estimate of construction project costs. During the planning and design stage of project development, fluctuations in the rate of inflation can lead to underestimation of project costs. Inflation may have been taken into account in the original cost estimates; however, if the inflation rate increases beyond the forecasted level during project implementation, then the original estimate will be surpassed.

### 5.1. Elimination of Multicollinearity among Input Variables.

In the realm of model development, multicollinearity can be described as a condition that one or more of the cost factors be interrelated with each other [17]. If the predictors or input variables are highly correlated, in the case of models involving more than one input variable, each account for a similar variance in the outcome. So, input variables should not negatively or positively correlate too highly [16]. Table 9 shows the SPSS output of the correlation matrix (R-matrix). It contains the Pearson correlation coefficient between all pairs of input variables. The results demonstrate that multicollinearity exists among most of the input variables as there is a relatively strong positive and negative relationship between them. The existence of multicollinearity will not only cause repeatability errors and trade-offs but also makes it difficult to assess the individual importance of an input variable or predictor [16, 17, 64].

To overcome collinearity problems in the model, e.g., multiple regression, a factor analysis based on principal component analysis was carried out on the input variables to reduce them to a subset of uncorrelated variables. In the present study, the SPSS software was run to perform a factor analysis and the input variables causing the multicollinearity were combined to form a component or variable group. One of the purposes of applying factor analysis in this study was to vanish the problem of multicollinearity. To ensure that the components are uncorrelated, the Anderson–Rubin method was used while carrying out factor analysis to determine the variable group scores which were later used to calculate the variable coefficients [65].

Furthermore, the correlation and covariance among the three variable groups were performed as shown in Table 10. Table 10 shows that the three variable groups or components have no correlations, with 0 Pearson correlation coefficient, and are independent from each other. Therefore, the above discussion can securely prove that the applying factor

TABLE 8: Selection order of input variables and distance to ideal input.

Variable group ( $V_i$ )	Symbol	Coefficient	BNP	Euclidean distance	Preferred order
Project size	$V_{11}$	0.1525	0.2834	1.1099	1
Project complexity	$V_{13}$	0.1417	0.2586	1.1341	2
Project scope	$V_{14}$	0.1274	0.1853	1.1938	3
Project type	$V_{12}$	0.1499	0.1294	1.2168	4
Project duration	$V_{15}$	0.1253	0.1216	1.2396	5
Project location	$V_{16}$	0.1044	0.0995	1.2701	6
<i>Variable group (<math>V_2</math>)</i>					
Number of bridges	$V_{23}$	0.2436	0.4366	0.9432	1
Existence of groundwater	$V_{21}$	0.3191	0.3393	0.9488	2
Bridge type	$V_{22}$	0.2804	0.2801	1.0179	3
<i>Variable group (<math>V_3</math>)</i>					
Inflation rate	$V_{31}$	0.3409	0.1293	1.0921	1
Site topography	$V_{32}$	0.2657	0.1442	1.1276	2
Soil type	$V_{33}$	0.2195	0.0798	1.2066	3

TABLE 9: Correlation matrix of the 12 input variables.

	$V_{12}$	$V_{13}$	$V_{16}$	$V_{14}$	$V_{11}$	$V_{32}$	$V_{22}$	$V_{23}$	$V_{21}$	$V_{33}$	$V_{31}$	$V_{15}$
$V_{12}$	1.000											
$V_{13}$	0.545	1.000										
$V_{16}$	0.287	0.270	1.000									
$V_{14}$	0.381	0.317	0.380	1.000								
$V_{11}$	0.425	0.338	0.347	0.713	1.000							
$V_{32}$	0.092	-0.001	0.249	0.247	0.102	1.000						
$V_{22}$	0.336	0.314	0.472	0.444	0.341	0.222	1.000					
$V_{23}$	0.405	0.360	0.394	0.438	0.383	0.242	0.768	1.000				
$V_{21}$	0.023	0.043	0.216	0.325	0.132	0.325	0.526	0.353	1.000			
$V_{33}$	0.000	0.059	0.113	0.164	0.079	0.324	0.181	0.225	0.388	1.000		
$V_{31}$	-0.059	-0.100	0.337	0.159	-0.006	0.306	0.122	0.092	0.136	0.349	1.000	
$V_{15}$	0.248	0.314	0.234	0.206	0.362	0.202	0.273	0.285	-0.013	0.025	0.083	1.000

TABLE 10: Variable group score correlation and covariance matrix.

