Modified Mamdani-fuzzy inference system for predicting the cost overrun of construction projects

Al-Nahhas, Yaman Saeid; Hadidi, Laith A.; Islam, Muhammad Saiful; Skitmore, Martin; Abunada, Ziyad

Published in:
Applied Soft Computing

DOI:
10.1016/j.asoc.2023.111152

Applied Soft Computing Journal 151 (2024) 111152

Contents lists available at ScienceDirect

Applied Soft Computing

journal homepage: www.elsevier.com/locate/asoc

Modified Mamdani-fuzzy inference system for predicting the cost overrun of construction projects

Yaman Saeid Al-Nahhas [a,1], Laith A. Hadidi [a,2], Muhammad Saiful Islam [b,*,3], Martin Skitmore [c,4], Ziyad Abunada [d,5]

[a] Department of Architectural Engineering and Construction Management, King Fahd University of Petroleum and Mineral, Dhahran 31261, Saudi Arabia
[b] Project Management Discipline, School of Engineering and Technology, Central Queensland University, 120 Spencer Steet, Melbourne, VIC 3000, Australia
[c] Faculty of Society and Design, Bond University, Robina, QLD 4226, Australia
[d] College of Project Management, Built Environment, Asset & Maintenance Management, School of Engineering and Technology, Central Queensland University, 120 Spencer Steet, Melbourne, VIC 3000, Australia

HIGHLIGHTS

- Robust Mamdani-type fuzzy inference model for predicting cost overrun amount.
- A small group of experts in a project can assess the cost overrun amount.
- The model can rank critical risks and predict cost overrun together.
- Only a factor's occurrence probability is required as input predicting cost overrun.
- 40 factors are listed to comprehensively understand cost overruns in large projects.

ARTICLE INFO

Keywords:
Cost overrun
Prediction
Construction project
Expert judgment
Fuzzy logic
Risk severity

ABSTRACT

Cost overruns are a common worldwide problem in the construction industry; improved proactive risk management and cost control are much needed. Several models have been proposed, but all have weaknesses, particularly in data demands and the severity of critical risks or uncertainties associated with expert judgment. In response, this study develops a new 3-part model based on the Mamdani-type fuzzy inference system (FIS) to predict the cost overrun of construction projects. The first part assesses the weight of each expert, evaluating the severity of cost overrun factors. The second part contains a list of 40 *in-built* cost overrun factors and their degree of severity, while the third part establishes the relationships of every factor's occurrence probability and severity to predict the cost overrun of *a specific project*. The severity of each factor is assessed based on a survey of 31 randomly selected experts in the Saudi Arabian construction industry. The model is demonstrated on two completed projects in Saudi Arabia. For each project, this involves a group of project-based experts rating the probability of occurrence of each factor on that project and applying this to the factor severity list to obtain a predicted cost overrun (PCO) for the whole project. The model is validated for robustness by sensitivity analysis comparing the predicted and actual whole project cost overrun and shown to be of practical value in assessing critical risks and predicting the likely amount of cost overrun. The model is equally applicable in the early project stages.

* Corresponding author.
E-mail addresses: g201474840@kfupm.edu.sa (Y.S. Al-Nahhas), lhadidi@kfupm.edu.sa (L.A. Hadidi), m.islam3@cqu.edu.au (M.S. Islam), mskitmor@bond.edu.au (M. Skitmore), z.abunada@cqu.edu.au (Z. Abunada).
[1] ORCID: 0000–0001-5987–267X
[2] ORCID: 0000–0003-4353–9156
[3] ORCID: 0000–0002-6414–1211
[4] ORCID: 0000–0001-7135–1201
[5] ORCID: 0000-0002-4143-1603

https://doi.org/10.1016/j.asoc.2023.111152
Received 25 May 2023; Received in revised form 9 November 2023; Accepted 10 December 2023
Available online 14 December 2023

1. Introduction

The cost overrun of large-scale infrastructure projects that cost more than USD 35 million to complete [1] is a worldwide issue [2,3]. Infrastructure projects, regardless of geopolitical location and functions, experienced an average of 60% or more final cost than their initially approved budgets [4–6]. Herrera et al. [7], for instance, find that construction projects in such developed countries as the United States, Australia, Holland, and South Korea experience a 16–95% cost overrun, with some European infrastructure projects experiencing a 200% cost overrun [8], while those in developing countries such as Saudi Arabia [9], Qatar [10], and Jordan [11] are from 70–200%. Frequent cost overruns diminish the project's viability by increasing production or service costs, forcing taxpayers to bear higher expenses for infrastructure services. It is essential to conduct a thorough risk analysis and assessment, develop risk management plans, and allocate contingency funds to manage cost overruns effectively.

From the contractor's perspective, avoiding cost overruns leverages a contractor's competency and meets client/owner expectations, creating a win-win scenario. On the other hand, experiencing a cost overrun could jeopardize the company's future [12]. Reducing cost overruns reduces the number of changes in a project, claims, disputes, and abandoned projects, ultimately improving project quality and the country's economy. Cost overrun reduction is a critical and challenging issue that needs an advanced level of integrated risk and cost analysis during the preliminary phases of a project, when little information is available, to address the potential issues that may be encountered in the execution phases [5]. Thus, predicting cost overruns through the analysis of critical risks and their associated cost impacts in similar prior projects is crucial during project cost estimation for having a realistic contingency fund included in the bidding cost to handle unexpected costs during the execution phases [13–15].

Many studies propose models for cost overrun risk assessment; however, only a few are concerned with predicting cost overrun. Among others, Williams and Gong [16] proposed a data mining technique for predicting cost overruns of large construction projects– a technique that depends on quantitative data sets and has a low predictive accuracy. Islam et al. [17] proposed an integrated genetic algorithm and Monte Carlo simulation approach to predict cost overruns using expert judgment. They used a statistical averaging method to measure the severity levels of cost overrun risks, which cannot address uncertainty in subjective data. Leu et al. [18] applied a dynamic Bayesian network and Markov method to predict the cost overrun of an ongoing project based on real-time cost trends and the interdependency of influencing factors. They did not assess critical risk factors and their influences in predicting cost overrun at the early stage of a project for contingency cost planning.

The main issue is that the risk assessment for cost overrun prediction at an early stage of projects mostly depends on a domain expert's judgment, which is qualitative, subjective, uncertain, and vague [6,19,20]. *Fuzzy theory* converts this imprecise, uncertain, and incomplete linguistic data into clear and precise predictions [21,22]. Several Fuzzy Inference Systems (FISs), including the Mamdani, Takagi–Sugeno–Kang (TSK), Tsukamoto, and Singleton models, are utilized in diverse scientific and technological domains to address expert judgment-based problems. Each model differs in terms of how rules are aggregated and defuzzified. For instance, the Takagi–Sugeno–Kang model requires the consequence of rules to be expressed as linear mathematical relationships, while in the Mamdani model, they are represented in such linguistic terms as "high" or "medium.".

Moreover, the Mamdani fuzzy system performs better with more factors and when their relationships to predict outcomes are complex and unclear [23,24]. Due to its capability to handle linguistic assessment terms and many factors (and their unclear relationships), the Mamdani FIS model is typically favored over other models. It is the most commonly used in various science and technological domains for risk and uncertainty modeling [23–28]. The closest to what is needed is Plebankiewicz's [29] *Mamdani FIS model* for predicting cost overruns using complex fuzzy max-min relationships. This finds the cost overrun probabilities of the most sensible activities regarding project cost changes. However, it is limited by requiring very detailed activity-level cost and risk data, not distinguishing between the different knowledge and experience of the experts involved, and not including critical cost overrun factors as input variables nor their importance based on their severity. Moreover, the model needs rebuilding from the beginning each time as it is very much dependent on the unique characteristics of each project.

