Canadian Journal of Civil Engineering

Performance Prediction of Construction Projects using Soft Computing Methods

Journal:	<i>Canadian Journal of Civil Engineering</i>
Manuscript ID	cjce-2018-0305.R1
Manuscript Type:	Article
Date Submitted by the Author:	16-Oct-2018
Complete List of Authors:	Fanaei, Seyedeh Sara; Concordia University, Dept. of Building, Civil and Environmental Engineering Moselhi, Osama; Concordia University, Dept. of Building, Civil and Environmental Engineering Alkass, Sabah; Concordia University, Dept. of Building, Civil and Environmental Engineering
Keyword:	Key Performance Indicators (KPIs), Neuro-Fuzzy, Performance Forecasting, Construction Project
Is the invited manuscript for consideration in a Special Issue? :	Not applicable (regular submission)
Note: The following files were submitted by the author for peer review, but cannot be converted to PDF. You must view these files (e.g. movies) online.	
Fig 1.vsd Fig 2.vsd Fig 4.vsd Fig 5.vsd	

SCHOLARONE™
Manuscripts

Performance Prediction of Construction Projects using Soft Computing Methods

Seyedeh Sara Fanaei¹, Osama Moselhi, Ph.D., P.Eng.², Sabah T. Alkass, Ph.D.³

¹ Ph.D. Candidate, Dept. of Building, Civil and Environmental Engineering, Concordia Univ., Montreal, QC, Canada H3G 1M8 (corresponding author). E-mail: s_fanaei@encs.concordia.ca

² Professor, Dept. of Building, Civil and Environmental Engineering, Concordia Univ., Montreal, QC, Canada H3G 1M8.

³ Professor, Dept. of Building, Civil and Environmental Engineering, Concordia Univ., Montreal, QC, Canada H3G 1M8.

Word count: 7700 words

ABSTRACT

Key Performance Indicators (KPIs) evaluate different aspects of projects and are used as a thermometer to determine the health status of projects. While there is considerable work on project quantitative performance prediction, less attention, however, has been directed towards qualitative performance prediction. This paper offers a novel framework for qualitatively measuring and predicting six important construction project KPIs using the neuro-fuzzy technique. Neuro-fuzzy models are developed to map the KPIs of three critical project stages to whole project KPIs. Subtractive clustering is utilized to automatically generate initial Fuzzy Inference System (FIS) models and the artificial neural network (ANN) technique is used to tune the parameters of the initial FIS models. The relative weight of each KPI is determined using a series of computing methods namely, Analytic Hierarchy Process (AHP) and Genetic Algorithm (GA), to generate the Performance Indicator (PI). The developed models are validated with real project data showing that the rate of error is reasonably low. The results show that the AHP

method is more accurate when compared to the GA method. This framework can be used in building construction projects to help decision-makers evaluate the performance of their projects.

Keywords: Key Performance Indicators (KPIs), Neuro-Fuzzy, Performance Forecasting, Construction Project

INTRODUCTION

Performance is well defined as the amount of efficiency and effectiveness in all project objectives (Nassar and AbouRizk 2014). Key Performance Indicators (KPIs) are a benchmark for assessing project performance. Dynamic project performance measurement refers to measuring the performance of a project when a project is in progress. In the construction industry, project performance prediction is vital for monitoring and controlling construction projects. In the Egan (1998) Report titled Rethinking Construction and published in 1998, Egan launched the KPI method for measuring performance. For a couple of years after, time and cost were the only KPIs used; nowadays, additional KPIs are needed to be recognized for understanding the true status of a project (Kerzner 2011). Defense Construction Canada (DCC) provides guidance on the development, measurement, and reporting of KPIs. DCC uses KPIs to measure the success of Canadian projects (DCC 2016). Furthermore, the Construction Industry Institute (CII) has introduced a performance assessment program for evaluating project performance in the United States (US) using KPIs (CII 2014).

Construction project KPIs have been measured and forecasted both quantitatively and qualitatively in the literature. Chua et al. (1999) assessed project achievement through three objectives: cost, schedule, and quality. Cha and Kim (2011) defined a quantitative performance measurement system for residential building projects using eighteen KPIs developed by a mathematical model. Nassar and AbouRizk (2014) used a quantitative approach to measure

project performance from the contractor's point of view during the construction phase of projects. KPIs were decided based on experts' opinion. Total project performance was calculated by summing up eight performance indicators with their weight, where the AHP (Analytical Hierarchy Process) method was used for calculating weights (w). Hanna et al. (2014) used a mathematical formula for the calculation of a performance metric that quantitatively assesses cumulative project performance using important key performance metrics. Reenu et al. (2017) forecasted project performance by developing an Artificial Neural Network (ANN) model and using four performance metrics: prediction of cost performance, schedule performance, quality performance, and satisfaction level. Leon et al. (2017) developed a system dynamics (SD) model for forecasting project performance using eight construction indicators.

While there is considerable work on quantitative performance prediction of the projects, less attention, however, has been directed towards measuring performance qualitatively. Examples of qualitative methods for measuring project performance are as follows: Dissanayake and Fayek (2008) developed a model for monitoring performance and identifying the causes of performance failures for Canadian projects using fuzzy, ANN, and Genetic Algorithm (GA) methods. Their proposed approach was daily based and for individual activities and did not consider the whole project level. Haponava and Al-Jibouri (2012) designed a system for measuring process performance in three main project stages. Their proposed measurement system considered a number of process-based KPIs. However, they did not consider the link between processes and overall performance. Wester (2013) recognized five KPIs in the design stage for advanced high-technological construction projects using a qualitative approach. However, their research only focused on the design stage of projects and did not consider other stages.