$V_i$	Correlation matrix			Covariance matrix		
	$V_1$	$V_2$	$V_3$	$V_1$	$V_2$	$V_3$
$V_1$	1			1.000		
$V_2$	0.000	1		0.000	1.000	
$V_3$	0.000	0.000	1	0.000	0.000	1.000

analysis based on the principal component analysis in the process of input variable selection can overcome the problem of collinearity among the input variables.

5.2. Dimensional Reduction and Categorization of the Input Variables. During the PCE process, utilizing a large amount of input variables makes the process very complicated. Moreover, Gransberg et al. [21] also stated that the data-driven conceptual cost-estimation models do not require to incorporate a large number of project attributes or variables to predict the cost of construction to reasonable accuracy at the early phase of project development. In the present study, a set of thirteen input variables was initially identified that can be used as model variables for PCE. If all these input variables are considered in the estimation process, it becomes time-consuming to collect the required data and

formulate the historical database as well as requires a large amount of effort in realising a realistic cost estimate. However, by applying the factor analysis in the input variable selection process, only a few variables can replace the large portion of the set of 12 variables so that the dimensionality of the data set can be reduced without significant loss of the information and simplicity to the whole process can be realised as well. Field [16] declared that any further analysis can be carried out on the variable group scores rather than the original data. The elements' factor (variable) loadings for a given variable group are used to calculate the input variable coefficients. Therefore, in this study, the determination of the coefficient of each input variable under each variable group was performed (see Table 3). The coefficient of a given input variable under a certain variable group indicates the level of effect on that variable group. At this stage, only three input variables, one

from each variable group, having high coefficient could be selected as input variables for PCE as they are not uncorrelated with each other. However, further analysis was still continued to deal with incomplete and uncertain knowledge and information during input variable evaluation, which enables it to meet the dual-objective goal proposed in this study.

**5.3. Handling Uncertainty and Vagueness.** In most cases, fuzziness and uncertainty exist in practice. Thus, fuzzy set theory integrated with AHP was used to effectively handle subjective perceptions and impreciseness, allowing the appropriate expression of linguistic evaluation. In this study, fuzzy AHP based on the geometric mean method was employed to determine the relative importance weight of the input variables with respect to each variable group in a fuzzy environment where the subjectivity and vagueness are handled with natural language expressions parameterized by triangular fuzzy numbers. This makes the input variable selection process more rational and effective in PCE.

## 6. Conclusions and Remarks

The aim of this study is to present a rational and systematic input variable selection approach for the PCE model in the highway construction industry. A set of 12 input variables in three categories, resulting from factor analysis, were involved in variable coefficient determination. In addition, these input variables located under each category were also evaluated by highway experts (decision makers) in terms of their relative importance, and the subjectivity and uncertainty of human assessment are taken into account through the fuzzy set theory in a fuzzy environment. The proposed approach in this study was aimed to meet the dual-objective goal during PCE and it was proposed that input variables be selected beginning with those which have high coefficient but also have high relative importance weight. This was realised by the use of Euclidean distance to calculate each distance from an ideal input variable. Based on the distance results, the input variables are suggested to be selected for the PCE model development to predict the construction cost in the order of starting with the shortest distance from the ideal input variable to the largest distance. Accordingly, each time as a new set of three input variables, one from each group, can be added to the PCE model. Based on the results of this study, particularly distance results, variables such as project size, number of bridges, and inflation rate from variable groups  $V_1$ ,  $V_2$ , and  $V_3$ , respectively, were found to be the first preferred input variables, and the PCE model can start with these three variables. As explained in the previous sections, the input variables which are categorized in the different group are not correlated with each other to realise the more accurate estimate. This is the reason why one input variable from each variable group is suggested to be selected at a time.