In response, the present study develops a new fuzzy-Mamdani inference model by first analyzing 13 directly relevant academic articles to identify an in-built list of the 40 most critical factors influencing cost overruns. Then, a questionnaire survey of 31 Saudi Arabian experts is used to rate the severity of each factor based on the respondents' severity scores weighted by each's knowledge and experience with large projects, experience in risk management, and academic qualifications. In applying the list to a specific project, expert judgment is used to estimate the probability of the occurrence of each listed factor, which is then combined with its associated severity value to predict the likely overrun amount for the whole project. This is demonstrated and validated in the Saudi Arabian construction industry, which represents the broader Middle Eastern construction industries, where cost overruns are the norm [9,30,31].

The remainder of this paper is organized as follows. Section 2 reviews the literature relating to the critical causes of cost overruns and potential predictive models. Section 3 describes the approach used to develop the in-built 40-factor severity list and data collection process for part 1 of the model. Section 4 describes part 2 for predicting the overrun of a specific project. Section 5 demonstrates the model's application to predict cost overruns of two real construction projects and a sensitivity analysis to test its robustness. Section 6 provides a discussion of the work and its major contributions. Finally, Section 7 delineates the conclusion and identifies the prospects for further studies.

2. Literature review

This section comprises two parts: identifying potential factors contributing to the cost overruns of large-scale construction projects and an overview of cost overrun prediction models aimed at identifying research gaps for further model development. To achieve this, several keywords, such as 'risk assessment', 'risk and uncertainty modeling, 'cost overrun prediction', 'large projects', 'fuzzy logic', 'fuzzy inference system', and 'expert judgment', were utilized in searches conducted on Google Scholar, Web of Science, the ASCE Library, and the online libraries of the authors' universities to retrieve the most relevant studies. The databases were filtered to include publications from 2001 to 2022, focusing exclusively on peer-reviewed journal articles, indexed conferences limited to IEEE, Scopus, and ASCE, as well as published books. This approach was taken to ensure the quality of the retrieved publications.

2.1. Potential cost overrun factors

In the construction industry, all project parties need to minimize cost overruns. They can cause projects to default, negatively affecting contractors by keeping them trapped with one project for a long time and losing their reputation. Client/owners may need help using the facility, and design and consultation fees could increase. Also, it can result in claims and disputes between the contracted parties.

Different studies from different countries worldwide have been carried out to find and analyze the reasons behind the occurrence of cost overruns to improve their understanding and control. As Table 1 shows, the main influencing factors have been identified in many countries, including Saudi Arabia [32], Bahrain [33], Jordan [34], Palestine [35], UAE [36], Vietnam [37], Egypt [38], and Malaysia [39]. Studies in

Table 1
Cost overrun factors obtained from the literature.

Factor (code)	References													Frequency
	Abusafiya and Suliman [33]	Alhomidan [32]	[34] Al-Hazim and Abusalem [34]	Alghonamy [42]	Creedy et al. [43] [40]	El-Karim et al. [38]	Allahaim and Liu [9]	Forcael et al.[41]	Mahamid and Dmaidi [35]	Johnson and Babu [36]	Kamaruddeen et al.[39]	[31] Seddeeq et al.[31]	Vu et al. [37]	
Poor communication between construction parties (F1)	X	X	-	-	X	-	-	-	X	-	-	X	-	5
Disputes between parties (F2)	X	-	X	X		-	-	-	X	-	-	X	-	5
Contractor’s poor site management and supervision skills (F3)	X	X	-	X	-	-	X	-	X	X	X	X	X	9
Inexperienced project manager for the owner (F4)	X	-	-	X	-	-	-	-	-	X	X	-	X	5
Poor consultant’s management skills (F5)	X	-	-	X	-	-	-	-	X	-	X	X	X	6
Poor productivity (F6)	X	-	-	X		X		X	X	X	-	X	-	7
Lack of knowledge and experience for laborers (F7)	X	-	-		X	X	X		-	-	-	-	-	3
Market conditions (availability and cost of materials, equipment, and labor) (F8)	X	-	X	X	X		X	X	X	-	X	-	X	9
Delays in material delivery (F9)	X	-	-	-	X	X	X	-	-	X	X	-	-	6
Labor insurance, work security, or health problems (F10)	-	-	-	X	-	-	-	-	X	-	-	X	-	3
Bid award for the lowest price (F11)	-	X	-	X	-	-	X	-	-	-	X	X	-	5
Frequent changes in design (F12)	X	X	X	X	X		X	-	X	X	-	X	-	9
Delays in progress payments (F13)	X	X	X	X	X	X	X	-	X	-	-	X	X	10
Delays in decision making (F14)	X	X	X	-	-	-	X	-	-	X	X	X	X	8
Undefined or change in the scope of the project (F15)	X	X	-	X	X		X	-	-	-	X	X		7
Unrealistic contract duration and requirements imposed (F16)	-	X	-	X	X		X	-	X	-	-	X	X	7
Adoption of a fast-track project delivery strategy (F17)	-	-	-	-	-	X	-	-	-	X	-	X		3
Many stakeholders (F18)	-	-	-	-	-	-	-	-	-	-	-	X		1
Poor planning and scheduling (F19)	X	X	X	X	X		X	-	-	-	-	X	X	9

(*continued on next page*)

Table 1 (*continued*)

Factor (code)	References													Frequency
	Abusafiya and Suliman [33]	Alhomidan [32]	[34] Al-Hazim and Abusalem [34]	Alghonamy [42]	Creedy et al. [43] [40]	El-Karim et al. [38]	Allahaim and Liu [9]	Forcael et al.[41]	Mahamid and Dmaidi [35]	Johnson and Babu [36]	Kamaruddeen et al.[39]	[31] Seddeeq et al.[31]	Vu et al. [37]	
Long period between design and time of implementation (F20)	-	-	-	X	-	-	-	-	X	-	-	-	-	2
Inadequate experience and comprehension of the scope of work and site condition (F21)	-	-	-	-	-	-	-	-	-	X	-	X	-	2
Financial status of contractors or sub-contractors (F22)	X	X	-	X	-	-	X		X	X		X	X	8
Delays in subcontractor's work (F23)	X	-	-	-	-	-	-	-	-	-	-	X	-	2
Number of projects the contractor is working on at the same time (F24)	X	-	-	X		X	-	-	X	-	-	-	-	4
Design errors (F25)	X	X	X	X		X	X	-	-	X	-	X	-	8
Delays in supplying and approving drawings (F26)	X	-	-	-	-	-	X	-	-	X	X	-	X	5
Inadequate or changes in material specifications and type (F27)	-	-	-	X	-	-	X	-	-	-	X	X	-	4
Poor quality control/ assurance (F28)	X	X	-	-	-	-		-	-	-	-	X	-	3
Lack of consultant/designer knowledge and experience (F29)	-	-	-	-	X		X	-	-	X	-	-	X	4
Deficiencies in cost predictions (F30)	X	X	X	-	X		X	-	X	X	-	X	-	8
Obstacles from the government (F31)	X	-	X	-	X		X	-	-	-	-	X	X	6
Project size and complexity (F32)	-	-	-	X		X	X	-	-	-	-	-	-	3
Inconvenient site access (F33)	-	X	-			X	X	-	-	-	-	-	-	3
Limited construction area (F34)	-	X	-				X	-	-	-	-	-	-	2
Project location and terrain condition (F35)	X	X	X					-	-	-	-	-	-	3
Social and cultural impacts (F36)	X	X	-	X			X	-	-	-	-	-	-	4
Inflation and taxes (F37)	X	X	X		X	X	X	-	-	-	-	X	-	7
Weather condition (F38)	X	X	X	X	X		X	-	-	-	-	X	-	7
High and inconsistent interest rates charged by bankers on loans (F39)	-	X	-	X	-	X	X	-	-	-	-	-	-	4
Level and number of competitors (F40)	X	-	-	X	-	-	-	-	X	-	-	-	-	3

Australia [40] and Chile [41] also identify the critical factors involved in different large construction projects. Table 1 lists the 40 most common of these and their frequency of mention in the papers.