More recently, Ngacho and Das (2015) developed a construction project performance assessment framework based on six KPIs identified through interviews and a literature review. They used a conceptual diagram to provide the relationship between critical success factors and overall project performance. However, predicting project performance was not part of their framework. Nilashi et al. (2015a) applied AHP and fuzzy logic to develop an expert system for assessing the performance level of green buildings. This model did not evaluate performance during the construction phase. Nilashi et al. (2015b) focused on finding factor importance and calculating weight as well as the interdependencies among selected criteria. They did not provide an approach for predicting KPIs based on current project status. Wan et al. (2017) used KPIs to forecast the performance of E-Commerce companies to indicate progress and confirm business goal attainment. The main objective was to enhance the forecasting of E-Commerce companies' KPIs using past data and their model did not consider construction projects. Shaikh and Darade (2017) focused on quality of activities by considering key indicators in the planning stage. The study found key indicators of an activity and their contribution to improve activity quality by using a Project Quality Plan (PQP) document. However, their framework did not predict performance and only focused on quality indicators without considering other KPIs.

Problem Statements

Based on the literature review, the following limitations emerge. First, most of the literature focuses on quantitative project performance forecasting and less on qualitative methods. However, some construction KPIs such as client satisfaction, quality, and safety are qualitative in nature and cannot be measured quantitatively. Second, most of the studies measure performance only after project completion and not during the construction phase (Haponava and Al-Jibouri 2012). The benefit of measuring performance during the project is that stakeholders

can use the measurements to propose corrective actions and forecast remaining project performance. In addition, comparing neuro-fuzzy, ANN, and multiple regression methods for prediction has been done for other domains. The literature demonstrates that the neuro-fuzzy technique is a more precise prediction method as compared to other methods (Esmaeili et al. 2014; Paul et al. 2015). In the construction area, Nilashi et al. (2017) applied machine learning techniques to develop a hybrid intelligent system for predicting heating and cooling loads of residential buildings. Adaptive Neuro-Fuzzy Inference System was compared with other techniques for predicting buildings' energy performance. The results indicate better prediction accuracy when using neuro-fuzzy. However, neuro-fuzzy has not been previously utilized for predicting construction performance indicators.

Research Contributions

This research offers a framework for qualitatively measuring and forecasting construction project performance. The most important KPIs are identified using two methods: an extensive review of the literature, and questioners distributed to expert. The proposed framework utilizes the qualitative data of KPIs during the critical stages of construction projects to develop neuro-fuzzy models for predicting overall project KPIs. In the proposed approach, the clusters radiuses are optimized in the neuro-fuzzy model with subtractive clustering. GA and AHP are used to calculate the weight of each KPI for calculating final project performance. The model is verified and validated using real construction project data.

RESEARCH METHODOLOGY

A framework is proposed for measuring and forecasting a set of qualitative KPIs using soft computing approaches. The proposed framework includes the following steps, shown in Fig. 1. The first step to measuring project performance is defining project objectives and goals.

Because of different stakeholder investments and benefits, it is important to first define from whose viewpoint project performance is measured. Then, the main project KPIs should be identified through both literature review and expert opinion. Afterward, how each KPI will be measured is defined. In this research, a qualitative method using a 1 to 7 scale based on a method proposed by Dissanayake and Fayek (2008) is used for measuring KPIs. The linguistic terms used for this scale are 1=Very Low, 2=Low, 3=Medium-Low, 4=Medium, 5=Medium-High, 6=High, 7=Very High. This is done since many construction KPIs such as client satisfaction, quality, and safety are qualitative in nature and cannot be measured quantitatively. Meanwhile, for other KPIs such as time and cost, confidentiality issues and lack of data that construction projects usually face complicating quantification. The KPIs are defined for three critical project stages: initial, middle, and finishing stage.

Prediction models are developed using the neuro-fuzzy technique for forecasting whole project KPIs. The proposed framework integrates the whole project KPIs to get the overall project Performance Indicator (PI) using a mathematical calculation shown in Equation 1, proposed by Nassar and AbouRizk (2014) and Cha and Kim (2011). Fig. 2 illustrates the components of the proposed framework which includes the KPI prediction models using neuro-fuzzy technique and PI prediction model.

$$PI = w_1KPI_1 + w_2KPI_2 + w_3KPI_3 + w_4KPI_4 + w_5KPI_5 + w_6KPI_6 \quad (1)$$

KPI_i : performance indicators w_i : weight of each indicator

Two different methods are used and compared to determine the weight of each indicator: AHP and GA. In AHP, the relative importance of each indicator is determined based on pair-wise comparisons. Then, actual data of selected KPIs is collected. This data serves as the model input.

After data is collected, the results of the model and the original plan should be compared. If a major deviation is found, the analysis should be performed to get the cause of the deviation and determine corrective actions for increasing project performance. Thus, this framework enables managing the project's performance by predicting project KPIs during three critical stages.

KPIs Identification

Table 1 presents a list of the KPIs used in the literature at the project level. The first six KPIs were chosen based on the frequency of use; these KPIs were used in approximately 50% or more of the reported studies. Furthermore, a questionnaire is designed for evaluating the importance of each of the KPIs based on expert opinion. The questionnaires were distributed to 34 owners of building construction projects who were asked to provide a score between 1 to 7 for each of the KPIs. Fig. 3 demonstrates the total score for each of the KPIs from the questionnaire. Based on the questionnaire outcomes, the first six KPIs got 40% to 90% of the score, with the seventh KPI receiving a score of 27, and the remaining KPIs receiving a score of 20% to 27%. Given the tangibly larger score ratio between KPI number six and seven in the questionnaire, it is fair to conclude that the experts found the first six KPIs more important than the others. Based on the literature review and the questionnaire, this study decided to select the first six KPIs: cost, time, quality, safety, client satisfaction, and project team satisfaction to develop the model. The selected KPIs comply with previous Canadian studies (DCC 2016; Dissanayake and Fayek 2008; Nassar and AbouRizk 2014). These six KPIs are used to design the questionnaire for data collection. All indicators are measured considering the profits and damages to owner qualitatively.