The proposed input variable selection approach in this study is just a tailor-made method for the highway construction projects which intends to present the ideas about

how to fix up the combined factor analysis and fuzzy AHP. When it is supposed to be applied to other projects or construction sectors, it is required to first study the project's characteristics and the sector's requirements. The integrated input variable selection process is designed to provide practitioners with a fuzzy point of view to the traditional variable selection technique for dealing with imprecision. The proposed method enables cost estimators or decision analysts to better understand the comprehensive input variable selection process. Moreover, this approach provides a more accurate, systematic, and rational decision support tool. In the end, further study can apply the other factor prioritization techniques under fuzzy environment and the results can be compared with the present method. The proposed input variable selection process has not been validated in this study. The practicability and validation of the proposed model will be verified upon the completion of the undergoing study by the present authors.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

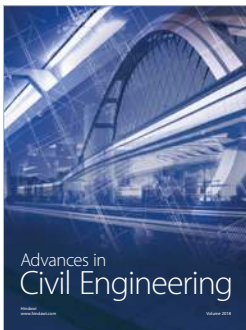
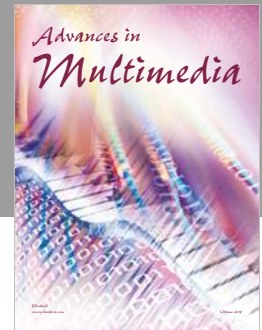
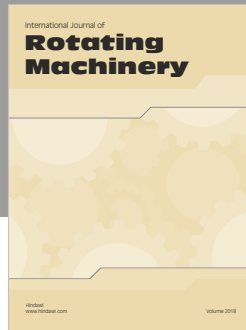
The authors declare that there are no conflicts of interest regarding the publication of this article.

## References

- [1] B. A. Tayeh, K. Al Hallaq, W. S. Alaloul, and A. R. Kuhail, "Factors affecting the success of construction projects in Gaza Strip," *Open Civil Engineering Journal*, vol. 12, no. 1, pp. 301–315, 2018.
- [2] M. A. Abdal-Hadi, "Factors affecting accuracy of pre-tender cost estimate: studies of Saudi Arabia," *International Journal of Applied Engineering Research*, vol. 9, no. 1, pp. 21–36, 2014.
- [3] A. Enshassi, S. Mohamed, and M. Abdel-Hadi, "Factors affecting the accuracy of pre-tender cost estimates in the Gaza Strip," *Journal of Construction in Developing Countries*, vol. 18, no. 1, pp. 73–94, 2013.
- [4] L. Xin-Lianga, Y. Taob, and W. Guo-Dongc, "Practical application study of Gaussian process model in construction project cost estimation," *Advanced Materials Research*, vol. 671–674, pp. 3100–3106, 2013.
- [5] P. O. Alumbugu, Department of Quantity Surveying, W. A. Ola-Awo, S. Ibrahim, A. Mustapha Muhammed, and A. Abdulazeez, "Assessment of the factors affecting accuracy of pre-tender cost estimate in Kaduna state, Nigeria," *IOSR Journal of Environmental Science, Toxicology and Food Technology*, vol. 8, no. 5, pp. 19–27, 2014.
- [6] A. Britto and B. A. K. S. Perera, "Factors Affecting the Accuracy of Pre-Tender Estimation of Road Construction in Sri Lanka," Faculty of Architecture Research Unit, University of Moratuwa, Moratuwa, Sri Lanka, 2013, <http://dl.lib.mrt.ac.lk/handle/123/9876>.
- [7] M. T. Hatamleh, M. Hiyassat, G. J. Sweis, and R. J. Sweis, "Factors affecting the accuracy of cost estimate: case of Jordan," *Engineering, Construction and Architectural Management*, vol. 25, no. 1, pp. 113–131, 2018.