Most studies in the Middle East region share common factors in that the construction industries in these countries have similar economic and development statuses [9]. For instance, the factors "poor planning and scheduling" and "contractor's poor site management and supervision skills" are found in nine different studies because clients do not invest sufficient funds in hiring qualified personnel for managerial positions [31]. Similarly, the factor "market conditions (availability and cost of materials, equipment, and labor)" is also found in nine different studies. This could be attributed to such reasons as having an unstable economy, unstable political situations, and border closures (e.g., Johnson and Babu, 2020). Another common factor, "frequent changes in design," is found in nine studies and identified as one of the most severe in Saudi Arabia, as most government agencies do not make precise plans for future projects in terms of capacity, services, and location [9]. A maximum of ten studies point to "delays in progress payments," which can be explained by owners not usually investing in the resources needed to evaluate the total cost and time-dependent cash flow sufficiently well before undertaking a project (e.g., Abusafiya and Suliman, 2017). In addition, "financial status of contractor or sub-contractors" is critical, as discovered by eight studies, because contractors in the Middle East do not utilize progress payments effectively to plan for the cash flow situation and manage their financial resources [30]. Also noteworthy is that consultant-related factors of "design error" and "deficiency in cost predictions" are also found in eight studies and, therefore, constitute critical cost overrun factors.

2.2. Overview of cost overrun prediction models

Many studies analyze the risks associated with construction cost overruns, and some develop associated prediction models. These can be categorized into *linear regression* [44], *probability distribution* [45], *simulation* [46], *artificial intelligence* [47], *data mining* [48], and *fuzzy logic*-based models [49]. Their applications and limitations are examined below to identify the research gaps and the need for further development.

[44] developed a *linear regression* model to predict the likelihood of cost overruns, where they used data for ten variables collected from 321 educational projects in Ghana. Unlike large projects, educational projects are homogeneous in their project-related characteristics and contexts except for location, which enabled them to obtain sufficient empirical data for model development and validation. [50] developed a multiple-regression model to quantify the anticipated project cost risk associated with change orders. They collected historical data for 140 projects of different types and sizes in Jordan. The model's independent variables comprised such project attributes as project types (buildings, infrastructure, heavy construction, etc.), size, job type (civil, mechanical, etc.), duration, and original project price. However, they did not conduct a comprehensive risk analysis, and the model demands considerable quantitative cost data from similar projects.

Love et al. (2013) conducted an empirical study to find the best-fit *probabilistic distribution* for predicting realistic cost overruns for construction projects at the time of contract award. They collected numerical cost overrun data from 276 Australian projects, justified the fit of some selected probability distributions using several non-parametric tests, and identified a three-parameter Fréchet probability function as the best-fit distribution. In another study, [48] found a log-logistic probability function to be the best-fit distribution for predicting the cost overruns of highway projects. Their study was also based on intensive numerical data sets, where the cost overrun records of 49 projects were examined. [51] demonstrated a probability model for predicting the cost overrun of a construction project against its contingency values. He also collected numerical cost overrun data by reviewing 34 project documents from the U.S. Army Corps.

The potential of Monte Carlo *Simulation* (MCS) has been demonstrated by [46], for example, in predicting contingency costs to help manage the cost overrun risks encountered in different project phases. The input variables were costs shared in the percentage of different work items of a project. They simulated and aggregated the component costs of a project to find 25 possible total costs with corresponding confidence levels. However, no critical risks or uncertainties were considered for risk scenario development and cost overrun modeling. Similarly, [52] proposed a powerful copula-based MCS model, accommodating different types of distribution patterns of cost variables in a single framework and predicting the outcome based on the best variable distribution pattern. While this overcomes the limitation of assuming cost variables to be randomly distributed, similar to other MCS models, it cannot appropriately address the uncertainties associated with humans [53].

As a form of *artificial intelligence*, the artificial neural net (ANN) model is a powerful machine learning tool that can accommodate many independent variables to predict a single outcome where there are nonlinear input-output relationships. [47] found the best ANN model for predicting cost overruns to be the principal component analysis-based ANN algorithm. They first identified critical cost overrun factors and measured the severity ranking of those factors using subjective judgment, after which 15 critical factors were used to predict cost overruns considering 56 project scenarios. They demonstrated the model in predicting cost overruns of highway projects, finding the basic limitation of ANN models is its uncontrolled architectures of the neurons with the selection of the best number of hidden layers, which significantly influences the prediction accuracy.

[54] have demonstrated a *data mining* model consisting of a bootstrapping-ensemble in ANN for predicting cost overruns at preliminary budgeting for construction projects with project cost data and associated information records from 1600 projects. Another data mining tool called knowledge discovery in database (KDD) was applied by [48] to predict the cost overruns of construction projects. KDD is a powerful tool to retrieve novel knowledge by analyzing historical data from similar previous projects for cost overrun predictions. They collected documents from 90 construction projects, finding baseline costs, actual cost overruns, project specifications (project type, owner type, function, etc.), scope changes, and change orders for cost overrun prediction under various project scenarios. The model's accuracy could be better; hence, its ability to predict the cost overrun of large projects in different company domains is questionable.

The basic limitation of all these models is their dependence on numerical data collected from many real-life similar projects from a particular region/location, while construction industries in different parts of the world have poor data records, and access to the projects' cost data sets is a challenge. Usually, the cost overrun risks and their corresponding cost impacts are assessed by eliciting expert judgment, which is subjective, uncertain, incomplete, and biased [55]. As a result, many studies have adopted fuzzy logic to improve risk assessment, as this can cope with such limitations. For instance, [56] developed a modified Fuzzy Group Decision Making Approach (FGDMA) to identify critical risks in different phases of construction projects. In a further study, [57] added the Bayesian belief network (BBN) theory (i.e., canonical model) with the FDGMA to address the causal relationships among the risks producing cost overruns in different execution phases of a complex project. Later, a risk-induced contingency cost estimation model was proposed, combining fuzzy set theory and BBN theory [5]. As mentioned earlier, Plebankiewicz (2018) proposed a fuzzy-Mamdani inference model for predicting cost overrun using complex fuzzy max-min relationships but limited by the cost overrun risks not being directly given as input variables, their severity importance disregarded, and not being generally applicable as contractors or consultants have to find a good record of each work item's cost as a percentage of a project's total budget.

To summarize, while each model has its advantages and disadvantages, and there is no catch-all model to deal with all risks, the fuzzy logic model is found to be more suitable since it uses the human experience and assumptions to handle the complex problems associated with construction projects [56]. In addition, previous studies mostly assess the potential and critical risks causing cost overruns. While some developed and demonstrated cost overrun prediction models, they mainly depend on quantitative data collected from extensive project records. Thus, *there is a need for a model that can use expert judgment for risk assessment to predict cost overruns and can address uncertainty, bias, and vagueness in expert judgment-based data sets*. Therefore, the present study aims to fulfill this need by developing a new Mamdani-type fuzzy inference model in recognition of its potential to predict cost overruns.

3. Development of the model part 1: the factor severity list

The study aims to develop and test a model on the Mamdani FIS technique [29] to assist decision makers in obtaining the probability of cost overruns for large-scale construction projects. This involves a new 2-part model, in which the first part contains an in-built list of 40 factors influencing cost overruns common for all projects and their degree of severity, while the second part contains each factor's probability of occurrence for a specific project. The 40-factor severity list is built in two steps. The first comprises identifying the factors from 13 highly relevant articles in the literature. The second then establishes the severity of each factor based on a survey of 31 randomly selected experts in the Saudi Arabian construction industry weighted using the Mamdani-fuzzy model by each's degree of knowledge/experience.

3.1. Identifying cost overrun factors

The 13 most relevant studies were used to find the potential and critical factors concerning cost overruns. Table 1 lists the 40 found with corresponding references. These are used to develop a questionnaire to collect the required data for demonstrating the proposed model.

3.2. Questionnaire survey

A structured questionnaire was designed based on the Table 1 factors to measure the severity of each factor (Appendix 1) by an online survey of experts. This comprises two sections. Section 1 solicits the respondents' demographic information, such as their years of experience with large and other construction projects, experience in risk management, and highest academic degree. Section 2 contains a 6-point scale for evaluating the severity of each cost overrun factor ranging from 0 (none) to 5 (very high) impact on project cost overruns.

