Hybrid GA-MANFIS Model for Organizational Competencies and Performance in Construction

Getaneh Gezahegne Tiruneh, Ph.D., A.M. ASCE[1], and

Aminah Robinson Fayek, Ph.D., P.Eng., M.ASCE[2]

Abstract

The majority of competency and performance modeling methods available in the literature are deterministic conceptual, statistical, and/or regression models that cannot capture the subjective uncertainty, complex, and nonlinear relationships inherent in construction, which makes accurate prediction difficult. Past studies utilized neuro-fuzzy system (NFS) models, such as adaptive neuro-fuzzy inference system (ANFIS), that combine the learning power of artificial neural networks and functionality of fuzzy systems to develop accurate predictive models. ANFIS is robust, fast, and effective in solving complex problems for a range of real-world construction engineering and management (CEM) applications. NFS models such as ANFIS have some limitations in handling multiple outputs common in construction industry problems, such as being prone to early convergence due to local minima entrapment. To address these limitations, this paper proposes a hybrid NFS combining the evolutionary optimization technique of genetic algorithm (GA) with multi-output adaptive neuro-fuzzy inference system (MANFIS) that can handle multi-input multi-output (MIMO) problems for CEM applications. The proposed modeling approach is demonstrated using a case study that showed good results in predicting multiple

[1] Ph.D., Department of Civil and Environmental Engineering, University of Alberta, 1-050 Markin/CNRL Natural Resources Engineering Facility, 9105 116 St. NW, Edmonton AB T6G 2W2, Canada.

[2] Tier 1 Canada Research Chair in Fuzzy Hybrid Decision Support Systems for Construction, NSERC Industrial Research Chair in Strategic Construction Modeling and Delivery, Professor, Hole School of Construction Engineering, Department of Civil and Environmental Engineering, University of Alberta, Edmonton, AB, Canada (corresponding author). aminah.robinson@ualberta.ca

organizational performance metrics using organizational competencies. The contributions of this paper are threefold: it (1) proposes a novel methodology of integrating different computing techniques for developing a GA-based multiple ANFIS (GA-MANFIS) model that can handle complex and nonlinear MIMO problems inherent in construction processes and practices, (2) relates organizational competencies to performance and predicts multiple organizational performance metrics, and (3) provides a GA-based feature selection approach that reduces data dimensionality, enabling identification of organizational competencies that significantly influence organizational performance. By uniquely integrating these techniques, this model enables construction organizations to evaluate their competencies and predict multiple organizational performance metrics simultaneously, and researchers can adapt it for a variety of construction contexts.

Keywords: Artificial intelligence; Construction; Hybrid neuro-fuzzy systems; Organizational issues; Organizational competency; Performance.

Introduction

Construction engineering and management (CEM) is an experience-based discipline – knowledge acquired from previous work plays a key role for successful performance in executing projects and organizational operations within a construction environment that is frequently complex and fraught with uncertainty (Cheng and Roy 2010). Adopting effective strategies and performance measurement methods is critical to improving the performance of construction organizations (Tiruneh and Fayek 2021). Because an organization's performance depends greatly on its people and their competencies, having competency-based performance measures is an important recourse for engendering the performance of CEM organizations (Altuncan and Tanyer 2018). However, the variables that capture CEM organizational competencies and performance are highly dimensional and both quantitative and qualitative in nature. Implementing a dimensionality reduction technique such as feature selection (FS) is critical to developing a concise and interpretable model with low complexity and high accuracy. Therefore, an FS using evolutionary algorithms such as genetic algorithm (GA) can yield better results and be computationally feasible (Tiruneh and Fayek 2019). Thus, organizational competency and performance modeling techniques that can handle both quantitative and qualitative variable types, uncertainty, complexity, and nonlinear relationships are needed.

Real-world CEM problems are characterized by their nonspecificity, uncertainty, complexity, dynamism, and nonlinearity, which creates challenges for construction management and makes accurate prediction difficult (Elbaz et al. 2020). Studies have indicated that lack of sufficient data (i.e., limitations in quantity and quality of data) and the subjective uncertainty associated with CEM problems make it difficult to explicitly represent such complex problems in a deterministic mathematical or statistical model (Cheng et al. 2015; Tokede et al. 2014). Therefore, one feasible approach to predicting performance is to use artificial intelligence (AI)-based models, such as hybrid neuro-fuzzy system (NFS), that combine the learning power of artificial neural networks (ANNs) with the functionality of fuzzy systems (i.e., improving reasoning and inference as well as representing knowledge explicitly) that are suitable to solving complex problems with nonlinear relationships and subjective uncertainty and also offer high accuracy and low cost (Cheng et al. 2015; Tiruneh et al. 2020). ANFIS and other NFS have been widely used for modeling a variety of CEM applications. For instance, Jin (2010, 2011) employed ANFIS, the most commonly used NFS, for decision-making processes in efficient risk allocation. ANFIS possesses the capability to handle the unspecificity, uncertainty, nonlinearity, and complexity involved in most risk-allocation decision-making processes. Elmousalami (2020) demonstrated the suitability of computational intelligence (CI) techniques – which combine fuzzy logic, neuro computing, and evolutionary computing – used for parametric cost prediction models. Bayram and Al-Jibouri (2016) demonstrated that radial basis function (RBF) is more suitable for detailed estimates compared to reference class forecasting and simple linear regression analysis. Their results showed that RBF performed better in forecasting estimated versus actual costs of building construction projects (Bayram and Al-Jibouri 2016). Rashidi et al. (2011) proposed a neuro-fuzzy genetic system to identify decision-making criteria for selecting qualified project managers in construction. Afshari (2017) combined a group fuzzy linguistic evaluation model with Delphi method for selecting the most suitable project managers in construction companies. Moon and Chowdhury (2021) demonstrated a prior information–based neural network (PI-NN) having a better prediction capability for the 28-day concrete compressive strength using a 3-day compressive strength as prior information compared to conventional ANN. Gunduz and Elsherbeny (2021) proposed a multidimensional fuzzy model to quantify the performance of construction contract administration processes at the project level where the practical implementations of the proposed model led to identification of the top strategies used to improve construction contract administration performance. Siraj et al. (2016) developed AI-based (i.e., ANN, ANFIS, and fuzzy rule-based) compressive strength predictive models for high-performance concrete (HPC). Nazari and Sanjayan (2015) proposed a hybrid model based on ANFIS and imperialist competitive algorithm capable of predicting compressive strength. Tayfur et al. (2014) demonstrated that performance of fuzzy logic and ANN models were comparable for predicting strength of HPC. Shahhosseini and Sebt (2011) applied ANFIS for selecting construction project employees based on competency. Adeli and Jiang (2003) presented an adaptive neuro-fuzzy logic model that provided more accurate estimates of work zone capacity compared to empirical equations, especially when data for factors impacting work zone capacity are only partially available. Shahtaheri et al. (2015) developed an ANFIS-based model for estimating baseline rates for on-site work categories in construction. Tokede et al. (2014) proposed a neuro-fuzzy hybrid cost model for predicting the final cost of small water infrastructure project. However, few past competency and performance studies used hybrid NFS (Georgy et al. 2005; Omar and Fayek 2016). Some hybrid fuzzy systems have limitations related to early convergence due to local minima entrapment and poor generalization (Elbaz et al. 2020; Yuan et al. 2014). Therefore, a combination of hybrid NFS and evolutionary optimization algorithms has been utilized to develop more accurate predictive models.

Many real-world engineering problems, particularly in CEM, are complex and nonlinear MIMO systems in which the multiple output variables may each depend on all input variables (Acampora et al. 2014; Fattahi et al. 2018). This strong interdependence among variables leads to highly complex and dynamic systems that make MIMO models too imprecise and uncertain to be trained using conventional system modeling approaches (Acampora et al. 2014; Fattahi et al. 2018). However, because conventional NFSs are configured as multi-inputs single-output (MISO) systems, such as ANFIS, and therefore have limitations in handling MIMO systems (Acampora et al. 2014; Cheng et al. 2002), various approaches have used improved ANFIS methods for learning the behavior of MIMO systems, such as MANFIS (Acampora et al. 2014; Das and Winter 2016). Because MANFIS is an extension and generalization of ANFIS for handling multiple outputs, ANFIS is the building block of MANFIS (Cheng et al. 2002), with several single-output neuro-fuzzy system (ANFIS) blocks being required and combined to develop a MANFIS. Some challenges in developing an effective MANFIS model are the choice of appropriate type of membership function (MF) (e.g., triangular, trapezoidal, Gaussian), clustering method (e.g., grid partition method, subtractive clustering, fuzzy *c*-means), and learning and/or optimization algorithm (e.g., gradient, hybrid, population-based) to be used for individual ANFIS. The type of MF selected must be suitable to the problem being modeled. The choice of clustering method for input data is critical, because it can impact the number of rules and generalization power of the model (Fattahi et al. 2018; Nayak et al. 2015). Researchers have recommended using fuzzy *c*-means (FCM) to avoid exponential growth of rules due to the number of input variables (Fattahi et al. 2018). The selected learning/optimization algorithm needs to improve the effectiveness of MFs and fuzzy rules in the model (Abd Elaziz et al. 2019; Elbaz et al. 2019; Elbaz et al. 2020). Thus, the configuration of ANFIS within a MANFIS model is critical to developing an effective and efficient predictive model.

NFS modeling techniques that can handle multiple outputs are common and widely used in non-construction research domains (Acampora et al. 2014; Das and Winter 2016). To date, a gap exists in addressing MIMO NFS modeling techniques for CEM problems, specifically for predicting multiple performance metrics. To address the need for developing modeling approaches that can handle complex, nonlinear MIMO performance prediction problems for construction applications, this paper proposes a novel methodology using a hybrid GA-MANFIS approach for modeling construction organizational competencies and simultaneously predicting multiple performance metrics. The objectives of this paper include (1) proposing a novel methodology for developing a hybrid GA-MANFIS modeling approach that can handle MIMO problems inherent in construction processes and practices, (2) relating organizational competencies to performance and predicting multiple organizational performance metrics, and (3) providing a GA-FS optimization approach that reduces dimensionality of data and enables identification of organizational competencies that significantly influence organizational performance. The results of a case study applying the proposed GA-MANFIS model are also presented.

The rest of this paper is structured as follows. First, an overview of past competency and performance modeling techniques, the application and limitations of ANFIS/MANFIS in construction problems and hybrid GA-based ANFIS and GA-MANFIS are discussed. Second, a novel methodology is presented for developing the proposed hybrid GA-MANFIS model for predicting organizational performance using organizational competencies. Third, the case study results are presented to illustrate the proposed methodology and show its application in CEM. Finally, conclusions and suggestions for future research are presented.

Literature Review

Overview of competency and performance modeling research

Competency models realize specific combinations of knowledge, skills, and other personal characteristics necessary for efficient execution of tasks (i.e., effective performance) in an organization (Tiruneh and Fayek 2021). Past competency and performance modeling methods available in the literature can be categorized into six groups: conceptual, SEM/correlation/regression, ANN, fuzzy systems, hybrid fuzzy methods, and NFS (Tiruneh and Fayek 2018). These modeling methods are discussed below.

Competency-based multidimensional conceptual models have been proposed to determine the performance of project managers. For instance, Suhairom et al. (2014) developed a conceptual competency model that relates personality and technical, non-technical, and career competency to superior work performance. Liu et al. (2010) proposed a model that showed a positive relationship between the levels of a project team's general task completion competency and the project team's performance. Moreover, conceptual models that link competencies to performance show the positive impact of competencies on performance (Ahadzie et al. 2009; Ahadzie et al. 2014); however, these conceptual models are generic and limited to specific aspects and hence do not capture industry and organizational contexts.

Structural equation models (SEM) and correlation/regression models have been used to analyze competencies and determine performance. Dainty et al. (2005) developed a statistical model to determine competencies defining superior management performance. Cheng et al. (2007) developed an empirical model using path analysis to examine the effects of competencies and job performance on overall project performance. Bolivar-Ramos et al. (2012) developed an SEM to determine organizational performance. Altuncan and Tanyer (2018) proposed a performance assessment methodology for conflict management based on competency theory; however, their model is limited in providing statistically generalizable results because of the unique characteristics of conflict in construction. Some studies employed regression models that correlate project managers' behavior with final project outcomes (Ling 2002, 2004). A regression model developed in past studies confirmed the impact of organizational competency on organizational performance (Liang et al. 2013; Levenson et al 2006; Liu et al. 2010). Liang et al. (2013) indicated that the variables of core competences are positively correlated with organizational performance. However, the SEM, correlation, and regression models discussed do not capture the complex relationships or subjective uncertainty inherent in CEM problems.

