Framework to Analyze Construction Labor Productivity Using Fuzzy Data Clustering and Multi-Criteria Decision-Making

Matin KAZEROONI[1], Mohammad RAOUFI[2], and Aminah Robinson FAYEK[3]

[1] MSc Student and Graduate Research Assistant, Hole School of Construction Engineering, Department of Civil and Environmental Engineering, University of Alberta, 7-203 Donadeo Innovation Centre for Engineering, 9211 116 Street NW, Edmonton, Alberta, T6G 1H9, Canada, email: kazeroon@ualberta.ca

[2] Postdoctoral Fellow, Hole School of Construction Engineering, Department of Civil and Environmental Engineering, University of Alberta, 7-381 Donadeo Innovation Centre for Engineering, 9211 116 Street NW, Edmonton, Alberta, T6G 1H9, Canada, email: mraoufi@ualberta.ca

[3] Professor, Director of the Construction Innovation Centre, Tier 1 Canada Research Chair in Fuzzy Hybrid Decision Support Systems for Construction, NSERC Industrial Research Chair in Strategic Construction Modeling and Delivery, Ledcor Professor in Construction Engineering, Hole School of Construction Engineering, Department of Civil and Environmental Engineering, University of Alberta, 7-232 Donadeo Innovation Centre for Engineering, 9211 116 Street NW, Edmonton, Alberta, T6G 1H9, Canada, PH: (780) 492-1205, email: aminah.robinson@ualberta.ca

ABSTRACT

Construction labor productivity (CLP) has a significant impact on the performance and profitability of construction projects. A construction project can benefit from improved labor productivity in many ways, such as a shorter project life cycle and lower project cost. However, budget and resource restrictions force construction companies to select and implement only the most effective CLP improvement strategies. Analyzing labor productivity in order to determine the most effective CLP improvement strategies is a difficult task because labor productivity is influenced by numerous subjective and objective factors. This paper presents a framework for ranking the factors affecting CLP according to their importance for CLP improvement; the framework uses an integration of fuzzy data clustering and multi-criteria decision-making methods. The proposed framework entails asking experts to weight determinant criteria for selecting CLP improvement strategies and then clustering CLP factors and ranking the clusters. This paper's major contribution is providing a systematic approach for analyzing and selecting CLP improvement strategies by identifying the CLP factors with the greatest impact on productivity improvement. The findings of this research will help establish a set of CLP improvement strategies in order to enhance CLP.

INTRODUCTION

Many activities in the construction industry are labor intensive. Therefore, improving construction labor productivity (CLP) is key for improving the overall performance of construction organizations in multiple areas, such as shortening project life cycles and lowering project costs. However, analyzing and improving labor productivity is difficult, as CLP occurs in a complex environment where numerous objective and subjective factors influence labor productivity.

Managers of construction companies apply a variety of CLP improvement strategies according to their knowledge and experience, but many do not use a systematic approach for considering the effectiveness of these strategies (Nasir et al. 2015). However, some limitations, such as the cost of implementation and limited resources, restrict companies from adopting multiple CLP improvement strategies. The main challenge for construction management teams is identifying the key factors influencing labor productivity in order to be able to prioritize CLP improvement strategies.

The objective of this paper is to provide a framework for ranking CLP factors according to their importance for CLP improvement in order to assist construction companies with the prioritization of CLP improvement strategies. The framework involves four steps. First, three determinant criteria for selecting CLP improvement strategies, namely strategy selection criteria (SSCs), are defined. Then, the weight of each criterion is evaluated using experts' opinions and the fuzzy analytic hierarchy process (FAHP) as a fuzzy multi-criteria decision-making (MCDM) method. Next, the fuzzy c-means (FCM) method is used to cluster the CLP factors based on their similarities and dissimilarities in terms of the SSCs. Finally, the Technique for Order Preference by Similarity to Ideal Solution (TOPSIS), a widely used MCDM method, is employed to rank the CLP clusters according to their importance for CLP improvement.