For the questionnaire survey, the exact population size of experts in Saudi Arabian large-scale projects was unknown as no source confirmed their exact number. Consequently, three primary sources were utilized - the King Fahd University alumni list, the Saudi Contractor Authority [58], and the Saudi Council of Engineers [59] to identify domain experts for the survey. A total of 52 experts from the Saudi Arabian construction industry, including project directors, project managers, cost accountants, project engineers, and designer/consultants, were randomly selected from the domain of large construction projects using those three sources and invited to participate in a structured questionnaire survey. Initially, the experts were contacted via email and sent a consent letter. Follow-up communication was carried out m over the telephone. Once their consent to participate in the questionnaire survey was obtained, they were emailed the questionnaire. Of the 52 experts invited, 31 valid responses were received. It is worth noting that Fuzzy Inference Systems (FIS) offers the advantage of making inferences with a small sample size, even fewer than 30. Many similar fuzzy models have been demonstrated to provide highly accurate outcome predictions when the sample size is close to 30 [24,60–62].

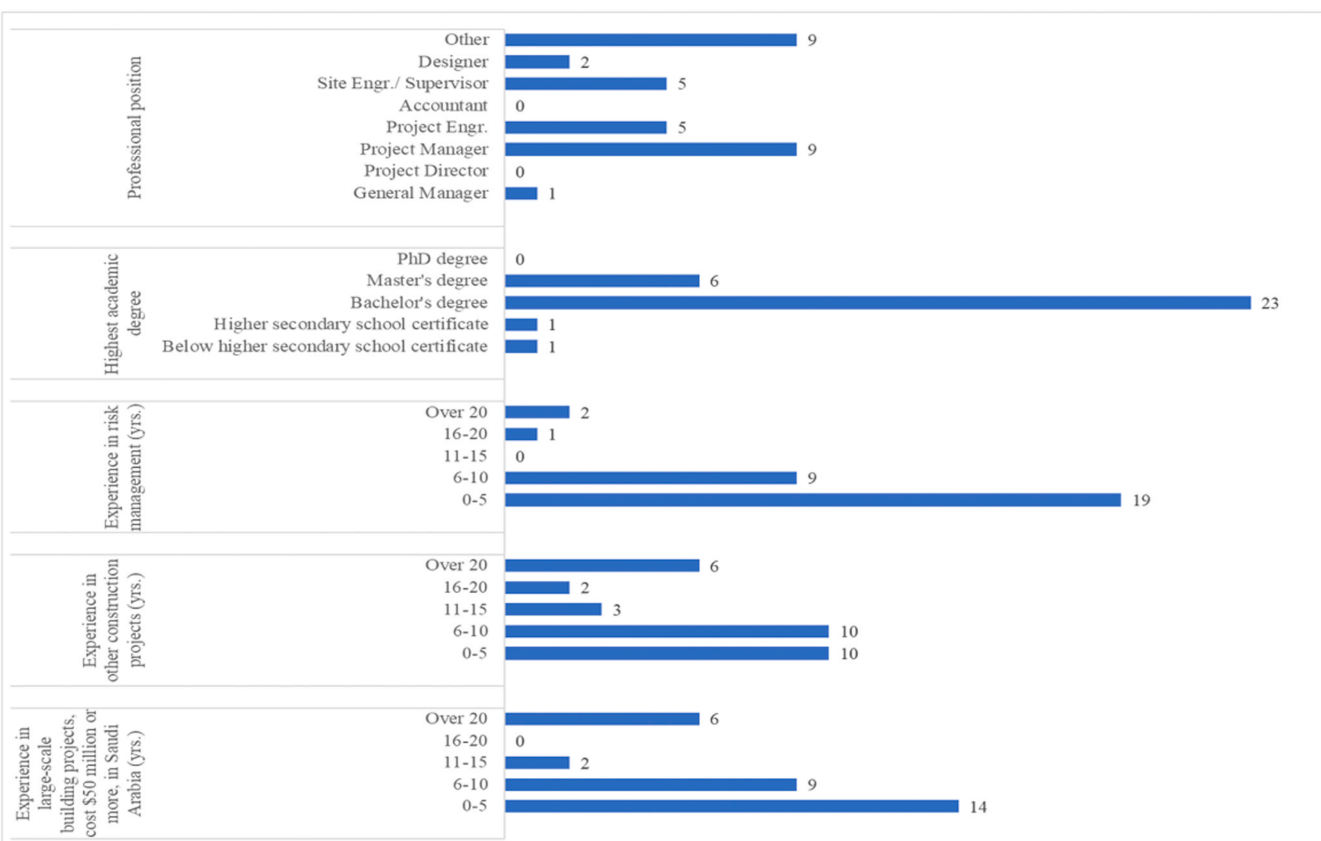

Fig. 1. Summary of respondent profiles.

Fig. 1 shows the respondents' profiles, with their professional positions ranging from general manager, project manager, project engineer, site engineer/supervisor, designer, and other engineering professions. Most hold a BSc Engineering degree with different years of experience in the Saudi construction industry. Having various respondents' knowledge and experience reduces subjective bias by evaluating the risks associated with complex projects [56].

3.3. Data consistency analysis

The consistency analysis reveals significant differences in risk evaluation among domain experts' responses. These variations in expert judgments can be attributed to such factors as knowledge gaps, differing judgment capabilities, biases, unclear comprehension of project uncertainty, and varying project contexts. In this case, the one-sample chi-square test was conducted, which is the most suitable test for the study as it deals with categorical or ordinal data. This data includes information collected from experts using linguistic variables to assess the risks associated with cost overruns (Islam et al., 2019; Ke et al., 2010).

In the Chi-Square test, the null hypothesis (H_0) posits that the experts' judgments are not consistent when evaluating a specific risk, while the alternative hypothesis (H_a) suggests that they are consistent in their assessment of a specific risk. The significance level (α-value) is conventionally set at 0.05 for hypothesis testing. Therefore, if the α-value is less than 0.05, the null hypothesis is rejected, indicating that the experts' risk assessments are consistent. In this study, a chi-square test was conducted to assess the experts' consistency in evaluating the severity of cost overrun risks. Appendix A shows the test results, indicating that the null hypothesis was rejected for 37 factors, with only three exceptions.

Further investigation revealed that, for factor F10, the average evaluation score was 2.03, with two experts rating the factor as 'very high' (5). When these two responses were removed, the p-value becomes 0.0034, less than α (0.05), leading to the rejection of the null hypothesis. Thus, these two responses were considered outliers. Similarly, for F33, the average score was 2.75, with one respondent providing a score of zero for the factor, which, upon removal, provided a p-value of 0.0181, also less than α (0.05) – leading to the rejection of the null hypothesis. Likewise, one respondent provided a score of zero for F36, and the p-value became 0.015 after removal. The revised data were used as inputs