A few studies used ANN (Elwakil et al. 2009) and fuzzy expert systems (FES) (Poveda and Fayek 2009) to determine and predict performance. However, ANN models lack inference and explicit knowledge representation, and FES lacks learning capability. Fuzzy hybrid models and NFS that combine ANN and fuzzy systems have also been developed to remedy the drawbacks of fuzzy systems and ANN models (Georgy et al. 2005; Omar and Fayek 2016).

Modeling techniques that relate construction organizational competency to performance and enable organizations to determine and predict performance are essential in the construction industry (Tiruneh and Fayek 2021). Moreover, predicting organizational performance helps researchers and organizations identify weak organizational processes and practices in order to improve performance (Georgy et al. 2005; Elwakil et al. 2009). However, most modeling techniques used in previous studies lack the ability to capture overall organizational competency and performance. Table 1 presents a summary of advantages and limitations of past competency and performance modeling methods.

The majority of competency and performance modeling methods presented in Table 1 are conceptual and/or correlation/regression models and thus subject to the limitations noted above. Although hybrid NFS that combine the learning power of ANN, functionality of fuzzy systems, and evolutionary optimization algorithms have been utilized previously to develop accurate predictive models, most hybrid NFS such as ANFIS cannot handle multiple outputs because of their MISO configuration. Thus, this paper proposes GA-MANFIS, a hybrid NFS modeling approach that can handle the multiple outputs inherent in real-world engineering problems.

As discussed above, very few studies have utilized hybrid NFS and evolutionary optimization algorithms. To the authors' knowledge, prediction of multiple performance metrics simultaneously using a hybrid GA-MANFIS has not been done in the construction domain. Therefore, this paper demonstrates the use of a hybrid GA-MANFIS model in CEM.

The following section presents the applications and limitations of GA-MANFIS and its components with respect to modeling organizational competencies and performance in CEM.

Table 1. Advantages and limitations of past competency and performance modeling methods

Method and References	Advantages	Limitations
Conceptual: Ahadzie et al. (2009, 2014); Liu et al. (2010); Suhairom et al. (2014)	• Clear distinction between competency and performance • Map competencies to performance	• Lack evidence-based relation; hence, needs validation
SEM, correlation, and/or regression: Altuncan and Tanyer (2018); Bolivar-Ramos et al. (2012); Cheng et al. (2007); Dainty et al. (2005); Levenson (2006); Liang et al. (2013); Ling (2002, 2004); Liu et al. (2010);	• Captures relationships between competency and performance • Establishes causal link between competencies and performance	• Generic and developed with limited data; hence, difficult to generalization • Lacks context
ANN: Adeli and Jiang (2003), Siraj et al. (2016), Tayfur et al. (2014)	• Capture complex and linear relationships • Possess learning capability	• Black box nature (lack transparency) • Do not capture subjective uncertainty • Lack interpretability
Fuzzy systems: Poveda and Fayek (2009), Siraj et al. (2016), Tayfur et al. (2014)	• Represent conditional relationships, i.e., rule-based knowledge • Use linguistic terms to assess the degree of interactions • Capture expert knowledge on casual factors • Capture subjective uncertainty • Inferencing ability	• Curse of dimensionality • Lack learning capability
Hybrid fuzzy systems: Omar and Fayek (2016)	• Capture subjective uncertainty • Knowledge representation • Inferencing ability	• Lack model flexibility for varying contexts • Need development of multiple models • Limited in handling high dimensional data attributes
NFS (conventional and hybrid):	• Model complex and non-	• Lack handling of multiple

Method and References	Advantages	Limitations
Chen et al. (2018), Chen et al. (2010), Georgy et al. (2005),	linear relationships • Capture both subjective and objective measures • Possess learning capability • Inferencing ability	outputs • May have high model complexity • High computational time

Application and limitations of ANFIS/MANFIS in CEM

ANFIS has been one of the most popular prediction models among NFS techniques capable of input-output mapping of complex and nonlinear relationships, and it has been widely and successfully used for various construction applications (Tiruneh et al. 2020); however, few uses of ANFIS in modeling competency and performance have been explored. Omar and Fayek (2016) proposed a fuzzy neural network to model construction project competencies and performance and showed that enhancing construction project competencies can improve project performance. Georgy et al. (2005) utilized neuro-fuzzy intelligent systems for predicting engineering performance in a construction project. Elbaz et al. (2019) proposed hybrid GA-ANFIS and Elbaz et al. (2019) proposed particle swarm optimization (PSO)-based ANFIS (PSO-ANFIS) models to predict performance for tunneling projects. Cheng et al. (2012) used an evolutionary fuzzy hybrid neural network for dynamic project success assessment in the construction industry. Most studies other than Omar and Fayek (2016) focus on performance prediction using various factors other than competency. Thus, a great potential still exists for using ANFIS in analyzing organizational competencies, relating them to performance, and predicting organizational performance.

AI-based models such as ANFIS have good performance with desirable accuracy compared to the mathematical or regression models in real engineering practice (Elbaz et al. 2019; Yuan et al. 2014). However, ANFIS has two important limitations: slow computational convergence and a potential for being trapped in local minima (Elbaz et al. 2020). Therefore, ANFIS may provide less accurate results and/or distorted or inadequate explanations for problems (Elbaz et al. 2019). To overcome these limitations, ANFIS needs to be optimized with a heuristic optimization technique such as GA, PSO, artificial bee colony (ABC), or ant colony optimization (ACO) (Elbaz et al. 2019; Elbaz et al. 2020). Furthermore, the application of ANFIS in construction research has limitations in handling multiple outputs. The configuration of the ANFIS architecture is only suitable for MISO problems. As such, a need exists for developing a modeling approach that can improve ANFIS so that it can handle complex and nonlinear MIMO problems in CEM.

Despite its broad applicability, ANFIS fails to directly deal with MIMO systems because of its MISO structure (Acampora et al. 2014; Cheng et al. 2002). So, various approaches used improved ANFIS that can handle MIMO systems, such as MANFIS (Acampora et al. 2014; Cheng et al. 2002; Das and Winter 2016). MANFIS can be viewed as an aggregation of many independent ANFISs and capable of modeling highly nonlinear and complex systems (Cheng et al. 2002; Das and Winter 2016). The core of the proposed model is a processing layer of ANFIS modal blocks that each correspond to and predict a single output (Cheng et al. 2002; Das and Winter 2016; Malik and Arshad 2011).

Past studies showed MANFIS's good performance in approximating multiple outputs with the desired precision (Das and Winter 2016; Malik and Arshad 2011). Malik and Arshad (2011) demonstrated the performance of MANFIS in modeling a nuclear power plant's multivariable primary pressure control system, which indicated excellent agreement between predicted and actual data, hence confirming the model's effectiveness in a real-world situation. Das and Winter (2016) utilized a MANFIS model to predict multi-output urban transport modes (bus, train, tram, and walking) with high accuracy. Agah and Soleimanpourmoghadam (2020) proposed a MANFIS model to predict the existence of pollutant heavy metals in the environment and showed that

MANFIS predicts with high accuracy the concentration of four heavy metals in mine drainages. However, MANFIS has similar limitations to those of ANFIS – slow computational convergence and potential of being trapped in local minima – which result in low accuracy and poor generalization.

A hybrid GA-MANFIS

Studies have shown that evolutionary algorithms (e.g., GA, PSO, ABC, ACO) have significant capability to improve the performance and accuracy of ANFIS in prediction models and solving real-world engineering and/or CEM problems (Elbaz et al. 2019; Kaveh et al. 2018), although many studies indicated that GA-ANFIS models had the best performance with the highest accuracy compared to other hybrid approaches to modeling nonlinear and complex real-world engineering problems (Kaveh et al. 2018). GA has been used successfully in solving CEM problems, because it has robustness in determining a global optimal solution (Abd Elaziz et al. 2019; Kumar and Hynes 2020; Yuan et al. 2014). Furthermore, hybridizing a robust optimization algorithm such as GA with ANFIS as its training algorithm improves the effectiveness of MFs and fuzzy rules in the model (Abd Elaziz et al. 2019; Elbaz et al. 2019; Elbaz et al. 2020). A trend toward heuristic-based ANFIS training algorithms for better performance has been addressed in recent published studies (Elbaz et al. 2019; Elbaz et al. 2020; Tiruneh et al. 2020). Thus, optimization of ANFIS using GA can be extended to MANFIS models to improve model performance in predicting multiple outputs.

Optimization of a multiple-output system is performed by integrating a MANFIS network and various evolutionary algorithms such as GA to improve prediction capacity (Cheng et al. 2002). Cheng et al. (2002) proposed a hybrid MANFIS neuro-fuzzy network that uses GA to optimize multiple-objective decision-making problems. Tahmasebi and Hezarkhani (2012) investigated the performance of integrated neural-fuzzy and GA for MIMO problems to predict the ore grade from boreholes of copper deposits and showed that their proposed approach has excellent performance for grade estimation. However, review of past studies shows that very few focused on MANFIS in general and incorporating evolutionary optimization methods, especially GA. Thus, the proposed GA-MANFIS model enables construction organizations to identify and evaluate their competencies that have significant impact on performance and to simultaneously predict multiple organizational performances. Additionally, the proposed GA-MANFIS model can serve as a reference to extend the scope of its application by researchers, practitioners, and different CEM organizations according to their context.

GA-MANFIS Modeling Methodology

Steps for developing the proposed hybrid GA-MANFIS model are: (1) identify organizational competencies and performance metrics, and collect data, (2) prepare the organizational competencies and performance metrics data, (3) select organizational competency features, (4) develop the GA-MANFIS model, and (5) verify and validate the GA-MANFIS model. The methodology is illustrated in Figure 1 and described below.

Identify organizational competencies and performance metrics and collect data

First, an initial list of organizational competencies and performance metrics was derived from existing research in both construction and non-construction domains. A total of 157 competencies were initially identified and grouped into two sets of organizational competencies: functional (how the organization operates and functions) and behavioral (individual/organizational attributes). The list of competencies was further refined to avoid redundancy and similarity. A total of 101 competencies (i.e., 58 functional and 43 behavioral competencies) were selected, and a total of 44

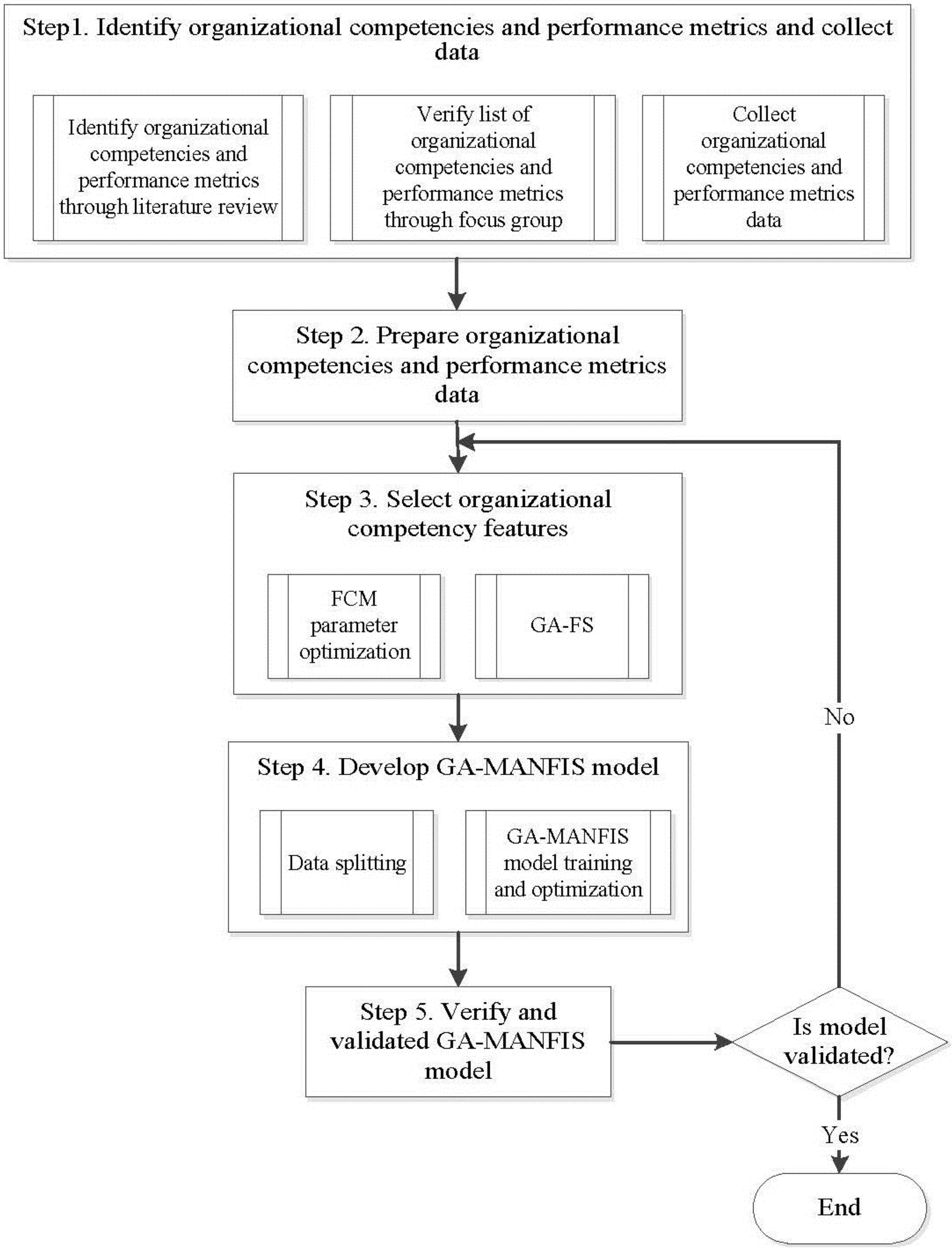


Fig. 1. GA-MANFIS modeling methodology for construction organizational competencies and performance.

organizational performance metrics were classified as key performance indicators (KPIs), key performance outcomes (KPOs), and perception measures (PerMs).