Although there are numerous studies on ranking the factors that affect labor productivity on different types of locations and projects, little research has been done on ranking CLP factors in terms of CLP improvement strategy selection. This study fills this gap in the research by presenting a new framework for ranking CLP factors that uses an MCDM method and fuzzy data clustering to determine the factors' importance for CLP improvement strategy selection. The outcomes of this study help construction management teams identify key CLP factors and implement improvement strategies in order to reinforce the factors that positively affect labor productivity and eliminate the factors that have a negative impact on labor productivity.

In this study, a CLP improvement strategy is a management strategy that comprises several management practices. According to Ghodrati et al. (2018), management practices are individual practices that are carried out by a construction management team to improve labor productivity. For instance, an incentive program is a CLP improvement strategy that consists of several management practices: performance-based incentives, health and safety incentives, incentives for no rework, etc.

This paper is organized as follows: First, a review of past research on the identification of key CLP factors and key CLP improvement strategies is provided. Second, the framework for ranking CLP factors is presented along with an example. The framework consists four phases: (1) component identification (i.e., list of factors influencing labor productivity and determinant criteria for selecting CLP improvement strategies); (2) data collection (i.e., development of survey questionnaires); (3) data preparation (i.e., weighting criteria, checking consistency of data, and aggregating survey data); and (4) data analysis (i.e., clustering data and ranking clusters).

LITERATURE REVIEW

Due to the importance of labor productivity for the overall performance of construction projects, a significant amount of research has been conducted to determine the most influential CLP factors and improve them by developing a variety of analysis models (Heravi and Eslamdoost 2015; Raoufi and Fayek 2018; Alaghbari et al. 2019; Kedir et al. 2019).

The factors that influence CLP are multilevel, ranging from the activity level to the organizational level and to national and global levels (Tsehayae and Fayek 2014; Gerami Seresht and Fayek 2019). The different perspectives of personnel (e.g., project managers, supervisors, craft workers and foremen) are therefore required in order to assess the importance of each CLP factor for CLP improvement. Several studies have incorporated the opinions of different project participants through interview and questionnaire surveys and categorized the CLP factors under different groups. For example, Alaghbari et al. (2019) categorized 52 factors under the four groups: human-labor, technical and technological, external, and management. In these studies, the most commonly used method for determining the rank of CLP factors is the relative importance index technique, which only considers one criterion, impact (I), when ranking factors.

Tsehayae and Fayek (2014) gathered 169 CLP factors from existing literature related to North American construction projects and investigated their influence on CLP by developing a protocol for collecting data from several construction companies. They not only focused on the impact (I) of CLP factors on labor productivity, but also considered another criterion, frequency or agreement (FoA), when ranking the CLP factors. FoA evaluates the extent to which each factor exists in a project setting.

The construction industry has had many opportunities to improve labor productivity by implementing innovative technologies and techniques. However, the influence of new innovations is not significant without efficient management strategies to control and support labor productivity (Nasir et al. 2015). Different studies have recommended several management strategies, such as training, incentive programs, and communication, to improve labor productivity, but only a few have used a systematic approach to evaluate the effectiveness of these strategies. For example, Nasir et al. (2015) and Ghodrati et al. (2018) developed statistical methods to ascertain the implementation level of some specific CLP improvement strategies. The results of their research reveal that construction projects with a high implementation of certain CLP improvement strategies have experienced higher labor productivity than construction projects with a lower level of implementation.

In spite of extensive research on the identification of key CLP factors and key CLP improvement strategies, few studies have attempted to develop a framework that investigates the importance of CLP factors in terms of CLP improvement strategy selection. By identifying the CLP factors with the most influence on CLP improvement, such a framework would help construction organizations select CLP improvement strategies more systematically. It would also help construction organizations allocate their limited budget and resources to those CLP improvement strategies that target the most important factors for improving labor productivity, rather than putting effort into strategies with minor or no influence on CLP (Ghoddousi et al. 2015).