DATA COLLECTION

To collect data on building construction projects, another questionnaire was designed and distributed to owners of building construction projects in Tehran, Iran. The experts were asked to qualitatively rank the performance of six selected KPIs in different construction projects. The questionnaire is used to collect project data in three critical stages: initial stage (0 to 30% project progress), middle stage (30% to 70% project progress), and finishing stage (70% to 100% project progress), as well as whole project KPIs and overall project PI. PI represents the overall health status of the projects. The data is collected using a 1 to 7 scale, where 1=Very Low, 2=Low, 3=Medium-Low, 4=Medium, 5=Medium-High, 6=High, 7=Very High (Dissanayake and Fayek 2008). Using quantitative scales allows soft computing methods to be applied to qualitative data. Two-hundred questionnaires were distributed to experts of which 119 with consistent results were selected for analysis. The questionnaires are used to collect the values of six KPIs at different stages. Table 2 shows the statistical descriptions of all parameters included in the questionnaire database. The quantities collected from the questionnaires are then used to train and develop a neuro-fuzzy model for predicting whole project KPIs and apply the AHP and GA methods for calculating project PI.

MODEL DEVELOPMENT

In the proposed framework, the neuro-fuzzy technique is applied to forecast the whole project KPIs automatically from data. The neuro-fuzzy technique is a combination of ANN and fuzzy logic used to resolve different research problems in construction management. Furthermore, an advantage of the neuro-fuzzy technique is that it does not assume a pre-defined mathematical expression. In addition, it also captures each variable's effect on the output without requiring a priori knowledge (Jang 1997).

182 This research generates a Sugeno-type FIS structure using a subtractive clustering method.
183 Gaussian type of membership function is used. MATLAB software is used to develop the
184 models. One of the limitations of using the neuro-fuzzy technique in MATLAB is that it provides
185 only one output (Jang 1997). However, this limitation does not impact the functionality of the
186 developed models by developing more prediction models. Multiple neuro-fuzzy models are
187 developed to predict each whole project KPI for three critical project stages. Six different neuro-
188 fuzzy models are developed for the initial stage. Each model in this stage has six inputs which
189 are the six KPIs of initial stage and one output. The output of each model is each of the six KPIs
190 of the whole project. For each middle and finishing stage, six different neuro-fuzzy models have
191 been developed as well, resulting in a total of 18 neuro-fuzzy models, as shown in Fig. 4.

192 The collected KPIs data as described in the previous section are used to develop the prediction
193 models. The inputs of the neuro-fuzzy model are the values of the six KPIs in three different
194 stages of the 119 responses to the questionnaires. Each KPI amount is a value obtained from the
195 questionnaires qualitatively (using a 1 to 7 scale). The data is divided into two groups: train and
196 test. The train data is used to train the models, while the test data is used to control the results.
197 The train dataset contains the experimental data utilized for tuning the FIS model parameters
198 during the training stage, whereas the test dataset includes new data that has not yet been
199 introduced to the FIS model.

200 Shahin et al. (2004) demonstrated that considering the statistical consistency between train and
201 test data can improve the performance of the prediction model. They concluded that considering
202 the statistical parameters between test and train data should be part of the data division processes.
203 Therefore, in this research, data is divided to test and train datasets to ensure statistical
204 parameters are as close to each other as possible. The mean, standard deviation, minimum, and

maximum are statistical parameters that are compared using trial and error to achieve this objective. The train dataset contains 70% of the collected data and the remaining 30% are used for the test dataset.

Development of neuro-fuzzy models consists of two main steps: the development of an initial Fuzzy Inference System (FIS) and the optimization of the initial FIS model using the ANN technique. To develop the initial FIS model, subtractive clustering is utilized. Subtractive clustering has been proposed as a reliable and precise method for developing prediction models. Subtractive clustering can be used to extract cluster centers that represent the FIS model (Chiu 1994). In subtractive clustering, the radius chosen for developing clusters impacts the number of fuzzy rules. A bigger radius results in a smaller number of fuzzy rules. On the other hand, a smaller radius results in a higher number of fuzzy rules but increases the chance of overfitting. Overfitting means a model corresponds too closely to the training dataset but can not precisely predict future data. Therefore, the cluster radius should be optimized to achieve optimum precision without overfitting the train dataset (Nasrollahzadeh and Basiri 2014).

In this research, the cluster radius of each FIS model is optimized by changing the cluster radius from 0 to 1 (the acceptable range in subtractive clustering) as proposed by Nasrollahzadeh and Basiri (2014). For each radius value, a FIS model is generated, resulting in various FIS models. The errors of the developed FIS models are measured for each cluster radius in two separate groups: train and test datasets. Several error measures including Mean Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE), and Coefficient of Variation (COV) are calculated between the model results of train and test datasets.

A model is optimum when the errors (MAPE, RMSE, and COV) calculated for the test dataset are at their lowest but also as close as possible to the train dataset. This approach ensures the

generalization capability of the model and prevents the problem of overfitting (Nasrollahzadeh and Basiri 2014). When two FIS models perform similarly regarding error measures, the simpler model with the smaller number of rules (smaller cluster radius) is preferred. This is because having more number of rules/clusters may result in less model generalization capability (Nasrollahzadeh and Basiri, 2014).

The above approach of changing radius values is used to select the best initial FIS models. The neuro-fuzzy technique is then utilized to optimize the initial FIS models' parameters to reduce model error as much as possible. The neuro-fuzzy inference system develops a FIS whose membership function parameters are tuned using a backpropagation algorithm in combination with the least squares method. This tuning allows the FIS model to learn from the data it is modeling (Jang 1997).