Next, a focus group was conducted to verify and validate the list and categorization of organizational competencies and performance metrics. Participants were experts who reviewed the list and proposed additional organization-level competencies and performance metrics they thought important. The initial list was updated to incorporate the experts' feedback and include proposed additional competencies backed by the literature. The resulting comprehensive list of organizational competencies and performance metrics not only considers the literature in construction and non-construction domains, but also captures the opinions of construction experts practicing in the industry. More details about the focus group results can be found in Tiruneh and Fayek (2021).

Then, data collection forms were based on the finalized list of organizational competencies and performance metrics from the focus group, and data was collected. The finalized list was pilot tested with a construction company to ensure that respondents could understand the forms and to check applicability of the evaluation and measurement scales and techniques of the data collection forms within CEM organizations. Final data collection forms consisted of 85 competencies and 42 organizational performance metrics were then prepared incorporating feedback from the pilot survey.

Two surveys – the senior management survey and the staff survey – were developed to collect data regarding organizational competencies influencing organizational performance. The surveys were distributed through Survey Monkey with a company's office and project personnel. Participants holding senior management positions completed the senior management survey, which addressed 85 competencies (48 functional and 37 behavioral competencies). All other participants, including project managers, field supervisors, and foremen, completed the staff survey, which addressed 63 competencies (34 functional and 29 behavioral competencies). The senior management survey addressed everything in the staff survey plus additional organizational competencies and performance metrics that can only be evaluated by senior management and were not known to the other respondent group.

Survey respondents evaluated organizational functional competencies based on maturity (the extent to which a specific competency exists in the organization) and impact on performance (the level of impact of a specific competency on overall performance of the organization). Respondents evaluated organizational behavioral competencies based on agreement (the extent to which the respondent agrees that a specific competency exists in the organization) and impact on performance. Maturity of functional competencies is measured on a scale ranging from 1 ("Informal") to 5 ("Optimized"). Agreement is measured on a scale ranging from 1 ("Strongly Disagree") to 7 ("Strongly Agree"). Impact on performance is measured on a scale ranging from 1 ("Extremely Low") to 7 ("Extremely High"). Actual company performance metrics data related to KPIs and KPOs were collected at the organizational (operational) and project levels using quantitative measures. Thus, performance data for KPIs and KPOs were extracted from relevant actual organizational/project documents. For performance, metrics related to PerMs were evaluated using a satisfaction scale ranging from 1 ("Extremely Unsatisfied") to 5 ("Extremely Satisfied"). Subjective performance measures related to KPIs and KPOs were evaluated using a scale ranging from 1 ("Very Low") to 5 ("Very High").

Prepare organizational competencies and performance metrics data

Data preparation or preprocessing techniques for modeling include data cleaning and data normalization (i.e., data transformation), which are usually implemented prior to any data-driven system modeling in order to eliminate responses or data instances that include outliers (i.e., noisy data), missing values, or bad data (Acampora et al. 2014; Cheng et al. 2015; Fattahi et al. 2018). These data preprocessing steps ensure that raw data collected or retrieved from the database and/or obtained from actual company and project documents are suitable for modeling.

Data cleaning

All online survey responses and performance data extracted from actual company and/or project documents were encoded to an Excel sheet. A total of 80 data instances were recorded and considered for model development. All survey responses and performance data were then checked for missing values, outliers, and inconsistencies. As part of the data cleaning step, survey responses and performance data with missing values and outliers were removed from the data. The data cleaning resulted in 62 data instances, which fell within the range of data instances used to develop MANFIS models in past studies (Agah and Soleimanpourmoghadam 2020; Cheng and Roy 2010; Cheng et al. 2015; Fattahi et al. 2018). Thus, the 62 data instances resulting from data cleaning for this study were considered sufficient, and were used to develop the proposed GA-MANFIS model.

Data normalization

Once cleaned, data are normalized using Equation (1), which transforms the dataset to the range of [0 1] in order to simplify and enhance training performance and improve prediction accuracy of the model. Normalizing the input-output data helps avoid domination of attributes in greater numeric ranges over smaller numeric ranges and to avoid numerical difficulties (Cheng and Roy 2010).

$$x_N = \frac{x_i - x_{min}}{x_{max} - x_{min}} \quad \textbf{(1)}$$

where, x_i and x_N are the original and normalized values, respectively, while x_{min} and x_{max} are the minimum and maximum values of x, respectively.

Select organizational competency features

After data cleaning and normalization, the number of input variables are reduced. High dimensionality of data makes it difficult to build a concise and efficient predictive model. Reducing data dimensionality helps reduce computational time and removes redundant or noisy attributes, thus improving model performance through increased predictive accuracy and interpretability. FS techniques, such as GA-FS, reduce data dimensionality and identify the best subset of data for which the predictive model has the highest accuracy in terms of the lowest root mean square error (RMSE), as expressed in Equation (2).

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_t - x_p)^2} \quad \textbf{(2)}$$

where, x_t and x_p are the actual/target and predicted values of x, respectively, while N is the number of data instances.

FCM parameter optimization for GA-FS

Prior to GA optimization, FCM parameters are optimized. FCM parameters include number of clusters, c, and fuzzification coefficient, m, which expresses the impact of the membership grades on the individual clusters. MATLAB programming language was used to develop a code that was run multiple times to find optimum values of c and m, with FCM performed on the cleaned and normalized input-output data. For FCM parameter optimization, c = 3 to 7 and m = 1.25 to 3.75 with 0.05 step were used. A total of 60 different runs were implemented, and the minimum RMSE and the FCM parameters for which the RMSE was minimum were recorded for each run. Next, a fuzzy inference system (FIS) was developed using the optimized FCM parameters. FCM-based FIS maps inputs to outputs using fuzzy logic or fuzzy set theory. Finally, the FIS was used to conduct GA-FS (Tiruneh and Fayek 2019), implemented through the following steps, as depicted in Figure 2:

1. Randomly generate an initial subset of the population, or system variables/features (i.e., organizational competencies), represented by binary chromosomes.
2. Evaluate the compatibility of each chromosome using RMSE as the fitness function.
3. Use selection, crossover, and mutation to create new generation or population based on fittest individuals from the previous generation.
4. New best offspring chromosomes partially or fully replace parents (i.e., old) chromosomes.
5. Repeat steps 2–4 until termination condition is satisfied. The chromosome with the highest accuracy in the last generation (i.e., organizational competencies) represented by ones are selected as the best subset of system variables for model development.

GA-FS to identify model input

Applying the optimum FCM parameters, an FIS was developed using the *genfisOptions* of MATLAB 2020b. Then, FS was conducted using binary-coded GA optimization on the FIS using the RMSE as the fitness function. The crossover and mutation probabilities were set as 0.8 and 0.1, respectively, while the number of generations was 100. The top five results were considered for the GA-FS step. For each result, a population of 50, 60, 80, and 100 was used, keeping the number of generations at 100. Therefore, 20 different combinations of GA-FS were conducted to identify the results with the best fitness (RMSE) values. After performing the FS using GA optimization, 19 competencies were selected out of the original 60.

Develop GA-MANFIS model

The proposed model development process has three steps: data splitting; model development and optimization; and model verification and validation. The model development steps depicted in Figure 3 are discussed below.

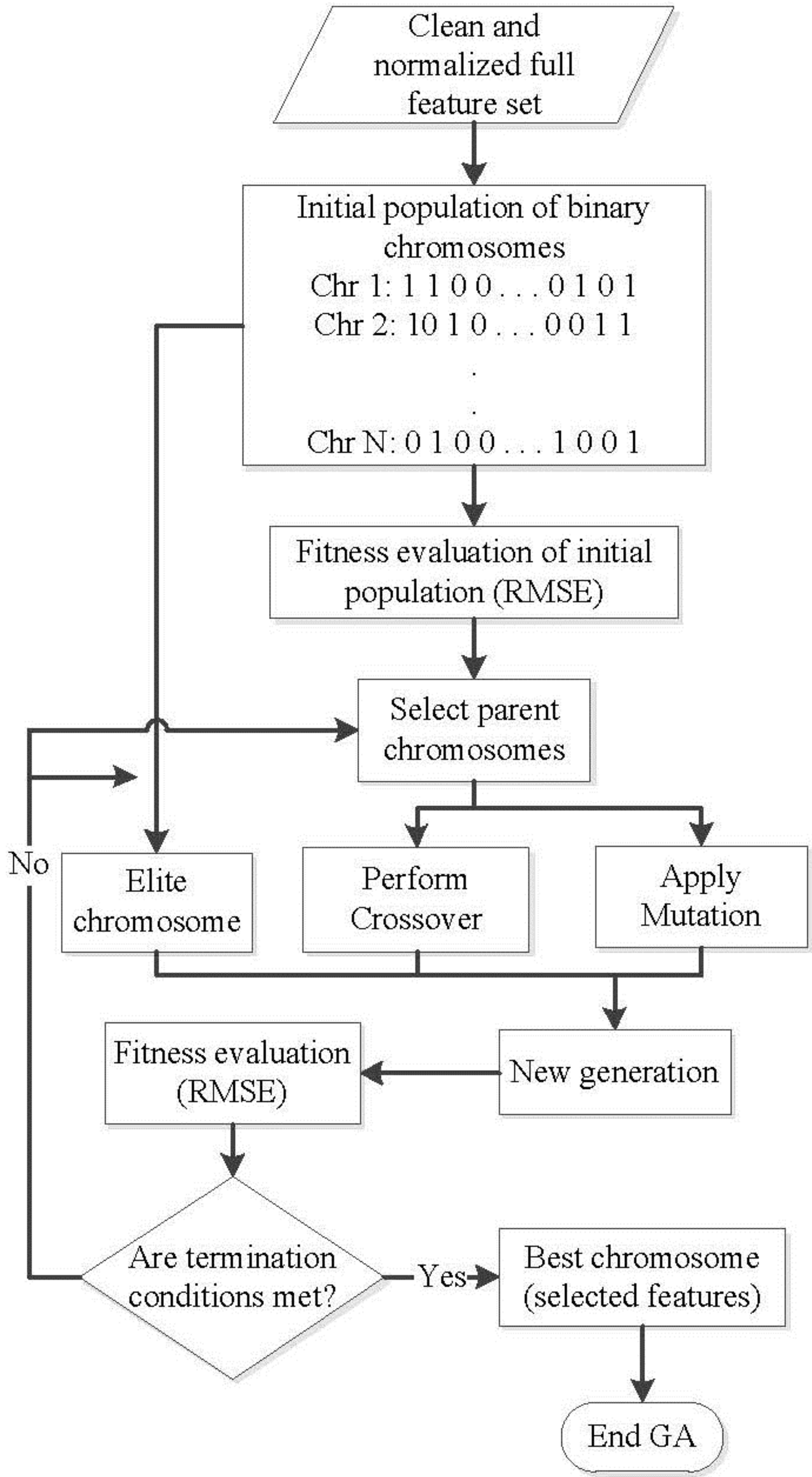


Fig. 2. FS using GA optimization.