FRAMEWORK FOR RANKING CLP FACTORS

This section presents the framework for integrating MCDM methods with fuzzy data clustering in order to evaluate the importance of CLP factors and rank them. Figure 1 illustrates the framework for clustering and ranking CLP factors.

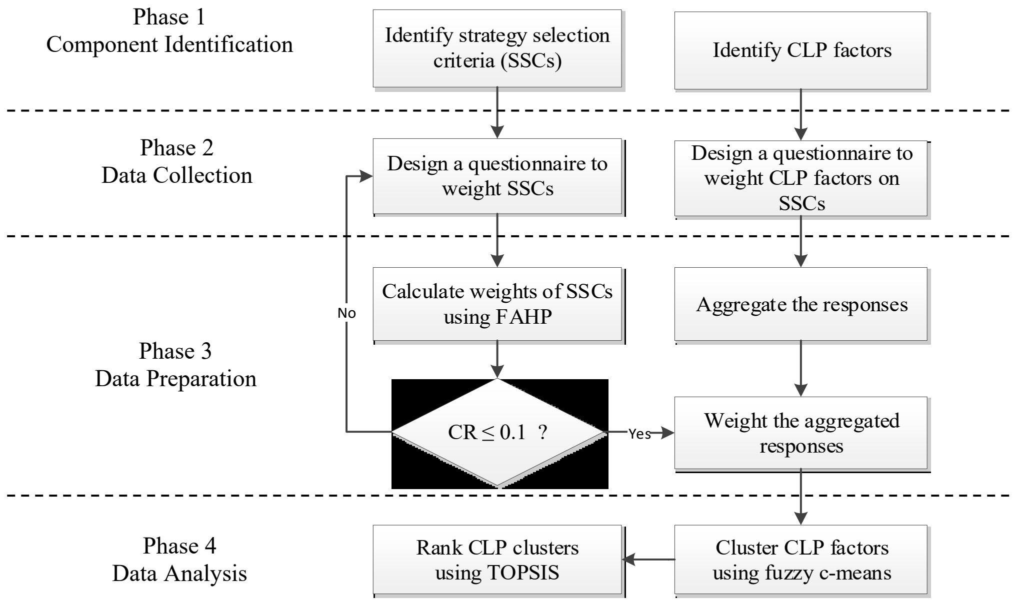

Figure 1. Framework for clustering and ranking CLP factors.

The proposed framework consists of four phases, as follows:

Phase 1: Component Identification

In the first phase, the two main components of the proposed framework are identified. First, a list of factors influencing labor productivity is elicited from existing literature related to construction projects. Second, three determinant criteria for selecting CLP improvement strategies (i.e., SSCs) are defined. These criteria are impact (I), frequency or agreement (FoA), and controllability (Ctrl). According to Tsehayae and Fayek (2014), the criterion I refers to the positive or negative influence of factors on CLP for the project under study and the criterion FoA shows the extent to which each factor exists in a project setting. CLP is a function of controllable and uncontrollable factors (Tsehayae and Fayek 2016). Therefore, selecting CLP improvement strategies is also influenced by the controllability of CLP factors. For instance, a construction company has no control over oil prices, so “volatility of oil prices” is an uncontrollable factor and no improvement strategy can improve it, whereas “job site orientation program for new craftsmen” is a controllable factor to some extent and can be improved by allocating a reasonable amount of time and cost. Accordingly, in this study, the criterion Ctrl is defined as the extent to which each factor can be controlled by a construction company in terms of cost and time.