- [8] I. Mahamid, "Factors affecting cost estimate accuracy: evidence from Palestinian construction projects," *International Journal of Management Science and Engineering Management*, vol. 10, no. 2, pp. 117–125, 2015.
- [9] A. N. Ofori-Boadu, "Exploring regression models for forecasting early cost estimates for high-rise buildings," *Journal of Technology, Management, and Applied Engineering*, vol. 31, no. 5, pp. 2–24, 2015.
- [10] A. M. Jarkas, S. A. Mubarak, and C. Y. Kadri, "Critical factors determining bid/no bid decisions of contractors in Qatar," *Journal of Management in Engineering*, vol. 30, no. 4, article 5014007, 2013.
- [11] T. T. Choon, L. C. Sim, T. Connie, A. K. Nita, A. G. Uche, and G. K. Chen, "Influential factors in estimating and tendering for construction work," *MATEC Web of Conferences*, vol. 47, article 04007, 2016.
- [12] G. Mahalakshmi and C. Rajasekaran, "Early cost estimation of highway projects in India using artificial neural network," *Lecture Notes in Civil Engineering*, vol. 25, pp. 659–672, 2019.
- [13] I. Mensah, T. Adjei-Kumi, and G. Nani, "Duration determination for rural roads using the principal component analysis and artificial neural network," *Engineering, Construction and Architectural Management*, vol. 23, no. 5, pp. 838–856, 2016.
- [14] T. Shehab and I. Meisami-Fard, "Cost-estimating model for rubberized asphalt pavement rehabilitation projects," *Journal of Infrastructure Systems*, vol. 19, no. 4, pp. 496–502, 2013.
- [15] O. Sweil, J. Gregory, and R. Kirchain, "Construction cost estimation: a parametric approach for better estimates of expected cost and variation," *Transportation Research Part B: Methodological*, vol. 101, pp. 295–305, 2017.
- [16] A. Field, *Discovering Statistics Using IBM SPSS Statistics*, SAGE Publications Ltd, Thousand Oaks, CA, USA, 4th edition, 2013.
- [17] K.-C. Lam, R. Tao, and M. C.-K. Lam, "A material supplier selection model for property developers using fuzzy principal component analysis," *Automation in Construction*, vol. 19, no. 5, pp. 608–618, 2010.
- [18] C.-Y. Chao and K.-F. Yang, "Combining fuzzy set theory and extent analysis to construct an integrated decision making approach in medical cosmetology industry," *Information Technology Journal*, vol. 10, no. 10, pp. 1950–1956, 2011.
- [19] S. S. Arage and N. V. Dharwadkar, "Cost estimation of civil construction projects using machine learning paradigm," in *Proceedings of the 2017 International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC 2017)*, pp. 594–599, Palladam, India, 2017.
- [20] D. D. Ahiaga-Dagbui and S. D. Smith, "Neural networks for modelling the final target cost of water projects," in *Proceedings of the 28th Annual ARCOM Conference*, pp. 307–316, Association of Researchers in Construction Management, Edinburgh, UK, September 2012.
- [21] D. D. Gransberg, H. D. Jeong, I. Karaca, and B. Gardner, "Top-down construction cost estimating model using an artificial neural network," Final Report FHWA/MT-17-007/8232-001, Montana Department of Transportation, Helena, MT, USA, 2017.
- [22] S. Kim and J. H. Shim, "Combining case-based reasoning with genetic algorithm optimization for preliminary cost estimation in construction industry," *Canadian Journal of Civil Engineering*, vol. 41, no. 1, pp. 65–73, 2014.
- [23] B. J. Gardner, D. D. Gransberg, and H. D. Jeong, "Reducing data-collection efforts for conceptual cost estimating at a highway agency," *Journal of Construction Engineering and Management*, vol. 142, no. 11, article 04016057, 2016.