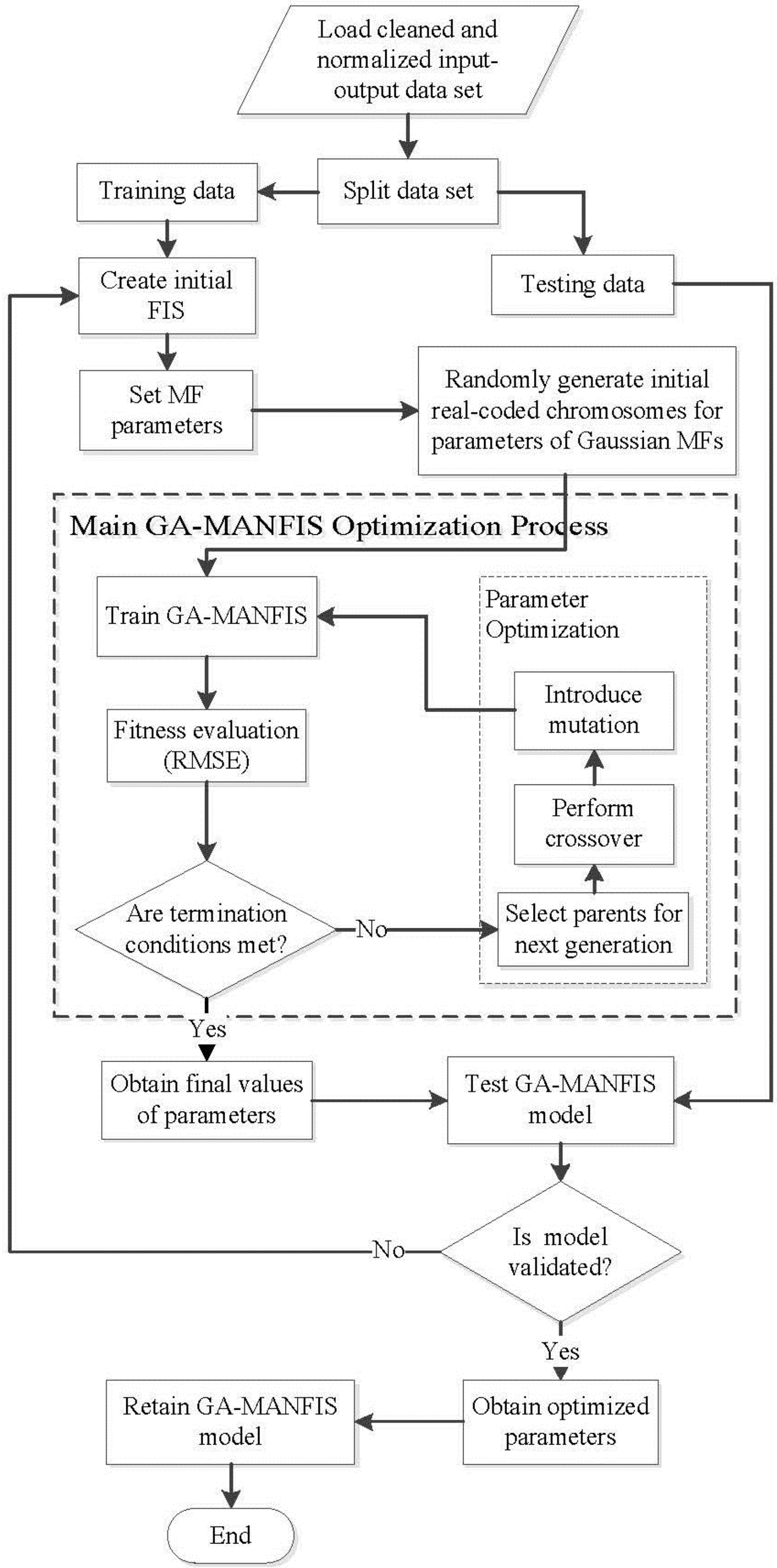


Fig. 3. GA-MANFIS model training and optimization.

Data splitting

To begin GA-MANFIS model development, all input variable data identified by GA-FS are shuffled in rows to ensure random arrangement before being divided into training and testing datasets. Past studies used different ratios of training to testing data depending on the availability of data. The most common ratio applied for model development (training dataset) to model validation (testing dataset) is 70/30. However, many studies that developed limited-data models used a ratio of 80/20 for training to testing data (Fattahi et al. 2018; Agah and Soleimanpourmoghadam 2020) or even 85/15 (Cheng and Roy 2010; Tahmasebi and Hezarkhani 2012). In this study, a ratio of 80/20 was used for model development.

Model development and optimization

The hybrid GA-MANFIS model was programmed in MATLAB R2020b. A Takagi-Sugeno FIS with Gaussian MF was applied to create the initial FIS to develop the GA-MANFIS model. Model architecture, input and output variables, development, training, and optimization procedures are discussed below.

The proposed GA-MANFIS model has three components, as shown in Figure 4. The input layer comprises the organizational competencies obtained from the GA-FS step. In the MIMO modal block layer, for *K* outputs, the model will have *K* number of ANFIS modal blocks (Cheng et al. 2002; Das and Winter 2016; Malik and Arshad 2011). So, each ANFIS block has a single input, is trained and optimized in parallel, and predicts a single output, ANFIS generates the number of MFs based on the FCM-based initial FIS, and the model predicts multiple outputs by using the same multiple inputs. Finally, the output layer comprises organizational performance metrics.

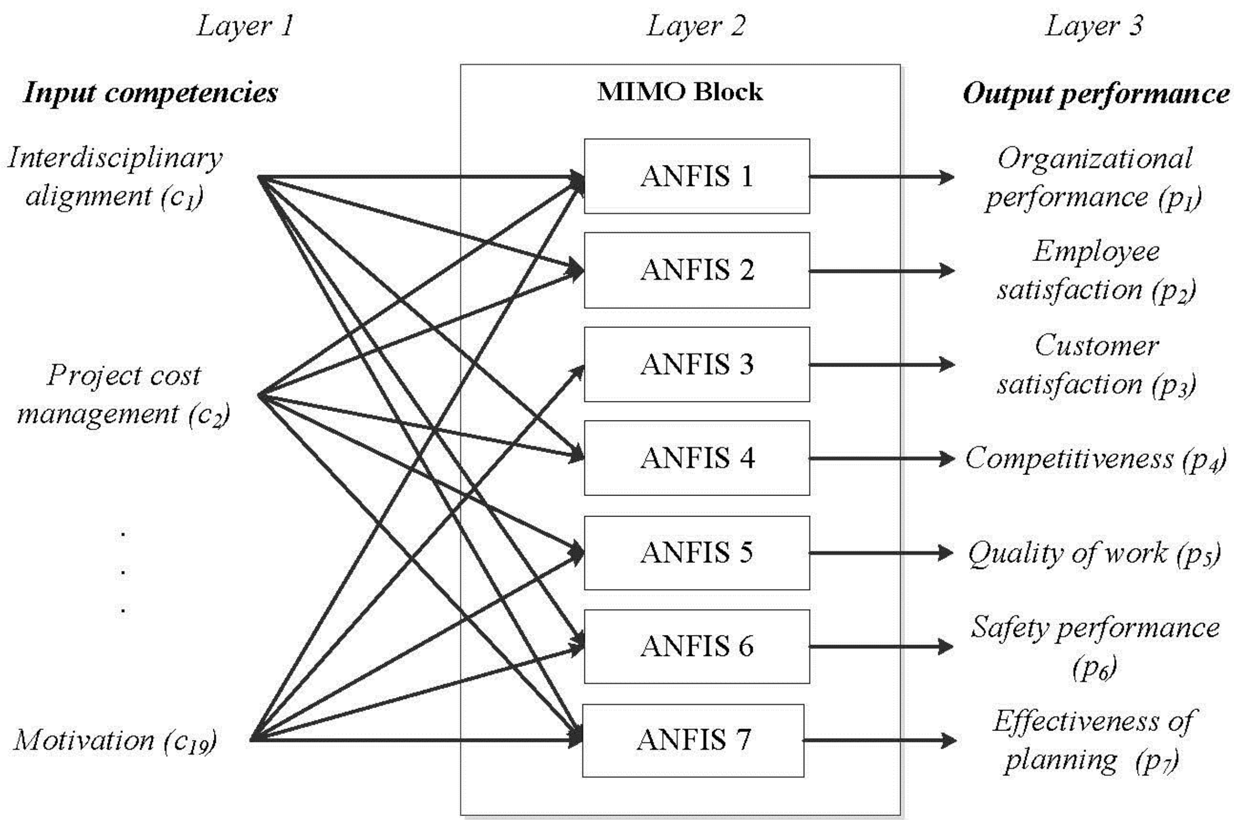


Fig. 4. GA-MANFIS model architecture for organizational competencies and performance.

As shown in Figure 4, organizational competencies served as the input variables for the GA-MANFIS model. Six performance metrics that include *Employee satisfaction* (p_2), *Customer satisfaction* (p_3), *Competitiveness* (p_4), *Quality of work* (p_5), *Safety performance* (p_6), and *Effectiveness of planning* (p_7) were identified as model outputs. A seventh model output, *Overall organizational performance* (p_1), was added by taking the average of the normalized values of the other six performance metrics to determine the overall organizational performance. Thus, the MIMO modal block of the MANFIS incorporated seven MISO ANFISs.

FCM clustering results in the development of a partition matrix ($\boldsymbol{U} = [u_{ik}]$) that includes the data points in each cluster (Pedrycz 2013). FCM clusters the input-output dataset into c numbers of clusters ($\boldsymbol{V} = [v_j]$) by determining a prototype (cluster center) for each cluster. Fuzzy partitioning is carried out through an iterative optimization by updating the partition matrix u_{ik} and cluster centers v_j using Equations (3) and (4), respectively (Pedrycz 2013).

$$u_{ik} = \frac{1}{\sum_{j=1}^{c}\left(\frac{\|x_k - v_i\|}{\|x_k - v_j\|}\right)^{2/m-1}}, \quad i = 1, \ldots, c, \quad k = 1, \ldots, N \tag{3}$$

$$v_j = \frac{\sum_{k=1}^{N} u_{ik}^{m} x_k}{\sum_{k=1}^{N} u_{ik}^{m}}, \quad i = 1, \ldots, c, \quad k = 1, \ldots, N \tag{4}$$

The FCM clustering algorithm maximizes the membership degree of each data point close to the cluster center, while minimizing the membership degrees of the data away from the cluster center (Elbaz et al. 2019). This method allows the development of data-driven FIS using rules for defining the relationships between input and output variables (Pedrycz 2013). Each cluster represents a fuzzy rule; thus, FCM clustering results in the development of *c* number of fuzzy rules in the form of "If X is A_i, then Y is B_i."

Two types of FIS (i.e., Mamdani and Takagi-Sugeno) have been widely used in various applications. Mamdani FIS are intuitive and have better interpretability (i.e., explicit knowledge representation). On the other hand, Takagi-Sugeno FIS have capability for numeric processing (i.e., accuracy of prediction). In this study, Takagi-Sugeno FIS was used because of its superior performance in terms of accuracy, and Gaussian MFs were used for representing model input variables. Studies have indicated that Gaussian MFs are a better option because they are efficient, with higher performance in prediction for their continuity and smoothness, simplicity in representation, ease of construction using a data-driven approach, faster convergence during MF optimization, and suitability for models that seek high-control accuracy (Elbaz et al. 2019; Siraj et al. 2016).

GA optimization enables the MANFIS training to optimize the parameters of input-output in the system. At this stage, real-coded parameters are used to represent model input variables instead of the binary coded strings used in the FS stage. Model input variables are further represented by a number of parameters or MFs and thus a real-code (i.e., real numbers) that encodes the MANFIS parameters with a corresponding range of input parameter values. The learning and parameter optimization process of the MANFIS network terminates when the fitness error measure, RMSE, between two sequential iterations or the maximum 100 iterations is reduced to a satisfied level, which is the set threshold of 10^{-5}.

In summary, the 62 data instances obtained from the data preprocessing stage were used for training and testing the GA-MANFIS model. As noted, 80% (50 instances) of the dataset were used for training and the remaining 20% (12 instances) were used for testing the model; an FCM-based Takagi-Sugeno FIS was used to develop each ANFIS modal block in the MANFIS MIMO block; and real-coded GA was used to train and optimize the premise and consequent parameters. Crossover and mutation probabilities were set as 0.8 and 0.1, respectively, and a roulette wheel selection method was used.

Model verification and validation

Verification is conducted to ensure that model components work as expected (Lucko and Rojas 2010). To verify the GA-MANFIS model, all mathematical equations and components of the model, such as MATLAB codes, are checked for their correctness. Further, the model is run multiple times to check for the replicability of its results, and tracing and plot graphs are used to track changes in model variables.

Validation determines how well a model reflects a real-world system. Conceptual validity and data validity were conducted on the GA-MANFIS model. For conceptual validity, the model was based on factors identified in the literature as validated by construction experts and practitioners through a focus group. Data was validated through pilot testing a data collection protocol, following a structured data collection methodology, testing for construct validity, and testing the reliability of the data-collection measures. The GA-MANFIS model performance was evaluated by comparing the model outputs (i.e., predicted results) against the testing dataset. RMSE was used as the fitness function to check the conformity of the predicted values with the actual observed or measured values with a minimum RMSE. Additionally, sensitivity analysis was conducted to determine whether the model behaves realistically, by changing model parameters and evaluating changes in the behavior of model output.