Phase 2: Data collection

In the second phase, two survey questionnaires are developed, one regarding CLP factors and one regarding criteria. To prevent biased results, survey respondents were randomly selected from a population of 505 construction experts with various positions, such as senior management, project management, and craftspeople in order to capture different perspectives (Tsehayae and Fayek 2014). In the first questionnaire (Table 1), namely the factor importance (FI) survey, the respondents indicate their opinions about CLP factors with respect to each strategy selection criterion using five-point Likert scales. The data collection effort produced a total of 141 FI surveys from construction experts with an average of 10 years of experience (Tsehayae and Fayek 2014). The second questionnaire (Table 2), namely the criterion importance (CI) survey, collects the respondents' opinions on the importance of one strategy selection criterion relative to another. Thus, the CI survey performs pairwise comparisons among I, FoA, and Ctrl by asking the respondents to select a preference term from "equal" to "absolute" when comparing the relative importance of one criterion over another. Each preference term in Table 2 is represented by a symmetric triangular fuzzy number in order to compute the criteria's weights in phase 3. The numbers of experts who selected a specific preference term when comparing the relative importance of one criterion over another are presented in Table 2. For instance, out of 12 experts who responded to the CI survey, seven experts selected the preference term "weak" on the right side of the questionnaire to express that the criterion Ctrl is weakly more important than the criterion I with respect to selecting CLP improvement strategies.

Table 1. Sample FI survey questionnaire.

	Frequency or Agreement					**Impact**					**Controllability**				
	Never	**Rarely**	**Sometimes**	**Often**	**Always**										
CLP Factor	**Disagree**	**Slightly Disagree**	**Neutral**	**Slightly Agree**	**Agree**	**None**	**Slight**	**Moderate**	**Strong**	**Very Strong**	**Very Low**	**Low**	**Moderate**	**High**	**Very High**
Power equipment breakdown	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
Instability of political system	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
Lack of protection from weather effect	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5

Table 2. CI Survey questionnaire including sample preferences.

Relative Importance of SSCs for Selecting CLP Improvement Strategies											
		Left Criterion is more important					**Right** Criterion is more important				
Question No.	Criterion	Absolute (4, 5, 6)	Very Strong (3, 4, 5)	Fairly Strong (2, 3, 4)	Weak (1, 2, 3)	Equal (1, 1, 1)	Weak (1, 2, 3)	Fairly Strong (2, 3, 4)	Very Strong (3, 4, 5)	Absolute (4, 5, 6)	Criterion
1	I				3	5	3	1			FoA
2	I					3	7	1	1		Ctrl
3	FoA				2	3	5	2			Ctrl

Phase 3: Data preparation

In the third phase, four steps are followed to prepare the collected data for clustering and ranking the CLP factors.

Step 1: Calculate the relative weight of importance for each criterion using an MCDM method. The triangular fuzzy preference numbers elicited from the CI survey responses must be processed through a fuzzy MCDM method in order to assess the relative importance of the SSC. Therefore, the FAHP method is applied in a manner similar to Perçin and Aldalou (2018), resulting in the weights of the SSCs shown in Table 3. W_I, W_{FoA}, and W_{Ctrl} refer to the weights of I, FoA, and Ctrl, respectively.

Table 3. SSC weights.

	W_I	W_{FoA}	W_{Ctrl}
Weight (0,1)	0.222	0.287	0.491

Step 2: Assess the consistency of the respondents' pairwise comparisons in the CI survey. This is done by calculating the consistency ratio (CR) of the matrix ($\widetilde{A}$), which includes the fuzzy preference numbers determined in step 1. First, the matrix $\widetilde{A}$ is defuzzified into two crisp matrices; the first matrix (A_1) includes the most likely values of the fuzzy numbers of the matrix $\widetilde{A}$ and the second matrix (A_2) includes the geometric mean of the lower and upper bounds. Then, based on the approach used by Saaty (1980), the CRs for matrices A_1 and A_2 are evaluated ($CR_{A_1} = 0.0030$ and $CR_{A_2} = 0.0047$). Since they are less than 0.1, no re-examination of the pairwise judgments of the CI surveys is required.