Depending on the stage of the project, initial stage (0 to 30% project progress), middle stage (30% to 70% project progress) and finishing stage (70% to 100% project progress), the developed models for initial, middle, or finishing stage can be utilized to calculate the whole project KPIs. For example, if the project is 45% in progress, the developed models for the middle stage can predict the KPIs.

The developed framework is implemented in MATLAB software to automatically generate 18 neuro-fuzzy models for predicting six projects KPIs in three critical stages. Codes are written to generate the models and the MATLAB toolbox is not used exclusively. The reasons for writing the codes are that firstly, this allows for the cluster radius amount to be optimized since by using the toolbox, the radius needs to be given by the user. Secondly, codes are written to perform data division into train and test based on the statistical characteristics of data. The flowchart presented in Fig. 5 summarizes the proposed framework for developing neuro-fuzzy models to predict

whole Project KPIs. A user interface able to get project information and predict the whole project KPIs is also developed.

Finding the weights of KPIs to calculate PI

After predicting project KPIs using the neuro-fuzzy models developed in the previous section, the project PI is calculated using Equation 1. In this section, the weights of each KPI in Equation 1 are calculated using both AHP and GA methods.

AHP Method: The AHP method is a methodology proposed for multi-criteria decision making (Saaty 1988). The method is able to derive priorities by ranking different alternatives to a problem. AHP uses a pairwise comparison matrix to rank different criteria and to evaluate the relative performance of alternatives. The consistency ratio (CR) of each of the questionnaires is calculated according to the Saaty (1988) method. From the 119 questionnaires, 35 questionnaires had an acceptable consistency ratio of 10% or less (Saaty and Vargas 1982). The pair-wise comparison matrix is developed using consistent questionnaires. These steps were adopted from Saaty (1988). For every questionnaire, the pair-wise comparison matrix is developed and then normalized. The resulting weight of each KPI can be seen in Table 3. Each cell in the pair-wise comparison matrix in Table 3 is a geometric mean of the 35 pair-wise comparison matrixes that are developed earlier. In the normalized matrix, each cell amount is calculated by dividing cell amounts by the sum of their columns. For defining the weight of each KPIs in Table 3, the average of each row is calculated.

GA Method: GA is a global optimization technique proposed by Holland (Holland 1975) suitable for complex and nonlinear problems. The main mechanism in GA is based on natural selection. GA searches a population of potential solutions and keeps the best solutions for the

generation of the next possible solutions. The GA method is used to optimize the weight of each KPI in Equation 1 by using the dataset collected from the questionnaire. GA toolbox in MATLAB environment is used for this purpose. The parameters to be optimized are the weights of KPIs in Equation 1 (W_1 to W_6), with the objective to optimize the fitness function value. The fitness function is defined as the RMSE error between the questionnaire PI and the overall PI calculated using Equation 1. The constraints were that each KPI weight should be positive and its sum equal to 1. Variables were the PI and KPIs amount obtained from the questionnaire. KPIs weights optimized by the GA method are shown in Table 4.

Sensitivity Analysis of Whole Project KPIs

A sensitivity analysis is carried out to evaluate the effect of the six KPIs of three project stages on whole project KPIs. For this purpose, the parameter under analysis is varied within its range of values, while other input parameters are kept constant (the average amount of collected questionnaire data). As shown in Fig. 6, the cost indicator value of each stage is changed to generate the impact on the whole project cost indicator. In the initial and middle stage, when the cost performance of the project increases, the amount of the whole project cost performance increases as well. However, in the finishing stage, when the cost indicator performance increases, the whole project cost performance increases slightly less. It can be seen that an increase in the cost performance indicator in the initial and middle stages will have more of an effect on the whole project cost performance. In the middle and finishing stage when the time indicator increases, the whole project time indicator increases as well, as shown in Figure 6. However, in the initial stage, when the time indicator increases from one to four, there is an increase in the whole project time indicator. By increasing the time indicator from four to seven, the whole project time indicator stays constant. This analysis was performed for the other KPIs

as well, as shown in Fig. 6. As can be seen below, whole project KPIs are improved by improving the amount of each KPI in the three critical project stages.

Comparing the model and questionnaire PIs

In this part, a comparison between the model PI and the PI amounts collected by the questionnaire is performed. The model PI is calculated by both AHP and GA methods using the KPIs predicted by neuro-fuzzy models. Fig. 7 illustrates the comparison of the calculated PI with the PI obtained from the questionnaires in three project stages. When the dots are closer to the diagonal line, the prediction is more accurate. Additionally, the coefficient of determination R^2 is provided in each of the diagrams in Figure 7. R^2 is a statistical analysis method; an R^2 value of near 1 means very good fit while a value closer to 0 means poor fit (Dissanayake and Fayek 2008).

As illustrated in Fig. 7, the data scattered in the AHP method is closer to the diagonal line as compared to the GA method. The R^2 values are closer to 1 using the AHP method; thus, the AHP method performs slightly better as compared to the GA method. MAPE for the GA method in the finishing stage is 2%, while the MAPE for the AHP method in the same stage is 1.8%. A comparison between the questionnaire results and the neuro-fuzzy model predictions indicates that the data scatterings around the diagonal line are very similar in each stage. The comparison shows that the accuracy of the proposed neuro-fuzzy models is quite satisfactory as the model predictions agree with the questionnaire data. In the finishing stage, using the AHP method, MAPE is 1.8%, while it is 2% and 4% in the middle and initial stages, respectively. Also, the data scattered in the finishing stage is closer to the diagonal line as compared to middle and initial stages and the R^2 value increases from the initial stage to the finishing stage. In other words, as one gets closer to the end of the project, the accuracy of predicting the project PI

improves. In Dissanayake and Fayek's (2008) results, the R^2 for predicting productivity using KPIs is equal to 0.95. In this research, the R^2 value in the finishing stage for AHP method is 0.93, which is very close to their reported value, showing the reasonable accuracy of the results.