- [24] F. Shutian, Z. Tianyi, and Z. Ying, "Prediction of construction projects' costs based on fusion method," *Engineering Computations*, vol. 34, no. 7, pp. 2396–2408, 2017.
- [25] B. Zhu, L.-A. Yu, and Z.-Q. Geng, "Cost estimation method based on parallel Monte Carlo simulation and market investigation for engineering construction project," *Cluster Computing*, vol. 19, no. 3, pp. 1293–1308, 2016.
- [26] A. O. Elfaki, S. Alatawi, and E. Abushandi, "Using intelligent techniques in construction project cost estimation: 10-year survey," *Advances in Civil Engineering*, vol. 2014, Article ID 107926, 11 pages, 2014.
- [27] Z. Ma, Z. Liu, and Z. Wei, "Formalized representation of specifications for construction cost estimation by using ontology," *Computer-Aided Civil and Infrastructure Engineering*, vol. 31, no. 1, pp. 4–17, 2016.
- [28] K. H. Hyari, A. Al-daraiseh, and M. El-mashaleh, "Conceptual cost estimation model for engineering services in public construction projects," *Journal of Management in Engineering*, vol. 32, no. 1, article 04015021, pp. 1–9, 2016.
- [29] W.-D. Yu and M. J. Skibniewski, "Integrating neurofuzzy system with conceptual cost estimation to discover cost-related knowledge from residential construction projects," *Journal of Computing in Civil Engineering*, vol. 24, no. 1, pp. 35–44, 2010.
- [30] V. Arabzadeh, S. T. A. Niaki, and V. Arabzadeh, "Construction cost estimation of spherical storage tanks: artificial neural networks and hybrid regression—GA algorithms," *Journal of Industrial Engineering*, vol. 14, no. 4, pp. 747–756, 2017.
- [31] S. Bayram, M. E. Ocal, E. L. Oral, and C. D. Atis, "Comparison of multi layer perceptron (MLP) and radial basis function (RBF) for construction cost estimation: the case of Turkey," *Journal of Civil Engineering and Management*, vol. 22, no. 4, 2015.
- [32] G. Agyekum—Mensah, "The degree of accuracy and factors that influence the uncertainty of SME cost estimates," *International Journal of Construction Management*, vol. 19, pp. 1–14, 2018.
- [33] Y. Zhang and R. E. Minchin, "Parametric cost estimation of resurfacing projects using ridge regression and a generalized feedforward neural network," in *Proceedings of the Construction Research Congress 2016*, pp. 718–728, San Juan, PR, USA, 2016.
- [34] I. Peško, V. Mučenski, M. Šešlija et al., "Estimation of costs and durations of construction of urban roads using ANN and SVM," *Complexity*, vol. 2017, Article ID 2450370, 13 pages, 2017.
- [35] A. Field, *Discovering Statistics Using SPSS*, SAGE Publications Ltd, Thousand Oaks, CA, USA, 3rd edition, 2009.
- [36] N. Weshah, W. El Ghandour, G. Jergeas, and L. C. Falls, "Factor analysis of the interface management (IM) problems for construction projects in Alberta," *Canadian Journal of Civil Engineering*, vol. 40, no. 9, pp. 848–860, 2013.
- [37] M. N. Omar and A. R. Fayek, "Modeling and evaluating construction project competencies and their relationship to project performance," *Automation in Construction*, vol. 69, pp. 115–130, 2016.
- [38] L. A. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, no. 3, pp. 338–353, 1965.
- [39] I. Dzitac, "The fuzzification of classical structures: a general view," *International Journal of Computers Communications & Control*, vol. 10, no. 6, pp. 772–788, 2015.