Results and Discussion

In this case study, the proposed hybrid GA-MANFIS model was used to analyze organizational competencies and simultaneously predict multiple organizational performance metrics for a company in the construction industry.

Data preparation and feature selection

Based on respondents' replies, a total of 60 organizational competencies (32 functional and 28 behavioral competencies) common to both the senior management and staff surveys were used for model development, as shown in Table 2. Six organizational performance metrics that had sufficient data variability were considered for modeling: *Employee satisfaction*, *Customer satisfaction*, *Competitiveness*, *Quality of work*, *Safety performance*, and *Effectiveness of planning*.

As noted in the methodology, data cleaning resulted in 62 data instances. Organizational competencies and performance metrics data were characterized as having 60 input features (i.e., competencies), 6 output features (i.e., performance metrics), and 62 data instances (i.e., complete survey responses or data points). Thus, the input data matrix was 62×60, the output data matrix was 62×6, and the overall input-output MIMO system data matrix was 62×66.

Table 2. Organizational competencies.

Group	Competencies
Functional	Commitment to safety; Communications management; Construction, production, and manufacturing; Construction technology and integration management; Cooperation and coordination; Customer/stakeholder focus; Delegation; Engagement; Goal orientation; Human resources management; Interdisciplinary alignment; Interface management; Management and support of diversity; Management experience and excellence; Materials management; Operations and maintenance; Planning and organizing of tasks/activities; Process engineering management; Project change management; Project cost management; Project finance management; Project integration management; Project quality management; Project risk management; Project safety management; Project schedule management; Project scope management; Quality of work; Resource management; Staff development; Technical innovation; Technical/job knowledge
Behavioral	Ability to build trust; Achievement drive; Adaptability/flexibility; Analytical ability; Assertiveness; Attention to detail; Communication; Competitiveness; Conflict and crisis resolution / issue management; Effectiveness; Influence; Innovation; Interpersonal skills; Judgment; Leadership; Motivation/commitment; Organizational awareness and culture; Perseverance / self-regulation and control; Problem-solving; Professionalism; Reasoning; Reliability/dependability; Resourcefulness; Responsiveness; Results orientation; Strategic thinking; Teamwork; Values and ethics

Given the size of the overall input-output matrix, the dimensionality of original raw data was very high and GA-FS was conducted. Table 3 shows the best results obtained from the FCM parameter optimization and the top five parameters used in GA-FS. The FCM results indicated that the RMSE tended to be minimum when the values of m were low, irrespective of the number of clusters, especially closer to 2.

Table 4 shows the results of the GA-FS ranking based on average fitness values, which further indicate that the FCM parameters that provided the best five GA-FS results with minimum error were for $c = 6$, $m = 1.45$ and $c = 3$, $m = 1.75$, respectively, and that the best optimum parameters identified in Table 3 – for $c = 6$, $m = 2.50$ – showed poor results in terms of the GA-FS fitness function. The poor performance results from the higher value of $m = 2.50$: as the m value

Table 3. FCM parameter optimization results

Code*	*c*	*m*	Minimum RMSE	Rank
Opt_S10	6	2.50	0.037141	1
Opt_Sug7	6	1.45	0.038779	2
Opt_Sug18	7	2.55	0.040223	3
Opt_S25	3	1.75	0.042300	4
Opt_Sug11	7	2.50	0.042737	5
Opt_Sug2	7	1.90	0.047960	6
Opt_Sug8	7	2.35	0.048160	7
Opt_S2	7	1.75	0.065234	8
Opt_S16	6	1.85	0.066222	9
Opt_S6	6	2.25	0.077706	10

* FCM parameter optimization run

gets higher, the MFs will become "spiky," meaning the membership grades are equal to 1 at the prototypes/cluster centers, and the values rapidly decline when moving away from the prototypes. With minimum overlap of adjacent MFs, the process therefore provides less accurate results. GA-FS results further showed that the number of features selected was lower as values of *m* used for FS increased. Moreover, results with the best fitness values provided almost similar numbers of features. For instance, four of the top five ranked results in Table 4 selected 19 features as a representative subset of the original data, while the remaining result obtained 18 features. For model development, the result with lower value of *c* and *m* value closer to 2 was considered. Pedrycz and Gomide (2007) recommended that a value of $m = 2.00$ or closer is appropriate for the application of FCM clustering. Therefore, $c = 3$, $m = 1.75$ is the optimum FCM parameter selected for GA-MANFIS model development.

GA-FS optimization selected 19 competencies out of the original 60: *Staff development* (c_1); *Goal orientation* (c_2); *Interdisciplinary alignment* (c_3); *Commitment to safety* (c_4); *Construction, production, and manufacturing* (c_5); *Project safety management* (c_6); *Project cost management* (c_7); *Project procurement management* (c_8); *Engagement* (c_9); *Ability to build trust*

Table 4. GA-FS results for the optimized FCM parameters

FCM parameter optimization values				**GA-FS result values**				
Code	*c*	*m*	Min. RMSE	Population	Selected features (no.)	Average fitness (RMSE)	Best fitness (RMSE)	Rank based on average fitness
Opt_S10	6	2.5	0.037141	50	18	0.047080	0.043382	14
				60	18	0.046710	0.043402	11
				80	16	0.046881	0.043083	12
				100	15	0.046983	0.043477	13
Opt_Sug7	6	1.45	0.038779	50	22	0.043843	0.040428	8
				60	**19**	**0.040227**	**0.037340**	**2**
				80	**18**	**0.040837**	**0.036858**	**3**
				100	**19**	**0.038646**	**0.035067**	**1**
Opt_Sug18	7	2.55	0.040223	50	15	0.045183	0.049491	9
				60	16	0.047307	0.043375	15
				80	17	0.046513	0.042393	10
				100	16	0.048640	0.044637	19
Opt_S25	3	1.75	0.042300	50	19	0.042157	0.040530	7
				60	**19**	**0.041731**	**0.039895**	**4**
				80	22	0.041968	0.039813	6
				100	**19**	**0.041967**	**0.040344**	**5**
Opt_Sug11	7	2.50	0.042737	50	15	0.048054	0.043580	17
				60	16	0.049209	0.045792	20
				80	16	0.048334	0.044543	18
				100	15	0.047802	0.042622	16

(c_{10}); *Organizational culture* (c_{11}); *Judgment* (c_{12}); *Values and ethics* (c_{13}); *Conflict resolution* (c_{14}); *Results orientation* (c_{15}); *Influence* (c_{16}); *Communications* (c_{17}); *Motivation* (c_{18}); and *Perseverance* (c_{19}). These 19 organizational competencies were used as input variables for model development, being the best subset of the original organizational competencies, and thus enabling development of a model that provided high accuracy.

GA-MANFIS model development

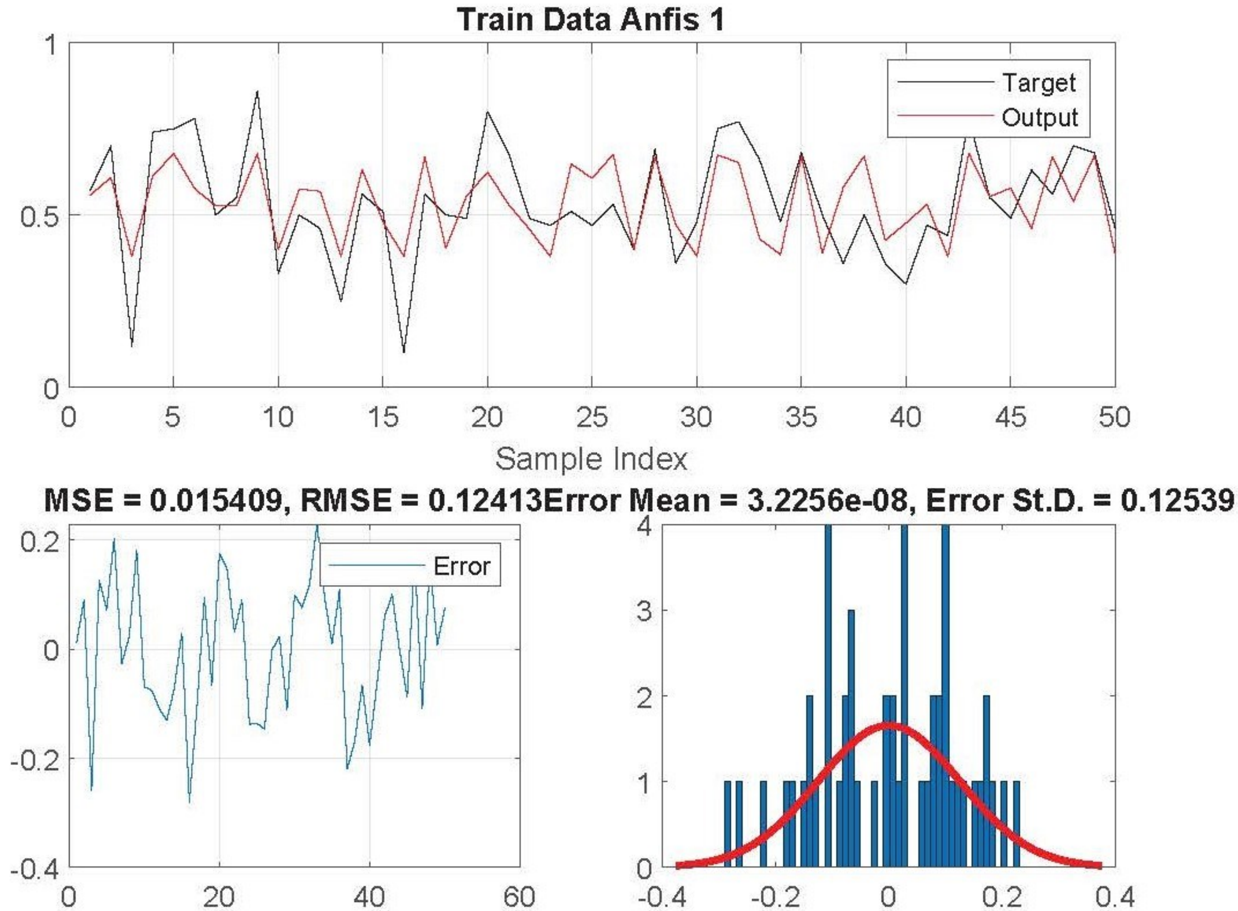
The 19 organizational competencies and 7 organizational performance metrics were the model input and output variables, respectively. In this study, 36 GA-MANFIS models were implemented for different GA parameters, such as population size and number of generations (iterations). For population size, values of 50, 60, 80, and 100 were tested; the number of generations used to run the model were 25, 50, and 100. Similarly, different initial FIS with 3, 5, and 7 clusters for developing the model were tested. The best optimized model is defined as the one that predicts the results of the test data with highest accuracy (i.e., minimum RMSE). The GA-MANFIS with a population size of 50, 100 generations, and 3 clusters was found to be the optimal model. The results indicate that increasing the number of clusters while developing a model with limited data reduces the model performance. Moreover, an investigation of the MFs obtained for models with $c = 7$ were very close to one another and could therefore be merged to obtain better model performance. Thus, models with fewer clusters provided the best result. Table 5 presents results of the most optimal GA-MANFIS model, which the values in Table 5 indicate can predict four of the seven organizational performance metrics with high accuracy. The highest prediction accuracy for the testing data with a minimum RMSE = 0.13784 was obtained for *Overall organizational performance*. The optimal GA-MANFIS model also predicted *Customer satisfaction*, *Employee satisfaction*, and *Effectiveness of planning* with a higher prediction accuracy. The prediction performance of the model for *Quality of work* was low compared to the other metrics, with RMSE = 0.32253. However, the predictions for *Competitiveness* and *Safety performance* showed better accuracy than *Quality of work*, with RMSE values of 0.24507 and 0.27596, respectively.