Validation

To validate the model, information from 16 real projects is used. The PI results from the model are compared with the PI obtained from real projects not used in generating the model. The PIs are calculated by both the AHP and GA methods using the KPIs derived from the model and the KPIs derived from the real projects. Different error calculation methods (COV, RMSE, and MAPE) are used to evaluate the results. The results of the three stages can be seen in Table 5. By comparing the GA and AHP method, the errors illustrate that the model works better with the AHP method.

Table 5 shows that the MAPE for the AHP method is between 10% and 20% in all three stages; according to Jia et al. (2015), this forecasted accuracy is reasonable. As illustrated in Fig. 8, most data points in the three stages are close to the diagonal line. Thus, the PIs calculated with the developed model agree with the real project PIs, confirming the generalization ability of the proposed framework. Moving from the initial and middle stages to the finishing stage, data points move closer to the diagonal line. This behavior is reasonable since, in the finishing stage, the model prediction ability is expected to be higher. Based on the attained results, the neuro-fuzzy approach is successfully applied for modeling the impact of six important KPIs on performance indicators of construction projects.

DISCUSSION

The proposed framework is analyzed by performing a sensitivity analysis that compared the PI of the models and questionnaires. This is further validated using real project data. The

results of the comparison between the model and questionnaires indicate the AHP method performs better in calculating KPI weights as compared to the GA method. Validation results based on 16 new project information indicate reasonable accuracy and generalization ability of the developed models.

CONCLUSION AND FUTURE WORK

This paper proposes a new framework for qualitatively measuring and predicting construction projects KPIs using the neuro-fuzzy technique. Six important KPIs (cost, time, quality, safety, client satisfaction, and project team satisfaction) are selected based on literature and expert opinion. The KPIs are measured qualitatively using a 1 to 7 scale. The KPIs of three critical project stages are mapped to the whole project KPIs through neuro-fuzzy models. For this purpose, 18 different neuro-fuzzy models are developed, six models for each of the three critical project stages. Neuro-fuzzy models are developed in two steps; first, initial FIS models are developed using subtractive clustering, where the cluster radius is optimized to achieve optimum precision without overfitting. Second, the optimization of the initial FIS model is performed using ANN. The weighted sum of the project KPIs is then used to calculate the overall project PI. The weight of each KPI is calculated using both AHP and GA methods. The resulting weights from these two methods are then compared to find the best method for calculating KPI weights.

The proposed framework is applied to building construction project data collected from 119 questionnaires. The developed models' analysis using collected data is performed using two procedures. First, a sensitivity analysis is carried out to evaluate the effect of the six KPIs of the three stages on whole project KPIs indicating the whole project KPIs are enhanced by improving the amount of each KPI in three project stages. Second, a comparison between the model PI

365 (using both AHP and GA methods) and the PI amounts collected by questionnaire is performed.
366 The results indicate that the AHP method performs better over the GA method. Furthermore, PIs
367 resulted from the questionnaire and PIs calculated by the model are very similar. To validate the
368 model, information from 16 real projects is used. The PIs calculated with the developed model
369 agree with the real project PIs showing that the rate of error is reasonably low.

370 The main contribution of this research is proposing a framework for forecasting construction
371 KPIs qualitatively. The proposed framework utilizes qualitative data of KPIs during the critical
372 stages of construction to develop a neuro-fuzzy model for predicting whole project KPIs.
373 Development of a methodology to optimize the clusters radiuses in developing the neuro-fuzzy
374 model with subtractive clustering is also provided.

375 A user interface is developed to get new information in each stage of the project and the model
376 can predict the whole project KPIs, as well as the overall project PI. The practical implication of
377 the proposed framework is that project stakeholders can predict project KPIs allowing them to
378 forecast the status of future work and perform necessary corrective actions. This can help to
379 improve project performance. The developed model assists owners and their agents to enhance
380 progress reporting of construction projects.

381 One of the limitations of the developed framework is that the user needs to manually enter the
382 data for six KPIs qualitatively. In the future, the model can be improved to become more user-
383 friendly by automatically deriving KPI values as model input at any stage of project based on
384 project information. Furthermore, future studies should look to connect the developed models to
385 the internet to provide an online prediction.

The developed model is based on the owners' point of view and the viewpoint of other stakeholders is beyond the scope of this study. In the future, the framework should be adapted to the viewpoint of other stakeholders. The proposed framework for developing a performance prediction model can be applied to data collection in other locations such as Canada or for other types of projects. The model can be improved by proposing corrective actions to project managers to improve project performance.

REFERENCES

- Alarcón, L.F., and Ashley, D.B. 1996. Modeling project performance for decision making. *Journal of Construction Engineering and Management* **122**(3): 265-273.
- Ashley, D.B., Lurie, C.S., and Jaselskis, E.J. 1987. Determinants of construction project success . Project Management Institute.
- Bassioni, H., Price, A., and Hassan, T. 2004. Performance Measurement in Construction. *Journal of Management in Engineering* **20**(2): 42-50.
- Cha, H.S., and Kim, C.K. 2011. Quantitative approach for project performance measurement on building construction in South Korea. *KSCE Journal of Civil Engineering* **15**(8): 1319-1328.
- Chan, A.P., and Chan, A.P. 2004. Key performance indicators for measuring construction success. *Benchmarking: An International Journal* **11**(2): 203-221.
- Cheung, S.O., Suen, H.C., and Cheung, K.K. 2004. PPMS: a web-based construction project performance monitoring system. *Automation in construction* **13**(3): 361-376.
- Chiu, S.L. 1994. Fuzzy model identification based on cluster estimation. *Journal of intelligent and Fuzzy systems* **2**(3): 267-278.
- Chua, D.K.H., Kog, Y.-C., and Loh, P.K. 1999. Critical success factors for different project objectives. *Journal of construction engineering and management* **125**(3): 142-150.