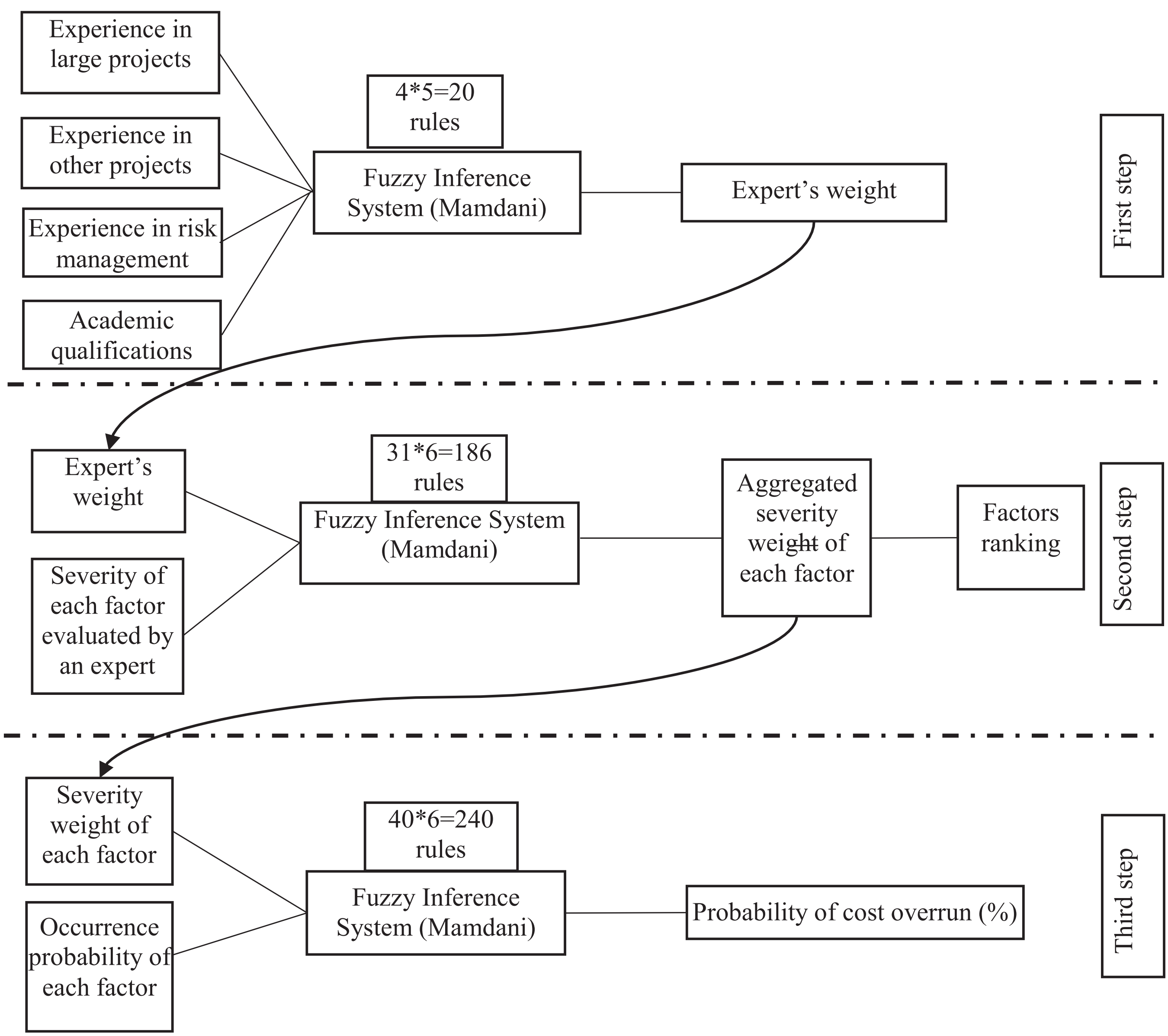

Fig. 2. Mamdani 3-step conceptual FIS model.

in the model to obtain more accurate results for assessing the severity of cost overrun factors.

The following sections illustrate and validate the model.

3.4. 3-step fuzzy model using FIS

A Mamdani Fuzzy Inference System (FIS) is developed for determining the factor-severity list using MATLAB in-built codes. This involves fuzzy logic, which assesses someone's "degree of belief (very high to very low)" instead of "true (1) or false (0)" measured by computers [63]. Fuzzy logic is used in many fields, including project management for time and cost optimization, speech recognition, disease detection (machine learning) processing, and interpreting domain experts' vague linguistic judgments for informed decision making in an uncertain work environment [64]. Following [65], this is given by

$$A = \{(X, \mu_A(X))\,|X \in A, \mu_A(X), \epsilon[0,1]\} \tag{1}$$

where the fuzzy set is denoted by A, discourseuniverse by X, and membership function $\mu_A(X)$, for $0 \le X \le 1$.

There are four basic steps in a Mamdani FIS includes defining the inputs and outputs of the model, selecting the appropriate membership function for the fuzzification of inputs, *if-then* rules, the FIS, and defuzzification (factor ranking by converting fuzzy numbers to a crisp value) [28,66].

Fig. 2 illustrates the conceptual Mamdani 3-step FIS model adopted from [24] to predict cost overruns.

3.4.1. Defining input and output variables

First, all the model's inputs and outputs are specified in the Mamdani FIS in MATLAB. The experts' experience with large and other projects, risk management, and academic qualifications are considered inputs in the first model, which has outputs of the experts' weightings from 0 to 1. The experts' weightings and assessments of each factor's severity are the inputs of the second model, where outputs are aggregated severity of each factor and factor ranking. The aggregated severity level of each factor and its probability of occurrence evaluated by the experts are the inputs of the third model for predicting the probability of cost overrun.

3.4.2. Defining the membership function

The second step in the Mamdani FIS is selecting an appropriate fuzzy membership function (FMF) among different types, such as triangular, rectangular, trapezoidal, and bell-shaped. The selected FMF converts the linguistic evaluation (none to very high) of an input variable to some mathematical values. It is important to select an appropriate fuzzy membership function, as varying outcomes can be desired depending on the project's nature. The triangular and trapezoidal FMFs are most common in terms of accuracy in outcomes. However, they sometimes do not provide reasonable outcomes in the presence of extreme boundaries in data sets. A g-bell-shaped membership function (gbmf) is suitable as uncertainty in expert judgments can be better accommodated [67][24]. Hence, the gbmf is used here since it has the advantage of being smooth and nonzero at all points to address possibly a higher level of uncertainty in data sets based on expert judgments [68].

3.4.3. If-Then rules and the inference system

In the Mamdani FIS model, the relationships among input variables are in *If-Then* statements and control the outcome accuracy. For example, *if an expert's experience is high with large, other construction projects and risk management, and they hold an M.Sc.,* ***then*** *their weight is "high"*. An example for obtaining the severity of a factor severity is "***If*** *the expert evaluates factor 1's severity as high, and that respondent's weight is high,* ***then*** *it has a high severity weight.*" Finally, the defuzzified scores of all factors are computed by their defuzzified severity weights and ranked accordingly. The following sub-sections present the basic fuzzy *If-Then* Mamdani FIS rules used.

Expert weighting:

$$\mathrm{If} E_1 \mathrm{is} a_1, E_2 \mathrm{is} a_2, \ \ldots, \mathrm{and} E_n \mathrm{is} a_k \mathrm{then} EW \mathrm{is} b_i (i = 1, 2, \ldots k = 5) \tag{1}$$

where the expert's attribute is denoted by E, type of attribute by n (= 4) is the, $a_{1,\,2,\,\ldots,\,k}$ denotes the degree of each attribute from 1 (very low) to 5 (very high), and EW is the fuzzy output weight (b_i) of the expert on the same scale. The fuzzy EW is then defuzzified to give a value from zero to unity. Each expert's defuzzified EW is further linguistically interpreted on the same 5-point scale (e.g., Islam et al., 2018).

Severity weight:

If x_1 is c_0, x_2 is c_1, …, and x_n is c_k (k = 0, 1,., k), the corresponding expert's weighting is

$$EW_{1,2,\ \ldots,k} \mathrm{then} SW \mathrm{is} d_i (i = 0, 2,\ \ldots, 5) \tag{2}$$

where the expert is denoted by x (n = 31) and $c_{0,\,1,\ldots k}$ denotes expert x's cost overrun factor's severity level assessed on a 6-point scale (Table 2) from 0 (none) to 5 (very high), EW denotes the weight of the expert on the same 5-point scale as before, and SW is the severity weight as the factor's severity weight (d_i) output on a 6-point fuzzy scale as above. Upon defuzzification, the severity levels of the factors are defined such as 0 to < 0.025 (not significant enough to produce cost overruns), 0.025 to < 0.10 (very little effect on cost overruns), 0.10 to < 0.30 (little significance to cost overruns), 0.30 to < 0.50 (moderate significance to cost overruns), 0.5 to < 0.70 (highly significant to cost overruns), and 0.7 to 1.0 (most significant to cost overruns) (e.g., Islam et al., 2018; [24]).