Table 5. Results of optimal GA-MANFIS model outputs

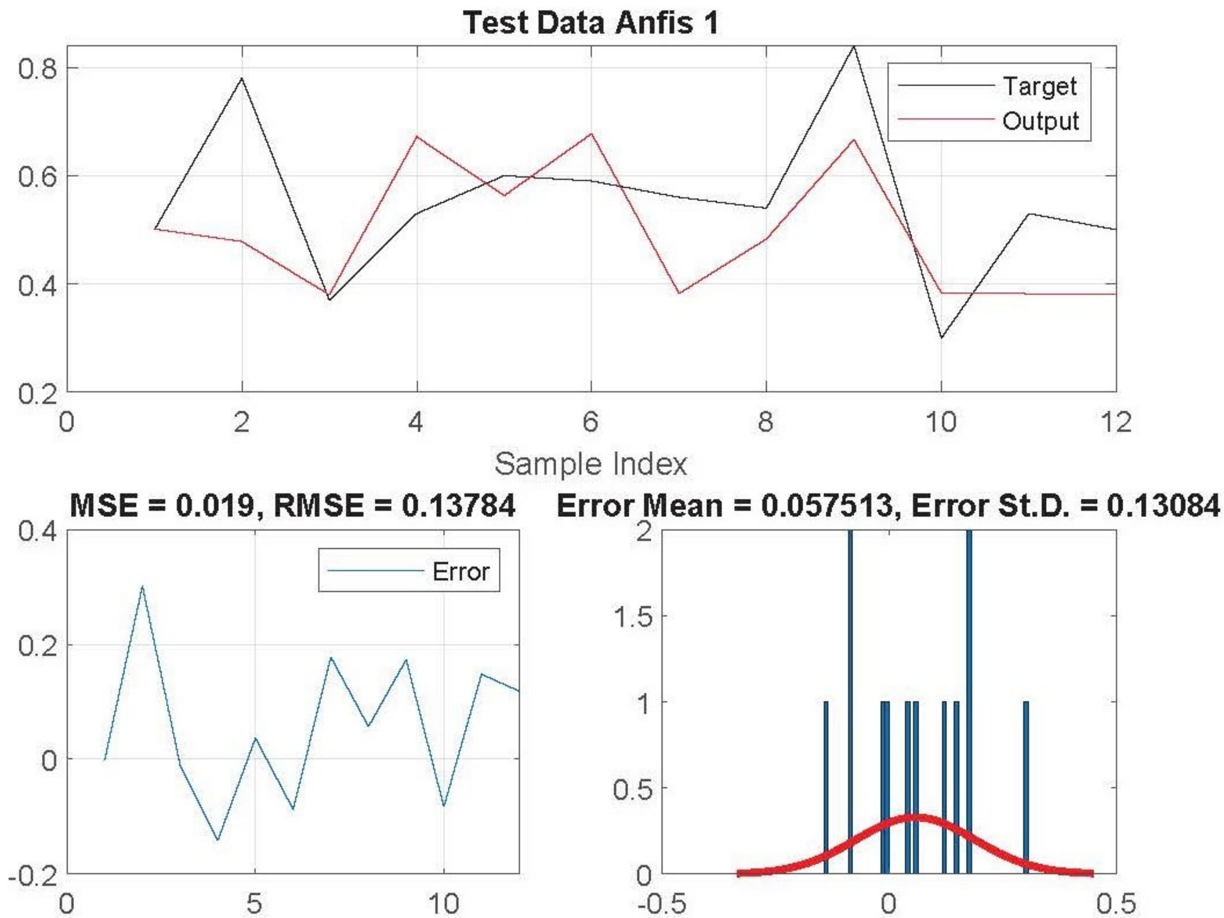
Organizational performance metrics	Training data			Testing data		
	RMSE	Error mean	Error st. d.	RMSE	Error mean	Error st. d.
Overall organizational performance	0.12413	3.22E-8	0.12539	**0.13784**	0.05751	0.13084
Employee satisfaction	0.20037	3.52E-8	0.20240	**0.18901**	0.00251	0.19740
Customer satisfaction	0.25376	0.09181	0.23896	**0.18078**	0.15063	0.10441
Competitiveness	0.21282	3.00E-8	0.21498	0.24507	0.11347	0.22688
Quality of work	0.41657	-0.27040	0.32010	0.32253	-0.12542	0.31037
Safety performance	0.29406	0.19591	0.22151	0.27596	0.13158	0.25336
Effectiveness of planning	0.23141	2.75E-8	0.22376	**0.19329**	-0.06933	0.18845

Comparison between the actual and predicted values of performance metrics by the best optimal GA-MANFIS model (i.e., with population = 50 and generations = 100) is depicted in Figure 5. As noted in the methodology, each ANFIS modal block (Figure 4) corresponds to the prediction of a single output. For instance, Figure 5 depicts the ANFIS 1 prediction of *Overall organizational performance* with RMSE = 0.26406, error mean = 0.057513, and standard deviation = 0.13084 for the training data. The prediction for testing data provided RMSE = 0.13784, error mean = 0.057513, and standard deviation = 0.13084. Figure 6 presents the prediction of ANFIS 2 for *Employee satisfaction* with RMSE = 0.18901, mean error = 0.0025, and error standard deviation = 0.1974 for the testing data. The plots of results showed a relatively good fit both for the training and testing data. Graphical methods such as residual analysis are advantageous for illustrating the relationship between model and data, and numerical or statistical methods (e.g., sum of square error, mean square error or residual mean square, and RMSE) for model validation. As such, results with mean square error (MSE) or RMSE value closer to 0 (zero) indicate a good fit that is useful for prediction. For instance, Figure 6(b) indicates that the model output value for *Employee satisfaction* follows the behavior of the target or actual values of the testing data, which appears to have a better fit visually when compared to Figure 5(b). However, the residual analysis plot indicates that the plots of results presented in Figure 6(b) with a residual mean square error MSE = 0.035725 have higher variability compared to Figure 5(b) with MSE = 0.019. Thus, the plots of results shown in Figure 5 have a good fit for prediction, and the GA-MANFIS is able to predict new observations of *Overall organizational performance* with a relatively high certainty compared to *Employee satisfaction*. Prediction for the remaining performance metrics was implemented in the same manner where the results showed a good fit, that is, followed the pattern of the actual target values.

In summary, the performance of the proposed GA-MANFIS was found to be excellent compared with the target goal. The performance curves, or graphical plots, for training and testing (Figures 5 and 6) are almost identical, which indicates that the model output shows a good fit that follows the patterns of the target results (actual values). Furthermore, the GA-MANFIS showed good performance in predicting four of the seven organizational performance metrics including *Overall organizational performance*, *Employee satisfaction*, *Competitiveness*, and *Effectiveness of planning*. The relatively poor fit for *Customer satisfaction*, *Quality of work*, and *Safety performance* results from the lack of adequate variability in the data.

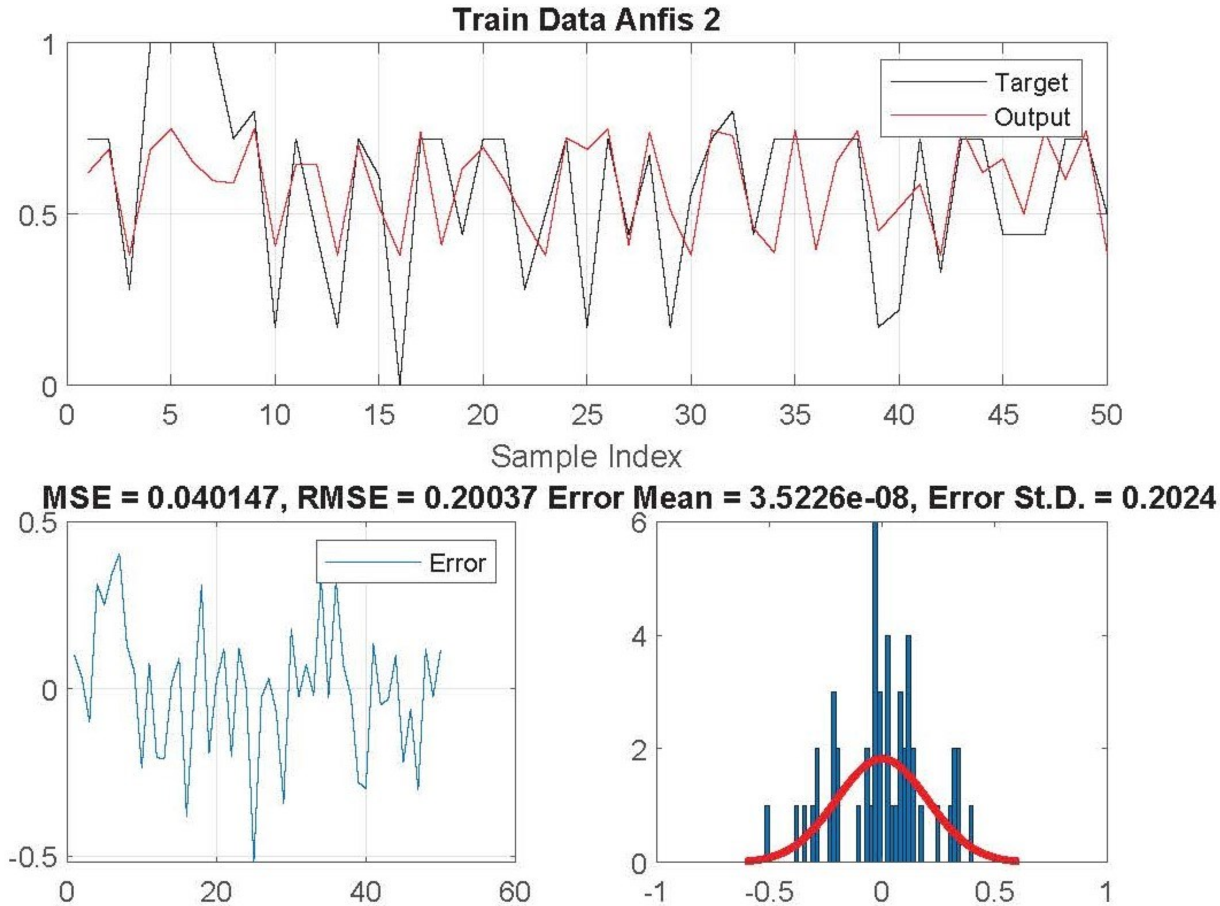

Train Data Anfis 1
Target
Output
Sample Index
MSE = 0.015409, RMSE = 0.12413Error Mean = 3.2256e-08, Error St.D. = 0.12539
Error

(a)

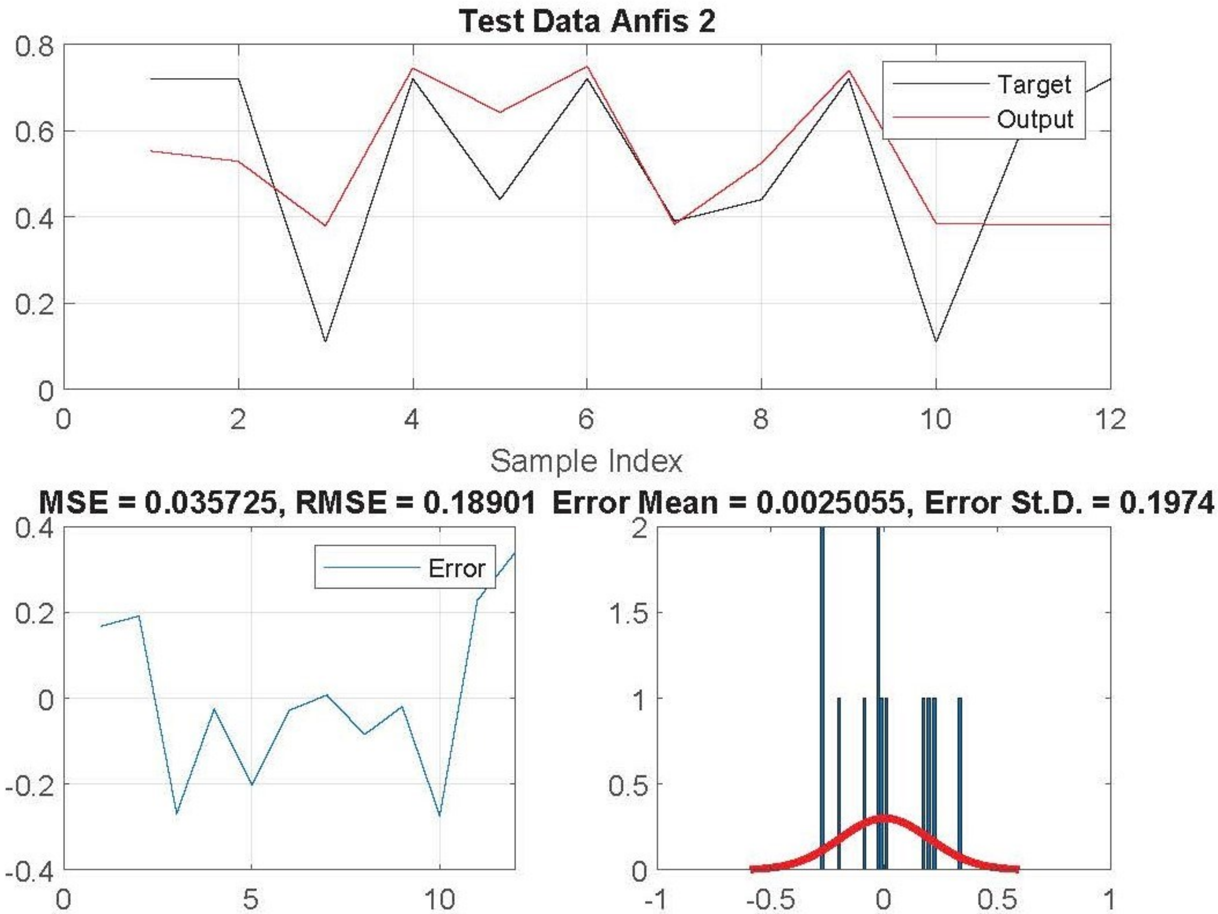


(b)

Fig. 5. Comparison of target, output, mean square error (MSE), root mean square error (RMSE), mean error, and standard deviation (st. d.) for *Overall organizational performance* – **(a)** training, **(b)** testing.