Step 3: Aggregate the respondents' opinions, elicited through the FI survey, by assessing the aggregated responses (ARs) (i.e., the levels of I, FoA and Ctrl) for each CLP factor. Eqs. (5), (6), and (7) compute AR with respect to the criteria I, FoA and Ctrl, respectively, where the maximum possible AR of the equations is 1.

$$AR_I = \frac{\sum_{i=1}^{5} i \times A_i}{5} \tag{5}$$

$$AR_{FoA} = \frac{\sum_{i=1}^{5} i \times B_i}{5} \tag{6}$$

$$AR_{Ctrl} = \frac{\sum_{i=1}^{5} i \times C_i}{5} \tag{7}$$

where A_i, B_i and C_i are the percentages of respondents in the FI survey who rated a particular factor as i in terms of I, FoA, and Ctrl, respectively. Table 4 shows sample ARs of 155 CLP factors derived from Tsehayae and Fayek (2014).

Step 4: Apply the relative importance of the SSCs for ranking CLP factors. This is done by calculating the weighted AR of each CLP factor by multiplying the aggregated responses from Table 4 with the corresponding SSC weights as computed in step 1. Table 5 shows sample weighted ARs with respect to the criteria I, FoA, and Ctrl for each CLP factor.

Table 4. Aggregated CLP data.

CLP factor no.	CLP factor	AR_I	AR_{FoA}	AR_{Ctrl}
1	Power equipment breakdown	0.4450	0.5721	0.9890
2	Instability of political system	0.1270	0.6272	0.0813
⋮	⋮	⋮	⋮	⋮
155	Lack of protection from weather effect	0.7928	0.7080	0.2037

Table 5. Weighted aggregated CLP data.

CLP factor no.	$W_I \times AR_I$	$W_{FoA} \times AR_{FoA}$	$W_{Ctrl} \times AR_{Ctrl}$
1	0.0988	0.1642	0.4915
2	0.0282	0.1800	0.0404
⋮	⋮	⋮	⋮
155	0.1760	0.2032	0.1012

Phase 4: Data analysis

In the last phase, a fuzzy data clustering technique and an MCDM method are used to cluster and rank the CLP factors, as explained below.

First, due to the overlapping nature of cluster boundaries, FCM is employed to partition the CLP factors (i.e., the data points) into clusters based on their similarities and dissimilarities in terms of the SSCs. The CLP factors are clustered in order to determine a set of CLP factors that are considerably more important than other factors for improving CLP. The identified set of key CLP factors assists construction companies with identifying a set of CLP improvement strategies that will have the most positive influence on CLP. Fuzzy partitioning is conducted through minimizing the objective function (Eq. 6) by the iterative update of the data points' memberships (u_{ij}^m) and the cluster centers (c_j) (Nayak et al. 2015).

$$J_m = \sum_{i=1}^{n} \sum_{j=1}^{l} u_{ij}^m d(x_i, c_j)^2 \qquad (6)$$

where n and l are the number of CLP factors and clusters, respectively. The parameter that controls the fuzziness of the clusters is $m \in (1, +\infty)$ and $d(x_i, c_j)$ is the Euclidean distance from the data point x_i (i.e., the i_{th} CLP factor) to the cluster center c_j. Shown below, u_{ij}^m is the degree to which the i_{th} CLP factor belongs to the j_{th} cluster.

$$u_{ij}^m = \frac{1}{\sum_{k=1}^{c} \left(\frac{d(x_i, c_j)}{d(x_i, c_k)}\right)^{\frac{2}{m-1}}} \qquad (7)$$

Cluster centers are initialized randomly and calculated as follows in the next iterations.

$$c_j = \frac{\sum_{i=1}^{n} u_{ij}^m . x_i}{\sum_{i=1}^{n} u_{ij}^m} \qquad (8)$$

Based on c_j, the value of u_{ij}^m for all i and j are computed and the iterations between these two equations are repeated until the minimum J_m or the following condition is achieved.

$$d\left(C^{(b)}, C^{(b+1)}\right) < \varepsilon \qquad (9)$$

where $C^{(b)}$ is the cluster center matrix in the b_{th} iteration step and ε is the predetermined level of accuracy. As a result of FCM, the CLP factors (i.e., the data points) are divided into ten clusters with their corresponding centers in three dimensions (C_I, C_{FoA}, and C_{Ctrl}), as presented in Table 6. The center of each cluster is the mean importance of its data points. For instance, cluster 1 includes 18 CLP factors with a mean importance of 0.1047, 0.2493, and 0.2663 for the criteria I, FoA and Ctrl, respectively.