Inference system:

The model's cost overrun prediction (i.e., the final output) is significantly affected by the aggregation techniques chosen for different fuzzy outputs. There are different aggregation techniques, with the most widely used being the max-min formulation [24,69]. The mathematical expression of the rule-based max-min Mamdani FIS fuzzy output (μ_A) membership function is [24]:

$$\mu_{A_k}(Z) = \max\left[\min\left[\mu_{1_k}(x_1), \mu_{2_k}(x_2), \ldots\ldots \mu_{n_k}(x_n)\right]\right],\ K = 1,\ 2,\ 3,\ r \tag{3}$$

where the input x_1's output membership function under rule k, and $\mu_{A_k}(Z)$ is denoted by μ_{1k}, meaning rule k's output membership function, where inputs are $x_1 to x_n$. K represents all *If-Then* rules.

3.4.4. Defuzzification

The defuzzification is the final step in the model development process. In this step, fuzzy outcomes are transferred into a single value. The centroid of area (COA) method is used for defuzzification, where the weighted average of a fuzzy set is computed, and the final cost overrun is estimated. The COA is the most commonly used defuzzification method, which produces the least error compared to other defuzzification methods [63,70]. The mathematical expression of the COA defuzzification method is shown in Eq. 6:

$$Z_{COA} = \frac{\int \mu_A(Z) \cdot Z dZ}{\int \mu_A(Z) dZ} \tag{4}$$

where Z_{COA} is the final output (i.e., the predicted cost overrun) and $\mu_A(Z)$ is the sum of all the K rules' membership functions computed by the

Table 2
Linguistic variables, defuzzification range, and descriptions.

Level of risk severity	Defuzzification	Description
Very high (5)	0.7 to 1.0	Most significant to cost overruns
High (4)	0.5 to < 0.70	Highly significant to cost overruns
Medium (3)	0.30 to < 0.50	Moderate significance to cost overruns
Low (2)	0.10 to < 0.30	Little significance to cost overruns
Very low (1)	0.025 to < 0.10	Very little effect on cost overruns
None (0)	0 to < 0.025	Not significant enough to produce cost overruns

fuzzy aggregation approach using Eq. 3.

3.5. Building the model

The model development process, including determining the respondents' knowledge/experience weights, factor severity weighting and ranking, and generating the predicted cost overrun (PCO), is discussed in the following.

3.5.1. Weighting experts

The first part of the questionnaire gathers personal profiles of the experts, including their experience in large-scale construction projects (i.e., a total cost of more than $50 million), experience in other construction projects, experience in risk management, and their highest academic degree. Each answer is assigned a 5-point score as before, with experience levels of 0–5 years, 5–10, 10–15, 15–20, and over 20 years, with a significance rating of 1, 2, 3, 4, and 5, respectively. The academic qualification of an expert is accorded a knowledge/experience weight of 1 when below higher secondary and 5 for a Ph.D. degree. Each expert's personal information is used as inputs. The gbmf defines the fuzzy membership functions. Fig. 3 shows each expert's received inputs represented by the gbmf. Twenty rules are formed from four person-related factors, each with five possible answers [24]. All personal questions are assumed to be equally important, and hence, all rules are assigned a weighting of unity. A sample of the rules and assigned weightings are shown in Table 2. Following the Eq. 1 (sub-subsection 3.3.3) concept, the *If-Then* rules are formed accordingly. These establish the expert weighting output (importance) and inputs (expert characteristics) relationships. Fig. 4 provides an example of applying the fuzzy *If-Then* rules to provide expert weightings as fuzzy scale outputs. The output fuzzy membership function is then found by applying Mamdani FIS (Eq. 5), with the COA model (Eq. 6) used for defuzzification to provide the expert weightings. The model's estimated respondent weighting defuzzified output ranges from zero to unity. The inference system is illustrated in Fig. 5, along with the defuzzification process used to obtain the respondents' estimated weightings based on their academic and professional profiles. All the 31 respondents' weightings are obtained by carrying out this process 31 times.

The model's output value (predicted respondent knowledge/experience weight) ranges from zero to unity. The respondents are further clustered based on their weightings into very low, low, medium, high, and very high groups for 0–0.2, 0.2–0.4, 0.4–0.6, 0.6–0.8, and 0.8–1.0, respectively, for input into the second step of the model (obtaining the severity weighting of each factor). The 31 respondents' ranks and weights are shown in Table 3. A total of 3, 15, 12, and 2 respondents have high (over 60%), medium, low, and very low weights, respectively, showing that. These values show the respondents to be high to moderate experienced and academically diversified. It can be noticed from Table 3 that the respondent ranked 31 has a poor profile score of 0.0939, which means that in the next step, their opinion in ranking the cost overrun factors will be less influential than respondent number 1's score of 0.637.

3.5.2. Severity weight and ranking cost overrun factors

The second step is ranking the most common cost overrun factors by severity weight. The severity level of each Table 1 cost overrun factor is assessed by the Saudi construction industry experts. All experts rated each factor's severity based on six potential answers on the 6-point none-very high scale. The answers are used as inputs, with a similar gbmf shown in Fig. 2 defining the membership functions to rank the cost overrun factors. Hence, like Table 2, 186 (=31 ×6) *If-Then* rules with their corresponding weights are used to describe the input variables'

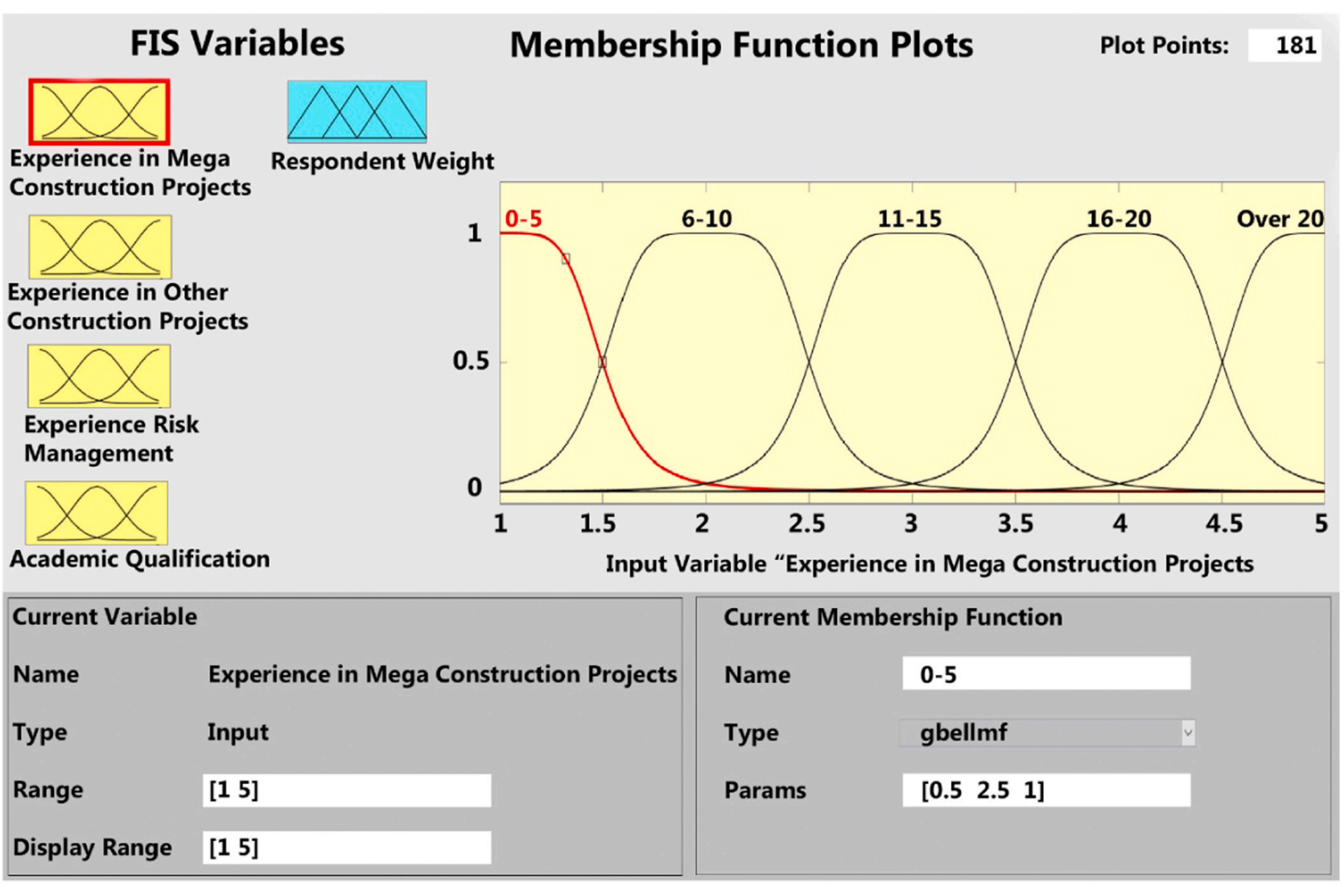

Fig. 3. "Expert Weight" membership function.