Train Data Anfis 2
Target
Output
Sample Index
MSE = 0.040147, RMSE = 0.20037 Error Mean = 3.5226e-08, Error St.D. = 0.2024
Error

(a)



(b)

Fig. 6. Comparison of target, output, MSE, RMSE, mean error, and st. d. for *Employee satisfaction* – **(a)** training, **(b)** testing.

GA-MANFIS model verification and validation

The sensitivity analysis results showed that the number of generations insignificantly effects model output, but changes in GA optimization population size significantly affects model outputs. Table 6 presents the sensitivity analysis of the optimal model with respect to changes in population size of the GA optimization.

The values in Table 6 reveal that as the model's population increases, prediction accuracy decreases and the search space for GA to find an optimal solution becomes large, which makes the optimization processes too complicated and much too time consuming. The model prediction patterns follow a similar trend to that of the optimal model, although with reduced prediction accuracy. Thus, population size is an important factor that needs to be chosen carefully in lieu of the data availability for model development.

Comparison of the seven independent MISO GA-ANFIS models developed for each organizational performance metric and overall organizational performance are shown in Table 7 and indicate that the GA-MANFIS performs better than the GA-ANFIS model in predicting five of the seven organizational performance metrics. For instance, GA-MANFIS showed a significant 27.62% improvement in prediction accuracy for *Effectiveness of planning* and 22.38% improvement for *Overall organizational performance*. GA-MANFIS obtained a better performance with 7.25% improvement in prediction accuracy for *Safety performance*, 5.16% for *Quality of work*, and 5.06% for *Employee satisfaction*. According to Benmiloud (2010), the increase in the number of weights of the GA-MANFIS model allows improvement of the prediction accuracy, explained by smaller prediction errors (i.e., RMSE). For instance, the RMSE of predicting *Overall organizational performance* is reduced from 0.16855 (GA-ANFIS) to 0.13784 (GA-MANFIS). Moreover, GA-MANFIS showed a significant prediction improvement

Table 6. Sensitivity analysis and comparison of best performing models

Organizational performance metrics	**RMSE**							
	Pop = 50		*Pop = 60*		*Pop = 80*		*Pop = 100*	
	Training	Testing	Training	Testing	Training	Testing	Training	Testing
Overall organizational performance	0.12413	0.13784	0.13035	0.11251	0.12925	0.09793	0.12947	0.09934
Employee satisfaction	0.20037	0.18901	0.19356	0.23850	0.18572	0.23174	0.19707	0.21448
Customer satisfaction	0.25376	0.18078	0.23836	0.25465	0.25457	0.25400	0.26481	0.26726
Competitiveness	0.21282	0.24507	0.22140	0.16173	0.22610	0.16677	0.20713	0.23690
Quality of work	0.41657	0.32253	0.40764	0.46990	0.37794	0.44287	0.39198	0.39410
Safety performance	0.29406	0.27596	0.27086	0.31270	0.31124	0.25224	0.31774	0.30462
Effectiveness of planning	0.23141	0.19329	0.23868	0.37885	0.27113	0.28610	0.27108	0.20100

for *Effectiveness of planning*, with RMSE = 0.19329 compared to RMSE = 0.24667 for GA-ANFIS. However, GA-ANFIS showed a better performance for *Competitiveness* with 16.07% improvement of prediction accuracy and 4.04% improvement for *Customer satisfaction*. The reason GA-ANFIS performs better in predicting *Customer satisfaction* and *Competitiveness* may be attributable to the nature of the data. The normalized values for *Customer satisfaction* and *Competitiveness* used for model training and testing have a greater numerical range compared to the remaining performance metrics. For example, most of the normalized values for *Competitiveness* include 0.00, 0.33, 0.67, and 1.00, compared to *Overall organizational performance* with better data variability (i.e., 0, 0.2, 0.25, 0.35, 0.45, 0.5, 0.55, 0.6, 07, 0.75, and 1). NFS models with a complex network such as GA-MANFIS can produce random oscillations between the training points to comply with a great numerical range or fast data variations where the training/optimization algorithm tends to produce high variance between target and predicted values (Carrano et al. 2008). So, although it needs further investigation and verification, the possible explanation for a better performance of GA-ANFIS in predicting *Customer satisfaction* and *Competitiveness* is owing to its less complex network compared to GA-MANFIS.

Table 7. Comparison of GA-ANFIS and GA-MANFIS model performance

Organizational performance metrics	RMSE for testing data		Prediction improvement (%)	
	GA-ANFIS	GA-MANFIS	GA-ANFIS	GA-MANFIS
Overall organizational performance	0.16855	0.13784	-	22.28
Employee satisfaction	0.19885	0.18901	-	5.06
Customer satisfaction	0.17348	0.18078	4.04	-
Competitiveness	0.20569	0.24507	16.07	-
Quality of work	0.33917	0.32253	-	5.16
Safety performance	0.29598	0.27596	-	7.25
Effectiveness of planning	0.24667	0.19329	-	27.62

Overall, the GA-MANFIS model showed a better prediction performance than the corresponding GA-ANFIS model. The higher prediction accuracy obtained from GA-MANFIS in predicting multiple organizational performance metrics allows construction industry organizations to determine realistic organizational performance by analyzing their competencies. In addition, GA-MANFIS also had greater capability to analyze multiple inputs (competencies), relate them to organizational performance metrics, and predict multiple organizational performance. Moreover, the GA-MANFIS model provides organizations and construction practitioners with insight into targeted areas for future investment and expansion strategies for improving organizational performance, which further helps them to make the best decisions. Thus, the proposed GA-MANFIS model has a great advantage over GA-ANFIS in that it can predict multiple organizational performance metrics at once rather than developing an independent model for each output.

Conclusions and Recommendations for Future Work

NFS models, specifically ANFIS, have previously been used to model real-world CEM problems because of their effective characteristics for solving nonlinear, dynamic, and complex problems. However, the application of ANFIS models have some limitations in handling multiple outputs. For example, the nonlinear multiple input-output relationships of real-world CEM problems inherently make them MIMO problems. To address this limitation and improve effectiveness in handling multiple outputs, this paper proposed a novel methodology to develop a hybrid GA-MANFIS model for application in CEM problems that was then used multiple organizational competencies as input variables to predict multiple organizational performance metrics. The proposed model was validated based on data collected from a company active in various industrial projects. The results showed that the optimal model for predicting organizational performance metrics with minimum RMSE is the GA-MANFIS model with 3 clusters, a population size of 50, and 100 generations. The proposed GA-MANFIS model showed a good performance with the highest accuracy in predicting multiple organizational performance metrics simultaneously. Sensitivity analysis identified the main parameters that affect model outputs. Accordingly, population size was found to have a significant impact on model outputs. Furthermore, Comparing GA-MANFIS and GA-ANFIS model outputs showed that the GA-MANFIS model performed better in predicting multiple organizational performance metrics simultaneously (Table 7) than individual, independent GA-ANFIS models for each performance metric.

This paper makes three main contributions. First, it provides a novel methodology for developing GA-MANFIS models that can handle MIMO systems inherent in construction processes and practices, thus addressing the issue of handling multiple outputs common in real-world CEM problems. Second, the proposed GA-MANFIS model has the capability to relate multiple construction organizational competencies to multiple organizational performance metrics, creating a more accurate prediction model than conceptual and regression models used in previous construction research. Third, this paper provides a GA-FS approach that is vital not only for dimensionality reduction, but also for identifying organizational competencies influencing performance by reducing model complexity and improving model prediction performance to obtain good results with high accuracy. By uniquely integrating these computing techniques, the proposed GA-MANFIS model enables CEM organizations to identify and evaluate competencies that have significant impact on performance as well as predict multiple organizational performances simultaneously. Moreover, the GA-MANFIS modeling approach does not require manual configuration; hence, it can serve as a reference for construction researchers for developing concise and accurate models that can predict multiple outputs, such as risk, cost, and schedule management, for other CEM disciplines. Additionally, the proposed GA-MANFIS modeling methodology in this paper is generalizable and can be adapted to different construction contexts for different industry groups such as owners, consultants, and contractors.

Future research will focus on exploring different evolutionary algorithms other than GA, such as PSO, ABC, and ACO, to train and optimize the GA-MANFIS model. The performance of the GA-MANFIS model optimized with various evolutionary algorithms will help researchers and practitioners compare performance of the model prediction accuracy and select the best performing model for a specific construction problem. Furthermore, the methodology will be extended to develop similar models applicable to other construction contexts. Data from more companies will be collected to expand the scope of applicability of the developed GA-MANFIS methodology, provide more insight into the most critical organizational competencies influencing performance, and analyze the relationship between competency and performance at the organization level. The GA-MANFIS model is developed at the organizational level; hence, it is a higher-level model. Therefore, it will be customized to other levels, such as business unit/department level, project level, and/or construction crew level.

Acknowledgements

This work was funded by the Natural Sciences and Engineering Research Council of Canada Industrial Research Chair in Strategic Construction Modeling and Delivery (NSERC IRCPJ 428226–15), which is held by Dr. Aminah Robinson Fayek. The authors gratefully acknowledge the support and data provided by industry partners and all personnel who participated in this study.

Data Availability Statement

All data, models, and code generated or used during the study appear in the submitted article.

References

Abd Elaziz, M., A. A. Ewees, and Z. Alameer. 2019. “Improving adaptive neuro-fuzzy inference system based on a modified salp swarm algorithm using genetic algorithm to forecast crude oil price.” *Nat. Res. Resear.*, 29 (4): 2671–2686. https://doi.org/10.1007/s11053-019-09587-1.

Acampora, G., W. Pedrycz, and A. V. Vasilakos. 2014. “Efficient modeling of MIMO systems through timed automata based neuro-fuzzy inference engine.” *Int. J. Approx. Reason.*, 55: 1336–1356. https://doi.org/10.1016/j.ijar.2014.02.003.

Adeli, H., and X. Jiang. 2003. “Neuro-fuzzy logic model for freeway work zone capacity estimation.” *J. Transp. Eng.*, 129 (5): 484–493. https://doi.org/10.1061/(ASCE)0733947X(2003)129:5(484).

Afshari, A. 2017. “Methods for selection of construction project manager: Case study.” *J. Constr. Eng. Manage.*, 143 (12): 06017003. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001400.

Agah, A., and N. Soleimanpourmoghadam. 2020. “Design and implementation of heavy metal prediction in acid mine drainage using multi-output adaptive neuro-fuzzy inference systems (ANFIS) – A case study.” *Int. J. Min. Geo-Eng.*, 54 (1): 59–64. http://doi.org/10.22059/ijmge.2019.278558.594794.

Ahadzie, D. K., D. G. Proverbs, P. O. Olomolaiye, and N. Ankrah. 2009. “Towards developing competency-based measures for project managers in mass house building projects in developing countries.” *Constr. Manage. Econ.*, 27: 89–102. https://doi.org/10.1080/01446190802621028.

Ahadzie, D. K., D. G. Proverbs, P. O. Olomolaiye, and I. Sarkodie-Poku. 2014. “Competencies required of project managers at the design phase of mass house building projects.” *Int. J. Proj. Manage.*, 32: 958–969. https://doi.org/10.1016/j.ijproman.2013.10.015.

Altuncan, İ. Ü., and A. M. Tanyer. 2018. “Context-dependent construction conflict management performance analysis based on competency theory.” *J. Constr. Eng. Manage.*, 144 (12): 04018112. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001581.

Bayram, S., and S. Al-Jibouri. 2016. “Efficacy of estimation methods in forecasting building projects’ costs.” *J. Constr. Eng. Manage.*, 142 (11): 05016012. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001183.