Table 6. CLP final clusters.

Cluster no.	Number of CLP factors	Cluster center		
		C_I	C_{FoA}	C_{Ctrl}
1	18	0.1047	0.2493	0.2663
2	18	0.1790	0.2234	0.3682
3	13	0.1067	0.1702	0.4390
4	18	0.0553	0.1784	0.1113
5	13	0.1423	0.0695	0.2872
6	17	0.0828	0.0501	0.0587
7	15	0.1652	0.0689	0.1056
8	16	0.0590	0.0503	0.3899
9	8	0.0903	0.1706	0.2050
10	19	0.1345	0.2171	0.0692

Second, the MCDM method used for ranking the CLP clusters obtained through the FCM technique is TOPSIS. This method identifies the CLP cluster that has the shortest distance to the positive-ideal solution (A^*) and the longest distance from the negative-ideal solution (A^-). The positive-ideal solution consists of the highest values (i.e., 0.1790 for C_I, 0.2493 for C_{FoA}, and 0.4390 for C_{Ctrl}) of SSC among CLP clusters, and the negative-ideal solution includes the lowest values (i.e., 0.0553 for C_I, 0.0501 for C_{FoA}, and 0.0587 for C_{Ctrl}). Hence, by considering the distances from the positive-ideal solution (S_i^+) and negative-ideal solution (S_i^-), the ranks of the clusters are calculated as follows:

$$RC_i = \frac{S_i^-}{S_i^+ + S_i^-} \qquad (10)$$

where RC_i is the relative closeness of the i_{th} CLP cluster to the positive-ideal solution. By implementing all the steps of the TOPSIS method, the relative closeness and, consequently, the rank of each CLP cluster is determined, as shown in Table 7. The highest value of relative closeness belongs to cluster 2, which consists of 18 CLP factors: (1) crew experience, (2) cooperation between craftsmen, (3) craftsman learning speed, (4) job site orientation program, (5) remuneration, (6) shortage of consumables, (7) material order tracking system, (8) waiting time for manlifts, (9) quality of work tools, (10) rework sources, (11) cleanliness of work area, (12) work conditions (noise, dust, and fumes), (13) foreman experience, (14) foreman skill, (15) adequacy of job instructions, (16) health and safety training, (17) materials management practices, and (18) zero accident techniques.

These factors have the highest importance for CLP improvement. This result will help construction companies implement a set of CLP improvement strategies in order to improve the identified key factors. For example, for foreman experience and foreman skill, which are among the identified CLP factors in the first ranked CLP cluster, one CLP improvement strategy would be to implement training programs to improve the experience and skills of foremen.

Table 7. Cluster ranks.

Cluster no.	RC	Rank
1	0.6082	3
2	0.8328	1
3	0.7895	2
4	0.2797	8
5	0.5073	5
6	0.0587	10
7	0.2419	9
8	0.5825	4
9	0.4236	6
10	0.3312	7