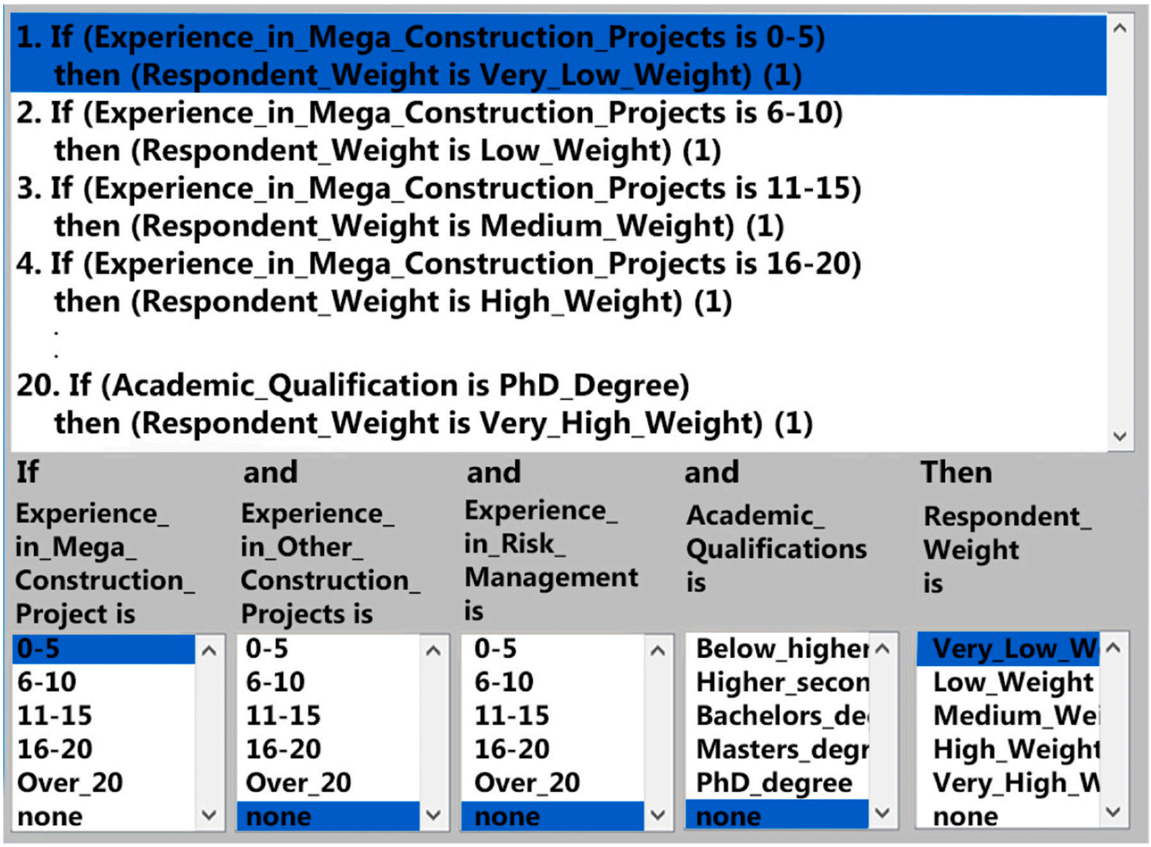

Fig. 4. Example application of *If-Then* rules to obtain the fuzzy linguistic scale "Respondent weighting".

relationships, i.e., six possible answers of each of the 31 respondents and the output. Then, the fuzzy output was defuzzified by the COA model to predict the severity of the required cost overrun factor, similar to the process shown earlier in Fig. 4. Based on the defuzzified severity weight, all the ranked factors are obtained. The final outcome generates the ranking of the 40 cost overrun factors with their severity weights in Table 4. The resulting "factor ranking" is used for weighting rules in the next step. The process of identifying the severity of the first cost overrun factor, "Poor communication between construction parties", is summarized in Fig. 6. It starts with inputting the first factor's severity rates collected from the 31 experts in the developed model. Then, the rules are formed to describe how the experts' (respondents) weights and the evaluation of the factor's severity affect the output, "severity weight." Then, the defuzzification process converts all the cumulated 31 fuzzy inputs into a single value output representing the first factor's overall severity weight. Fig. 7.

The predicted factor severity weight ranges from zero to unity to 1, which can be defined for understanding the severity level of each factor. The factors are classified from the scale in Section 3.3.3, based on their weighting, from 0 to < 0.17, 0.17 to < 0.30, 0.30 to < 0.50, 0.50 to < 0.70, and 0.7 to 1.0 as "very low", "low", "medium", "high", and "very high" severity factors. Table 4 shows the severity ranking and level of the factors identified, indicating 37 factors to be 'high severe' as they have severity weights ranging from 0.50 to 0.70.

4. Development of the model part 2: generating the PCO

The occurrence probability of each cost overrun factor listed in Table 4, assessed by decision makers *for a specific project*, is given as the input in part 2 of the model to predict the cost overrun of the whole project as a percentage of the initial budget. As with the severity level, the occurrence probability is assessed using a 6-point scale ranging from 0 to 5, indicating none to very high respectively. Similarly, following the preceding steps presented in Figs. 2–4, the gbmf defines the input and output membership functions. All potential correlations between the 40 input variables (occurrence probability of each factor) and the output variable (i.e., the PCO) are represented by $40 \times 6 = 240$ rules since each of the 40 factors has six possible outcomes (i.e., occurrence probability). The "Factor's severity weight" obtained in part 1 of the model is used to weigh the 240 rules. Finally, the defuzzification process using the "centroid of area" model is applied to predict the project's likely cost overrun. Fig. 8 summarizes the procedure involved, starting with inputting the factors' occurrence probabilities assessed by the decision-maker/risk management team of a project. Then, the rules are developed to establish the relationship between the severity weight of each factor and its occurrence probability to predict cost overrun. Finally, the defuzzification process transforms the fuzzy outputs into a single value output that represents the PCO of a specific construction project. The basic fuzzy If-then rule for the PCO is as follows:

$$\mathrm{If} F_1 \mathrm{is} O_1, F_2 \mathrm{is} O_2, \ldots \mathrm{and} F_n \mathrm{is} O_k \mathrm{and the respective factor weight is} SW_{1,\ldots,k} \mathrm{successively, then PCO is} C_i (i = 0, 2, \ldots k = 6) \quad (5)$$