Benmiloud, T. 2010. “Multioutput adaptive neuro-fuzzy inference system.” In *Proc., 11th WSEAS International Conferences on Neural Networks, Evolutionary Computing, and Fuzzy Systems*, 94–98. Stevens Point, Wisconsin, USA: World Scientific and Engineering Academy and Society. ISBN: 978-960-474-195-3.

Bolivar-Ramos, M. T., V. J. Garcia-Morales, and E. Garcia-Sanchez. 2012. “Technological distinctive competencies and organizational learning: Effects on organizational innovation to improve firm performance.” *J. Eng. Technol. Manage.*, 29: 331–357. https://doi.org/10.1016/j.jengtecman.2012.03.006.

Carrano, E. G., R. H. C. Takahashi, W. M. Caminhas, and O. M. Neto. 2008. “A genetic algorithm for multiobjective training of ANFIS fuzzy networks.” In *Proc., IEEE Congress on Evolutionary Computation*, 3259–3265. https://doi.org/10.1109/CEC.2008.4631239.

Cheng, C.-B, C.-J. Cheng, and E. S. Lee. 2002. “Neuro-fuzzy and GA in multiple response optimization.” *Comp. Math. Appl.*, 44: 1503–1514. https://doi.org/10.1016/S0898-1221(02)00274-2.

Cheng, E. W. L., H. Li, and P. Fox. 2007. “Job performance dimensions for improving final project outcomes.” *J. Constr. Eng. Manage.*, 133 (8): 592–599. https://doi.org/10.1061/(ASCE)0733-9364(2007)133:8(592).

Cheng, M., and A. F. V. Roy. 2010. “Evolutionary fuzzy decision model for construction management using support vector machine.” *Expert Syst. Appl.*, 37: 6061–6069. https://doi.org/10.1016/j.eswa.2010.02.120.

Cheng, M., H. Tsai, and E. Sudjono. 2012. “Evolutionary fuzzy hybrid neural network for dynamic project success assessment in construction industry.” *Autom. Constr.*, 21: 46–51. https://doi.org/10.1016/j.autcon.2011.05.011.

Cheng, M., D. K. Wibowo, D. Prayogo, and A. F. V. Roy. 2015. “Predicting productivity loss caused by change orders using the evolutionary fuzzy support vector machine inference model.” *J. Civ. Eng. Manage.*, 21 (7): 881–892. https://doi.org/10.3846/13923730.2014.893922.

Dainty, A. R. J., M.-I Cheng, and D. R. Moore. 2005. “Competency-based model for predicting construction project managers’ performance.” *J. Manage. Eng.*, 21 (1): 2–9. https://doi.org/10.1061/(ASCE)0742-597X(2005)21:1(2).

Das, R. D., and S. Winter. 2016. “Detecting urban transport modes using a hybrid knowledge driven framework from GPS trajectory.” *Int. J. Geo-Info.*, 5 (11): 207. http://doi.org/10.3390/ijgi5110207.

Elbaz, K., S. Shen, W. Sun, Z. Yin, and A. Zhou. 2020. “Prediction model of shield performance during tunneling via incorporating improved particle swarm optimization into ANFIS.” *IEEE Access*, 8: 39659–39671. https://doi.org/10.1109/ACCESS.2020.2974058.

Elbaz, K., S. Shen, A. Zhou, D. Yuan, and Y. Xu. 2019. “Optimization of EPB shield performance with adaptive neuro-fuzzy inference system and genetic algorithm.” *Appl. Sci.*, 9 (4): 780. https://doi.org/10.3390/app9040780.

Elmousalami, H. H. 2020. "Artificial intelligence and parametric construction cost estimate modeling: State-of-the-art review." *J. Constr. Eng. Manage.*, 146 (1): 03119008. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001678.

Elwakil, E., M. Ammar, T. Zayed, M. Mahmoud, A. Eweda, and I. Mashhour. 2009. "Investigation and modeling of critical success factors in construction organizations." In *Proc., ASCE Construction Research Congress 2009*, edited by S. T. Ariaratnam and E. M. Rojas, 350–359. Construction Institute of ASCE; University of Washington. https://doi.org/10.1061/41020(339)36.

Fattahi, H., A. Agah, and N. Soleimanpourmoghadam. 2018. "Multi-output adaptive neuro-fuzzy inference system for prediction of dissolved metal levels in acid rock drainage: A case study." *J. AI and Data Min.*, 6 (1): 121–132. https://doi.org/10.22044/jadm.2017.975.

Georgy, M. E., L. Chang, and L. Zhang. 2005. "Prediction of engineering performance: A neurofuzzy approach." *J. Constr. Eng. Manage.*, 131 (5): 548–557. https://doi.org/10.1061/(ASCE)0733-9364.

Gunduz, M., and H. A. Elsherbeny. 2021. "Critical assessment of contract administration using multidimensional fuzzy logic approach." *J. Constr. Eng. Manage.*, 147 (2): 04020162. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001975.

Jin, X. 2010. "Neurofuzzy decision support system for efficient risk allocation in public-private partnership infrastructure projects." *J. Comput. Civ. Eng.*, 24 (6): 525–538. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000058.

Jin, X. 2011. "Model for efficient risk allocation in privately financed public infrastructure projects using neuro-fuzzy techniques." *J. Constr. Eng. Manage.*, 137 (11): 1003–1014. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000365.

Kaveh, A., S. M. Hamze-Ziabari, and T. Bakhshpoori. 2018. “Feasibility of PSO-ANFIS-PSO and GA-ANFIS-GA models in prediction of peak ground acceleration.” *Int. J. Optim. Civ. Eng.*, 8 (1): 1–14. http://ijoce.iust.ac.ir/article-1-321-en.html.

Kumar, R., and N. R. J. Hynes. 2020. “Prediction and optimization of surface roughness in thermal drilling using integrated ANFIS and GA approach.” *Eng. Sci. Technol.*, 23 (1): 30–41. https://doi.org/10.1016/j.jestch.2019.04.011.

Levenson, A. R., W. A. van der Stede, and S.G. Cohen. 2006. “Measuring the relationship between managerial competencies and performance.” *J. Manage.*, 32 (3): 360–380. https://doi.org/10.1177/0149206305280789.

Liang, C., Y. Lin, and H. Huang, 2013. “Effect of core competence on organizational performance in an airport shopping center.” *J. Air Transp. Manage.*, 31: 23–26. https://doi.org/10.1016/j.jairtraman.2012.11.005.

Ling, F. Y. Y. 2004. “How project managers can better control the performance of design-build projects.” *Int. J. Proj. Manage.*, 22: 477–488. https://doi.org/10.1016/j.ijproman.2003.09.003.

Ling, Y. Y. 2002. “Model for predicting performance of Architects and Engineers.” *J. Constr. Eng. Manage.*, 128 (5): 446–455. https://doi.org/10.1061/(ASCE)0733-9364(2002)128:5(446).

Liu, J. Y., H. H. Chen, J. J. Jiang, and G. Klein. 2010. “Task completion competency and project management performance: The influence of control and user contribution.” *Int. J. Proj. Manage.*, 28: 220–227. https://doi.org/:10.1016/j.ijproman.2009.05.006.

Lucko, G., and E. M. Rojas. 2010. “Research validation: Challenges and opportunities in the construction domain.” *J. Constr. Eng. Manage.*, 136 (1): 127–135. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000025.

Malik, A. H., and F. Arshad. 2011. "Design of multi-input multi-output hybrid adaptive neuro-fuzzy intelligent system for primary pressure control system of pressurized heavy water reactor." *Proc. the Pakistan Academy of Sciences*, 48 (2): 65–77. ISSN: 0377–2969.

Moon, S., and A. M. Chowdhury. 2021. "Utilization of prior information in neural network training for improving 28-day concrete strength prediction." *J. Constr. Eng. Manage.*, 147 (5): 04021028. https://doi.org 10.1061/(ASCE)CO.1943-7862.0002047.

Nayak, V., Y. P. Banjare, and M. F. Qureshi. 2015. "Multioutput adaptive neuro-fuzzy inference system-based modeling of heated catalytic converter performance." *Int. J. Innov. Resear. Sci. Eng. Tech.*, 4 (2): 604–615. https://www.ijirset.com/upload/2015/february/80_27_Multioutput.pdf.

Nazari, A., and J. G. Sanjayan. 2015. "Modeling of compressive strength of geopolymers by a hybrid ANFIS-ICA approach." *J. Mater. Civ. Eng.*, 27 (5): 04014167. https://doi.org/10.1061/(ASCE)MT.1943-5533.0001126.

Omar, M. N., and A. R. Fayek. 2016. "Modeling and evaluating construction project competencies and their relationship to project performance." *Autom. Constr.*, 69: 115–130. https://doi.org/10.1016/j.autcon.2016.05.021.

Pedrycz, W. 2013. *Granular computing: Analysis and design of intelligent systems*. Boca Raton, Florida, USA: CRC Press.

Pedrycz, W., and F. Gomide. 2007. *Fuzzy systems engineering: Toward human-centric computing.* Hoboken, New Jersey, USA: John Wiley & Sons.

Poveda, C. A., and A. R. Fayek. 2009. "Predicting and evaluating construction trades foremen performance: Fuzzy logic approach," *J. Constr. Eng. Manage.*, 135 (9): 920–929. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000061.

Rashidi, A., F. Jazebi, and I. Brilakis. 2011. “Neurofuzzy genetic system for selection of construction project managers.” *J. Constr. Eng. Manage.*, 137 (1): 17–29. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000200.

Shahhosseini, V., and M. Sebt. 2011. “Competency-based selection and assignment of human resources to construction projects.” *Scientia Iranica*, 18 (2): 163–180. https://doi.org/10.1016/j.scient.2011.03.026.

Shahtaheri, M., H. Nasir, and C. T. Haas. 2015. Setting baseline rates for on-site work categories in the construction industry.” *J. Constr. Eng. Manage.*, 141 (5): 04014097, https://doi.org/10.1061/(ASCE)CO.1943-7862.0000959.

Siraj, N. B., A. R. Fayek, and A. A. Tsehayae. 2016. “Development and optimization of artificial intelligence-based concrete compressive strength predictive models.” *Int. J. Str. Civ. Eng. Resear.*, 5 (3): 156–167. https://doi.org/10.18178/ijscer.5.3.156-167.

Suhairom, N., A. H. Musta’amal, N. F. M. Amin, and N. K. A. Johari. 2014. “The development of competency model and instrument for competency measurement: The research methods.” *Procedia - Social Behav. Sci.*, 152: 1300–1308. https://doi.org/10.1016/j.sbspro.2014.09.367.

Tahmasebi, P., and A. Hezarkhani. 2012. “A hybrid neural networks-fuzzy logic-genetic algorithm for grade estimation.” *Comput. Geosci.*, 42: 18–27. https://doi.org/10.1016/j.cageo.2012.02.004.

Tayfur, G., T. K. Erdem, and Ö. Klrca. 2014. “Strength prediction of high-strength concrete by fuzzy logic and artificial neural networks.” *J. Mater. Civ. Eng.*, 26 (11): 04014079. https://doi.org/10.1061/(ASCE)MT.1943-5533.0000985.

Tiruneh, G. G., and A. R. Fayek. 2018. “A framework for modeling organizational competencies and performance.” In *Proc., ASCE Construction Research Congress 2018*, edited by C. Wang,

C. Harper, Y. Lee, R. Harris, and C. Berryman, 712–722. https://doi.org/10.1061/9780784481271.069.

Tiruneh, G. G., and A. R. Fayek. 2019. "Feature selection for construction organizational competencies impacting performance." In *Proc., FUZZ-IEEE 2019 International Conference on Fuzzy Systems*, 1–5. IEEE. https://doi.org/10.1109/FUZZ-IEEE.2019.8858820.

Tiruneh, G. G., and A. R. Fayek. 2021. "Competency and performance measures for organizations in the construction industry." *Can. J. Civ. Eng.* 48: 716–728. http://doi.org/10.1139/cjce-2019-0769.

Tiruneh, G. G., A. R. Fayek, and V. Sumati. 2020. "Neuro-fuzzy systems in construction engineering and management research." *Autom. Constr.*, 119: 103348. https://doi.org/10.1016/j.autcon.2020.103348.

Tokede, O., D. Ahiaga-Dagbui, S. Smith, and S. Wamuziri. 2014. "Mapping relational efficiency in neuro-fuzzy hybrid cost models." In *Proc., ASCE Construction Research Congress 2014*, edited by D. Castro-Lacouture, J. Irizarry, and B. Ashuri, 1458–1467. https://doi.org/10.1061/9780784413517.149.

Yuan, Z., L. Wang, and X. Ji. 2014. "Prediction of concrete compressive strength: Research on hybrid models genetic based algorithms and ANFIS." *Adv. Eng. Software*, 67: 156–163. http://doi.org/10.1016/j.advengsoft.2013.09.004.