CONCLUSIONS AND FUTURE RESEARCH

Budget and resource restrictions force construction organizations to prioritize CLP improvement strategies according to their effect on CLP factors. Hence, it is necessary to rank CLP factors based on their importance for CLP to help construction organizations identify the most effective CLP improvement strategies. This study proposed a hybrid model for ranking CLP factors using the integration of fuzzy data mining and MCDM methods. Two questionnaires were designed to measure the weights of SSCs based on experts' opinions and collect experts' opinions about each CLP factor with respect to the selection criteria. The first contribution of this paper is defining a new criterion, controllability, which influences the selection of CLP improvement strategies. The second contribution is using a fuzzy MCDM method for aggregating SSCs for CLP improvement strategies. The third contribution of this paper is presenting a framework of MCDM and fuzzy data clustering for ranking CLP factors based on their importance for CLP improvement. This importance is measured in terms of three criteria that influence the selection of CLP improvement strategies. The last contribution is providing a systematic approach for analyzing and selecting CLP improvement strategies by identifying the most effective CLP factors. The results of this paper will help construction organizations identify key CLP factors and implement a set of CLP improvement strategies in order to improve the identified key factors. The findings of this work provide a basis for future research, including ranking CLP factors with actual collected data and using a fuzzy MCDM method such as fuzzy TOPSIS.

REFERENCES

Alaghbari, W., Al-Sakkaf, A. A., and Sultan, B. (2019). "Factors affecting construction labour productivity in Yemen." *Int. J. Constr. Manage.,* 19(1), 79–91.

Gerami Seresht, N., and Fayek, A. R. (2019). "Factors influencing multifactor productivity of equipment-intensive activities." *Int. J. Product. Perform. Manage.*, ahead-of-print.

Ghoddousi, P., Poorafshar, O., Chileshe, N., and Hosseini, M. R. (2015). "Labour productivity in Iranian construction projects: Perceptions of chief executive officers." *Int. J. Product. Perform. Manage.,*, 64(6), 811–830.

Ghodrati, N., Wing Yiu, T., Wilkinson, S., and Shahbazpour, M. (2018). "Role of management strategies in improving labor productivity in general construction projects in New Zealand: Managerial perspective." *J. Manage. Eng.,* 34(6), 04018035.

Heravi, G., and Eslamdoost, E. (2015). “Applying artificial neural networks for measuring and predicting construction-labor productivity.” *J. Constr. Eng. Manage.,* 141(10), 04015032.

Kedir, N. S., Raoufi, M., and Fayek, A. R. (2019). “Integrating fuzzy agent-based modeling and multi-criteria decision-making for analyzing construction crew performance.” *Proc., Int. Conf. Comput. Civ. Eng. 2019: Vis., Inform. Mod., and Sim.,* ASCE, Reston, VA, 569–576.

Nasir, H., Haas, C. T., Caldas, C. H., and Goodrum, P. M. (2015). “An integrated productivity-practices implementation index for planning the execution of infrastructure projects.” *J. Infrastruct. Syst.,* 22(2), 04015022.

Nayak, J., Naik, B., and Behera, H. (2015). “Fuzzy c-means (FCM) clustering algorithm: A decade review from 2000 to 2014.” H. M. Behera and D. P. Mohapatra (Eds.), *Computational Intelligence in Data Mining–Volume 2,* Springer, 133–149.

Perçin, S., and Aldalou, E. (2018). “Financial performance evaluation of Turkish airline companies using integrated fuzzy AHP fuzzy TOPSIS model.” *Uluslararası Iktisadi Ve Idari Incelemeler Dergisi,(18.EYİ Özel Sayısı),* 583–598.

Raoufi, M., and Fayek, A. R. (2018). “Fuzzy agent-based modeling of construction crew motivation and performance.” *J. Comput. Civ. Eng.*, 32(5), 04018035.

Saaty, T. L. (1980). The analytic hierarchy process for decision in a complex world. Pittsburgh: RWS Publications.

Tsehayae, A. A., and Fayek, A. R. (2016). “Developing and optimizing context-specific fuzzy inference system-based construction labor productivity models.” *J. Constr. Eng. Manage.,* 142(7), 04016017.

Tsehayae, A. A., and Fayek, A. R. (2014). “Identification and comparative analysis of key parameters influencing construction labour productivity in building and industrial projects.” *Can. J. Civ. Eng.,* 41(10), 878–891.