- [40] S. Nadaban, S. Dzitac, and I. Dzitac, "Fuzzy topsis: a general view," *Procedia Computer Science*, vol. 91, pp. 823–831, 2016.
- [41] A. Emrouznejad and W. Ho, *Fuzzy Analytic Hierarchy Process*, CRC Press, Taylor & Francis Group, Boca Raton, FL, USA, 2018.
- [42] P. Sirisawat and T. Kiatcharoenpol, "Fuzzy AHP-topsis approaches to prioritizing solutions for reverse logistics barriers," *Computers & Industrial Engineering*, vol. 117, pp. 303–318, 2018.
- [43] S. Kaganski, J. Majak, and K. Karjust, "Fuzzy AHP as a tool for prioritization of key performance indicators," *Procedia CIRP*, vol. 72, pp. 1227–1232, 2018.
- [44] Y. O. Ouma, J. Opudo, and S. Nyambenya, "Comparison of fuzzy AHP and fuzzy TOPSIS for road pavement maintenance prioritization: methodological exposition and case study," *Advances in Civil Engineering*, vol. 2015, Article ID 140189, 17 pages, 2015.
- [45] G. Polat, E. Eray, and B. N. Bingol, "An integrated fuzzy MCGDM approach for supplier selection problem," *Journal of Civil Engineering and Management*, vol. 23, no. 7, pp. 926–942, 2017.
- [46] N. Prascevic and Z. Prascevic, "Application of fuzzy Ahp for ranking and selection of alternatives in construction project management," *Journal of Civil Engineering and Management*, vol. 23, no. 8, pp. 1123–1135, 2017.
- [47] S. Rajaprakash, R. Ponnusamy, R. Jaichandran, K. Karthik, and K. Somasundaram, "Determining the school students stress factors using fuzzy AHP process," *International Journal of Pure and Applied Mathematics*, vol. 118, no. 5, pp. 889–900, 2018.
- [48] M. Sharma, R. Gupta, and P. Acharya, "Prioritizing the critical factors of cloud computing adoption using multi-criteria decision-making techniques," *Global Business Review*, vol. 20, no. 5, article 097215091774118, pp. 1–20, 2017.
- [49] K. Sharawat and S. K. Dubey, "An approach to vendor selection on usability basis by AHP and fuzzy topsis method," *Advances in Intelligent Systems and Computing*, vol. 584, pp. 595–604, 2018.
- [50] T. Yaghoobi, "Prioritizing key success factors of software projects using fuzzy AHP," *Journal of Software: Evolution and Process*, vol. 30, no. 1, pp. 1–11, 2018.
- [51] J. J. Buckley, "Fuzzy hierarchical analysis," *Fuzzy Sets and Systems*, vol. 17, no. 3, pp. 233–247, 1985.
- [52] C.-C. Sun, "A performance evaluation model by integrating fuzzy AHP and fuzzy topsis methods," *Expert Systems with Applications*, vol. 37, no. 12, pp. 7745–7754, 2010.
- [53] A. Meier, E. Portmann, K. Stoffel, and T. Tamayo, *The Application of Fuzzy Logic for Managerial Decision Making Processes: Latest Research and Case Studies*, Springer International Publishing, New York, NY, USA, 2017.
- [54] E. Elbeltagi, O. Hosny, R. Abdel-razek, and A. El-fitory, "Conceptual cost estimate of Libyan highway projects using artificial neural network," *International Journal of Engineering Research and Applications*, vol. 4, no. 8, pp. 56–66, 2014.
- [55] B. J. Gardner, D. D. Gransberg, and J. A. Rueda, "Stochastic conceptual cost estimating of highway projects to communicate uncertainty using bootstrap sampling," *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, vol. 3, no. 3, article 05016002, 2017.
- [56] *Cost Estimating Manual for Projects*, Washington State Department of Transportation (WSDOT), Washington DC, USA, 2015.
- [57] K. Kimiafar, F. Sadoughi, A. Sheikhtaheri, and M. Sarbaz, "Prioritizing factors influencing nurses' satisfaction with hospital information systems," *CIN: Computers, Informatics, Nursing*, vol. 32, no. 4, pp. 174–181, 2014.
- [58] S. Kaleem, M. Irfan, and H. F. Gabriel, "Estimation of highway project duration at the planning stage and analysis of risk factors leading to time overrun," in *Proceedings of the Second Congress of the Transportation and Development Institute of ASCE*, pp. 612–626, Orlando, FL, USA, June 2014.
- [59] I. Mahamid, "Effects of project's physical characteristics on cost deviation in road construction," *Journal of King Saud University—Engineering Sciences*, vol. 25, no. 1, pp. 81–88, 2013.
- [60] J. S. Shane, K. R. Molenaar, S. Anderson, and C. Schexnayder, "Construction project cost escalation factors," *Journal of Management in Engineering*, vol. 25, no. 4, pp. 221–229, 2009.
- [61] K. Y. Chan, C. K. Kwong, and T. S. Dillon, "An enhanced fuzzy AHP method with extent analysis for determining importance of customer requirements," *Computational Intelligence Techniques for New Product Design*, vol. 403, pp. 79–93, 2012.
- [62] P.-E. Danielsson, "Euclidean distance mapping," *Computer Graphics and Image Processing*, vol. 14, no. 3, pp. 227–248, 1980.
- [63] I. Mahamid, "Early cost estimating for road construction projects using multiple regression techniques," *Construction Economics and Building*, vol. 11, no. 4, pp. 87–101, 2011.
- [64] A. Forghani, S. J. Sadjadi, and B. F. Moghadam, "A supplier selection model in pharmaceutical supply chain using PCA, Z-TOPSIS and MILP: A case study," *PLoS One*, vol. 13, no. 8, Article ID e0201604, 2018.
- [65] B. G. Tabachnick and L. S. Fidell, *Using Multivariate Statistics*, Pearson Education Limited, London, UK, 6th edition, 2013.



**Hindawi**

Submit your manuscripts at  
[www.hindawi.com](http://www.hindawi.com)