- 409 CII. 2014. General Large Metrics Definition-Performance Metric Formulas and Definitions. CII-
410 Construction Industry Institute.
- 411 Cox, R.F., Issa, R.R., and Ahrens, D. 2003. Management's perception of key performance
412 indicators for construction. *Journal of construction engineering and management* **129**(2): 142-
413 151.
- 414 CURT, C.u.r. 2005. Construction Measures: Key Performance Indicators. UP-101, The
415 Construction Users Roundtable.
- 416 DCC. 2016. Corporate Plan Summary 2016–2017 to 2020–2021. Defense Construction Canada,
417 Canada.
- 418 Dissanayake, M., and Fayek, A.R. 2008. Soft computing approach to construction performance
419 prediction and diagnosis. *Canadian Journal of Civil Engineering* **35**(8): 764-776.
- 420 Egan, J. 1998. The Egan report-rethinking construction. report of the construction industry task
421 force to the deputy prime minister. London.
- 422 Esmaeili, M., Osanloo, M., Rashidinejad, F., Bazzazi, A.A., and Taji, M. 2014. Multiple
423 regression, ANN and ANFIS models for prediction of backbreak in the open pit blasting.
424 *Engineering with computers* **30**(4): 549-558.
- 425 Freeman, M., and Beale, P. 199 .2Measuring project success, Project Management Institute.
- 426 Griffith, A.F., Gibson, G.E., Hamilton, M.R., Tortora, A.L., and Wilson, C.T. 1999. Project
427 success index for capital facility construction projects. *Journal of Performance of Constructed*
428 *Facilities* **13**(1): 39-45.
- 429 Group, K.W., and Britain, G. 2000. KPI Report for the Minister for Construction. Department of
430 the Environment, Transport and the Regions.

- 431 Hanna, A.S., Lotfallah, W., Aoun, D.G., and Asmar, M.E. 2014. Mathematical Formulation of
432 the Project Quarterback Rating: New Framework to Assess Construction Project Performance.
433 Journal of Construction Engineering and Management **140**(8): 04014033.
- 434 Haponava, T., and Al-Jibouri, S. 2012. Proposed System for Measuring Project Performance
435 Using Process-Based Key Performance Indicators. Journal of Management in Engineering **28**(2):
436 140-149.
- 437 Heravi, G., and Ilbeigi, M. 2012. Development of a comprehensive model for construction
438 project success evaluation by contractors. Engineering, Construction and Architectural
439 Management **19**(5): 526-542.
- 440 Holland, J. 1975. Adaption in natural and artificial systems. Ann Arbor MI: The University of
441 Michigan Press.
- 442 Jang, J.R. 1997. MATLAB: Fuzzy logic toolbox user's guide: Version 1. Math Works.
- 443 Jia, C., Wei, L., Wang, H., and Yang, J. 2015. A hybrid model based on wavelet decomposition-
444 reconstruction in track irregularity state forecasting. Mathematical Problems in Engineering.
- 445 Kerzner, H.R. 2011. Project management metrics, KPIs, and dashboards: a guide to measuring
446 and monitoring project performance. John Wiley & Sons.
- 447 Khosravi, S., and Afshari, H. 2011. A success measurement model for construction projects.
448 KPIzone, C.f.C.I. 2014. KPIzone. Constructing Excellence Industry Benchmarks.
- 449 Lauras, M., Marques, G., and Gourc, D. 2010. Towards a multi-dimensional project performance
450 measurement system. Decision Support Systems **48**(2): 342-353.
- 451 Leon, H., Osman, H., Georgy, M., and Elsaid, M. 2017. System Dynamics Approach for
452 Forecasting Performance of Construction Projects. Journal of Management in Engineering **34**(1):
453 04017049.

- 454 Li, J., Moselhi, O., and Alkass, S. 2006. Forecasting project status by using fuzzy logic. Journal
455 of construction engineering and management **132**(11): 1193-1202.
- 456 Ling, F.Y.Y., Chan, S.L., Chong, E., and Ee, L.P. 2004. Predicting performance of design-build
457 and design-bid-build projects. Journal of Construction Engineering and Management **130**(1): 75-
458 83.
- 459 Marques, G., Gourc, D., and Lauras, M. 2011. Multi-criteria performance analysis for decision
460 making in project management. International Journal of Project Management **29**(8): 1057-1069.
- 461 Menches, C.L., and Hanna, A.S. 2006. Quantitative measurement of successful performance
462 from the project manager's perspective. Journal of construction engineering and management
463 .1293-1284 :(12)132
- 464 Nasrollahzadeh, K., and Basiri, M.M. 2014. Prediction of shear strength of FRP reinforced
465 concrete beams using fuzzy inference system. Expert Systems with Applications **41**(4): 1006-
466 1020.
- 467 Nassar, N., and AbouRizk, S. 2014. Practical Application for Integrated Performance
468 Measurement of Construction Projects. Journal of Management in Engineering: 04014027.
- 469 Ngacho, C., and Das, D. 2015. A performance evaluation framework of construction projects:
470 insights from literature. International Journal of Project Organisation and Management **7**(2): 151-
471 173.
- 472 Nilashi, M., Dalvi-Esfahani, M., Ibrahim, O., Bagherifard, K., Mardani, A., and Zakuan, N.
473 2017. A soft computing method for the prediction of energy performance of residential
474 buildings. Measurement **109**: 268-280.

- 475 Nilashi, M., Zakaria, R., Ibrahim, O., Majid, M.Z.A., Zin, R.M., Chughtai, M.W., Abidin, N.I.Z.,
476 Sahamir, S.R., and Yakubu, D.A. 2015a. A knowledge-based expert system for assessing the
477 performance level of green buildings. *Knowledge-Based Systems* **86**: 194-209.
- 478 Nilashi, M., Zakaria, R., Ibrahim, O., Majid, M.Z.A., Zin, R.M., and Farahmand, M. 2015b.
479 MCPCM: a DEMATEL-ANP-based multi-criteria decision-making approach to evaluate the
480 critical success factors in construction projects. *Arabian Journal for Science and Engineering*
481 **40**(2): 343-361.
- 482 Parsanejad, M. 2013. Applying Association Rules to Explore Relationships among Project
483 Success Criteria. *Journal of Industrial and Intelligent Information* Vol **1**.(1)
- 484 Paul, S.K., Azeem, A., and Ghosh, A.K. 2015. Application of adaptive neuro-fuzzy inference
485 system and artificial neural network in inventory level forecasting. *International Journal of*
486 *Business Information Systems* **18**(3): 268-284.
- 487 Rad, P.F. 2003. Project success attributes. *Cost Engineering-Morgantown* **45**(4): 23-29.
- 488 Reenu, M., Kumar, R., and Babu, S. 2017. Construction Project Performance Model Using
489 Artificial Neural Network. *International Journal of Recent Trends in Engineering & Research*
490 (IJRTER).
- 491 Roberts, M., and Latorre, V. 2009. KPIs in the UK's Construction Industry: Using System
492 Dynamics to Understand Underachievement. *Revista de la Construcción* **8**(1): 69-82.
- 493 Saaty, T.L. 1988. What is the analytic hierarchy process? *In* *Mathematical models for decision*
494 *support*. Springer. pp. 109-121.
- 495 Saaty, T.L., and Vargas, L.G. 1982. *The logic of priorities: Application in business, energy,*
496 *health, and transportation*. Nijhoff, Boston.

- 497 Shahin, M.A., Maier, H.R., and Jaksa, M.B. 2004. Data division for developing neural networks
498 applied to geotechnical engineering. *Journal of Computing in Civil Engineering* **18**(2): 105-114.
- 499 Shaikh, S.S., and Darade, M.M. 2017. Key Performance Indicator for Measuring and Improving
500 Quality of Construction Projects. *International Research Journal of Engineering and Technology*
501 *)IRJET.(*
- 502 Swan, W., and Kyng, E. 2005. An introduction to key performance indicators. Manchester:
503 Centre for Construction Innovation.
- 504 Tucker, R.L., Shields, D.R., and Thomas, S.R. 2003. Measurement of Construction Phase
505 Success of Projects. *In Construction Research Congress*. pp. 1-8.
- 506 Wan, C.C., Hoogendoorn, M., and Muhonen, V. 2017. Forecasting E-commerce Key
507 Performance Indicators, Vrije Universiteit Amsterdam.
- 508 Wang, Y.-R., Yu, C.-Y., and Chan, H.-H. 2012. Predicting construction cost and schedule
509 success using artificial neural networks ensemble and support vector machines classification
510 models. *International Journal of Project Management* **30**(4): 470-478.
- 511 Wester, S. 2013. Identify and Develop Key Performance Indicators for High-Technological
512 Engineering and Construction Projects-A Case Study of Kockums AB, Lunds University.
- 513 **Fig. 1.** Flowchart of research methodology
- 514 **Fig. 2.** Components of the proposed PI prediction framework
- 515 **Fig. 3.** KPIs scores from the questionnaire
- 516 **Fig. 4.** Developing 18 neuro-fuzzy models for predicting six KPIs for three critical stages

517 **Fig. 5.** Flowchart of the neuro-fuzzy model

518 **Fig. 6.** Sensitivity analysis of whole project KPIs based on each project KPI in three stages

519 (1=Very Low, 2=Low, 3=Medium-Low, 4=Medium, 5=Medium-High, 6=High, 7=Very High)

520 **Fig. 7.** Comparison between PI of the model and PI of questionnaires in three stages

521 **Fig. 8.** Comparing PI derived from the proposed model and validation data

Draft

<https://mc06.manuscriptcentral.com/cjce-pubs>

Statistics of input parameters for construction projects used in questionnaire

	Parameter	Min.	Max.	Mean	COV (%)
Initial stage	<i>Cost</i>	2.00	7.00	4.69	34.94
	<i>Time</i>	1.00	7.00	4.73	35.59
	<i>Quality</i>	3.00	7.00	5.65	18.02
	<i>Safety</i>	3.00	7.00	5.54	19.13
	<i>Client satisfaction</i>	3.00	7.00	5.66	17.02
	<i>Project team satisfaction</i>	3.00	7.00	5.53	16.04
Middle stage	<i>Cost</i>	2.00	7.00	4.95	23.70
	<i>Time</i>	1.00	7.00	5.08	24.52
	<i>Quality</i>	3.00	7.00	5.60	18.58
	<i>Safety</i>	3.00	7.00	5.26	19.48
	<i>Client satisfaction</i>	2.00	7.00	5.51	19.52
	<i>Project team satisfaction</i>	3.00	7.00	5.30	16.14
Finishing stage	<i>Cost</i>	2.00	7.00	5.27	23.04
	<i>Time</i>	2.00	7.00	5.14	27.99
	<i>Quality</i>	2.00	7.00	5.79	19.63
	<i>Safety</i>	2.00	7.00	5.53	21.72
	<i>Client satisfaction</i>	2.00	7.00	5.61	21.60
	<i>Project team satisfaction</i>	2.00	7.00	5.33	20.52
Whole project	<i>Cost</i>	2.00	7.00	4.98	23.70
	<i>Time</i>	2.00	7.00	4.76	26.56
	<i>Quality</i>	3.00	7.00	5.76	18.64
	<i>Safety</i>	3.00	7.00	5.42	20.51
	<i>Client satisfaction</i>	2.00	7.00	5.61	20.06
	<i>Project team satisfaction</i>	2.00	7.00	5.31	20.39
	<i>Performance indicator (PI)</i>	3.00	7.00	5.29	17.02

Note: Very low=1, Low=2, Medium low=3, Medium=4, Medium high=5, High=6, Very high=7

Pair-Wise Comparison Matrix						
KPIs	Cost	Time	Quality	Safety	Client Satisfaction	Project Team Satisfaction
Cost	1.000	1.015	0.454	0.841	0.488	0.527
Time	0.986	1.000	0.393	0.897	0.515	0.499
Quality	2.205	2.544	1.000	2.502	2.394	2.394
Safety	1.190	1.115	0.400	1.000	0.912	0.986
Client Satisfaction	2.051	1.940	0.418	1.096	1.000	1.246
Project Team Satisfaction	1.898	2.002	0.418	1.015	0.803	1.000

Pair-Wise Comparison Matrix - Normalized						
Cost	Time	Quality	Safety	Client Satisfaction	Project Team Satisfaction	Weight Of KPIs using AHP
0.107	0.106	0.147	0.114	0.080	0.079	0.106
0.106	0.104	0.128	0.122	0.084	0.075	0.103
0.236	0.265	0.325	0.340	0.392	0.360	0.320
0.128	0.116	0.130	0.136	0.149	0.148	0.134
0.220	0.202	0.136	0.149	0.164	0.187	0.176
0.203	0.208	0.136	0.138	0.131	0.150	0.161

<i>KPIs</i>	<i>Cost</i>	<i>Time</i>	<i>Quality</i>	<i>Safety</i>	<i>Client Satisfaction</i>	<i>Project Team Satisfaction</i>
<i>Weight Of KPIs</i>	0.027	0.193	0.292	0.100	0.062	0.327

Draft

Initial stage				
No.	PI (AHP) Questionnaire	PI (AHP) Model	PI (GA) Questionnaire	PI (GA) Model
1	4.63	4.66	5.81	4.72
2	6.42	6.59	6.33	6.21
3	1.00	1.18	5.53	1.00
4	4.63	4.66	5.04	4.72
5	5.13	5.59	5.88	5.08
6	3.37	3.79	6.45	3.77
7	4.82	3.85	5.71	6.15
8	4.63	4.66	5.41	4.72
9	4.62	4.66	5.44	4.48
10	4.63	4.66	5.58	4.72
11	5.66	5.56	4.50	6.23
12	5.16	4.38	4.31	5.01
13	1.00	1.00	3.78	1.00
14	4.79	4.89	3.00	4.87
15	4.79	4.89	3.00	4.87
16	6.20	6.30	5.35	6.32
Mean: 0.9991		Mean: 1.4769		
COV(%): 9.36%		COV(%): 85.92%		
RMSE: 0.3574		RMSE: 1.7869		
MAPE(%): 5.52%		MAPE(%): 30.49%		

Middle stage				
No.	PI (AHP) Questionnaire	PI (AHP) Model	PI (GA) Questionnaire	PI (GA) Model
1	5.73	4.35	5.81	4.37
2	6.43	6.26	6.33	5.99
3	5.44	5.65	5.53	5.63
4	5.22	4.94	5.04	4.69
5	6.00	5.00	6.01	5.00
6	5.83	5.34	6.45	5.55
7	5.60	3.94	5.71	3.90
8	5.28	4.54	5.41	4.36
9	5.50	4.79	5.44	4.69
10	5.24	5.08	5.46	5.10
11	4.51	4.19	4.37	4.12
12	4.40	5.10	4.25	5.10
13	3.79	4.25	3.78	4.25
14	3.00	4.92	3.00	4.90
15	3.00	4.30	3.00	4.20
16	5.62	4.41	5.35	4.27
Mean: 1.0520		Mean: 1.0691		
COV(%): 19.79%		COV(%): 20.23%		
RMSE: 0.9609		RMSE: 1.0111		
MAPE(%): 17.79%		MAPE(%): 18.85%		

No.	Finishing stage			
	PI (AHP) Questionnaire	PI (AHP) Model	PI (GA) Questionnaire	PI (GA) Model
1	5.73	5.12	5.81	5.16
2	6.43	6.32	6.33	6.29
3	5.44	6.04	5.53	5.96
4	5.22	4.69	5.04	4.63
5	5.65	5.39	5.88	5.37
6	5.83	5.07	6.45	5.26
7	5.60	4.53	5.71	4.47
8	5.28	4.65	5.41	4.64
9	5.50	5.07	5.44	5.10
10	5.59	5.41	5.58	5.46
11	4.86	5.50	4.50	5.44
12	4.57	5.20	4.31	5.14
13	3.79	3.59	3.78	3.65
14	3.00	3.68	3.00	3.66
15	3.00	4.33	3.00	4.34
16	5.62	4.32	5.35	4.32
Mean: 1.0291			Mean: 1.0281	
COV(%): 15.00%			COV(%): 15.87%	
RMSE: 0.7158			RMSE: 0.7727	
MAPE(%): 13.48%			MAPE(%): 14.27%

Draft

168x83mm (300 x 300 DPI)

111x164mm (300 x 300 DPI)

115x164mm (300 x 300 DPI)