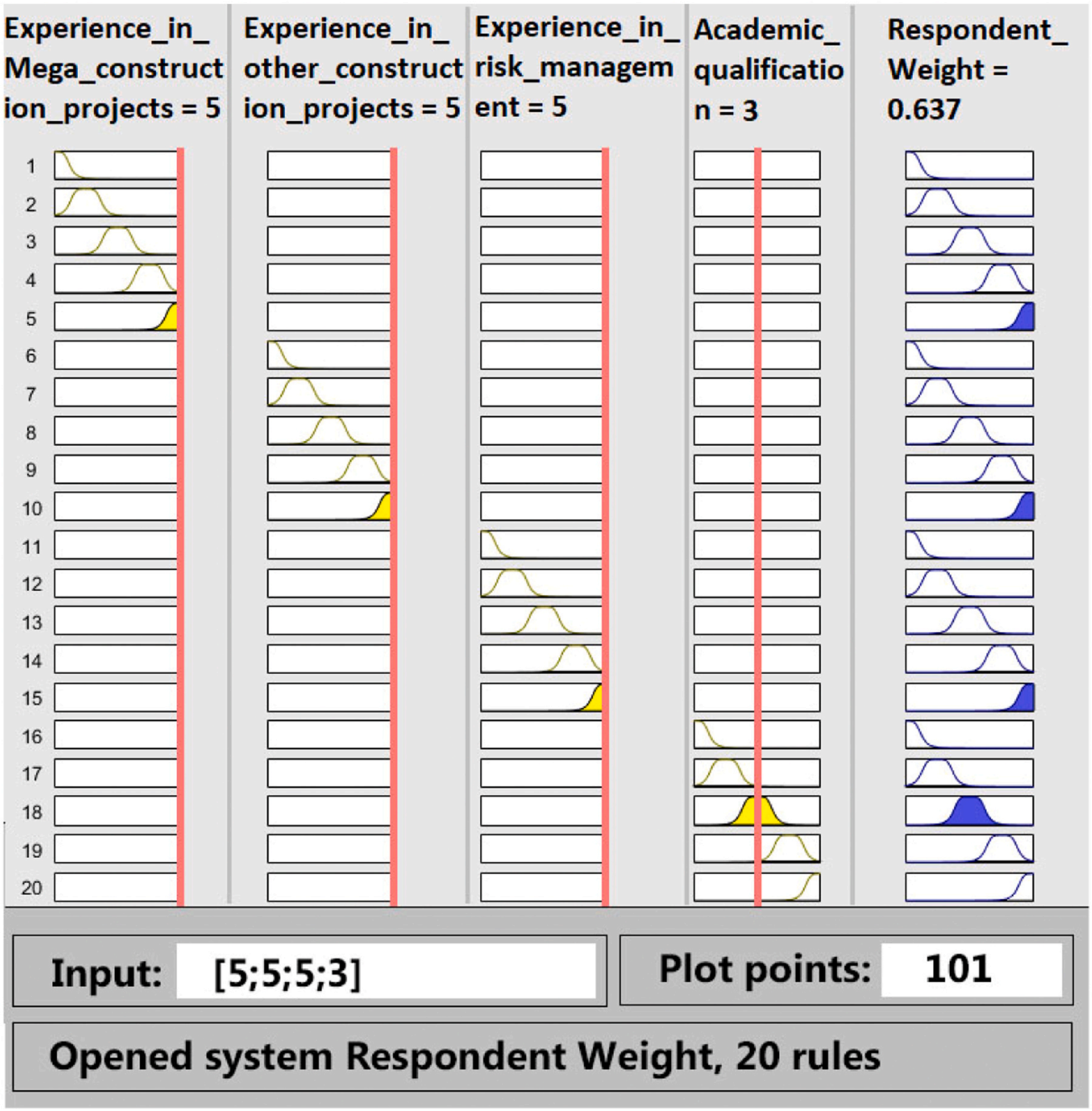

Fig. 5. "Respondent weighting" FIS and defuzzification.

where F is the cost overrun factor, n is the number of identified factors in the model, O is the level of occurrence probability in a 6-point scale ranging from none to very high, SW is the severity weight of a factor found from the first part of the model, the PCO is C in a fuzzy form of none to very high cost overruns.

5. Model demonstration and testing

The model is demonstrated by inputting the probability of cost overrun risks in two real-life projects. Afterward, to check the internal consistency of the model's outcomes, a sensitivity analysis is conducted. The following sections briefly demonstrate the model predicting cost overruns of two projects, followed by a sensitivity analysis.

5.1. Cost overrun prediction of two projects

Once the individual respondent knowledge/experience weight and the severity weight of each factor are provided in the Mamdani-type FIS in MATLAB, then the model creates *in-built* weights of each factor in the system. Now, the sole input to the model is the occurrence probability of each cost overrun for a specific project the experts assess. Suppose the model is fed with the occurrence probabilities of all cost overrun factors. In that case, it will predict the project's cost overrun considering the fuzzy *If-Then* rules developed in Part 2 (Section 4). Accordingly, a study of two actual Saudi Arabian construction projects was carried out to demonstrate the model's capability of predicting cost overruns. The experts were interviewed to rate the occurrence probability of each cost overrun factor considering a recently completed construction project. Also, they were requested to apply traditional practices with their experience to predict the cost overrun in the same project.

The first project was a large-scale slaughterhouse (factory building) in Riyad, the capital city of Saudi Arabia, that has a capacity for hundreds of thousands of cattle. The project's initial estimate was 470 million SAR, and the actual cost of the project was 600 million SAR. This means that the project has experienced a cost overrun of 27.7%. The project manager (contractor's side), senior cost estimator (contractor), and the senior design consultant were requested to evaluate the occurrence probability of selected 40 factors. That was a group interview, and they evaluated the occurrence probability of each factor based on their consensus. The interviewed experts thought that the project would have a 35% to 50% chance of experiencing cost overrun, considering the history of the Saudi construction industry. The information provided for

Table 3
Sample "Expert Weight" fuzzy rules.

Rule	Inputs	Output
	Antecedent	Consequence
1	"Very high "experience in large (mega) and other construction projects and risk management is, with Ph.D.	The respondent has a very high weighting
2	"High" experience with large construction projects.	The respondent has a high weighting
3	"Medium experience with construction projects.	The respondent has a medium weighting
4	"Low" experience with construction projects.	The respondent has a low weighting
5	"Very low" experience with construction projects.	The respondent has a very low weighting
……	……………………………….	…………………………………
……	……………………………….	…………………………………
16	Ph.D. degree	The respondent has a very high weighting
17	Master's degree	The respondent has a high weighting
18	Bachelor's degree	The respondent has a medium weighting
19	Higher secondary school certificate	The respondent has a low weighting
20	Below higher secondary	The respondent has a very low weighting

Table 4
Respondent rank and weight.

Rank	Weight	Rank	Weight
1	0.637	17	0.431
2	0.63	18	0.383
3	0.617	19	0.363
4	0.515	20	0.363
5	0.5	21	0.363
6	0.5	22	0.328
7	0.5	23	0.328
8	0.5	24	0.328
9	0.5	25	0.328
10	0.5	26	0.328
11	0.5	27	0.328
12	0.5	28	0.328
13	0.5	29	0.328
14	0.437	30	0.264
15	0.437	31	0.0939
16	0.431	–	–

the occurrence probability of each factor was input into the model, which predicted a likely project cost overrun of 47.7%. Fig. 9 provides a screenshot of the defuzzification process involved. According to the experts, the model provides a reliable prediction of the cost overrun as its outcome is within the 35%− 50% expected limit set by the experts. Also, it is reasonable for a project that has ended up with a cost overrun of 27.7%. Table 5 presents the experts' answers for the occurrence probability of each cost overrun factor in the studied project.

The second was a smaller-scale project to test the model's reliability with smaller construction projects. The project is called "Site Development in Faculty Housing Area" within a public university campus in Eastern Province, Saudi Arabia. It was a 3500 m^2 (total building area) housing project. The predicted project cost was 14 million SR, while the actual cost at completion was 17 million SR, meaning that the project has a cost overrun of 21.4%. The project manager (contractor side), cost accountant, and risk manager were invited to participate in a group interview to evaluate the probability of 40 cost overrun factors. According to their experience, the experts assessed a 30%− 45% chance that the project would experience cost overrun. Based on their inputs, the model predicts a cost overrun of 44.2%. A screenshot from the model's defuzzification process is presented in Fig. 10. The experts deemed the model's cost overrun prediction acceptable as the project ended up experiencing a cost overrun of more than 21%. The resulting probability falls below the upper predicted limit. Therefore, the model can predict the cost overrun of small-scale projects, too.

5.2. Sensitivity analysis

The model was tested in six different scenarios to evaluate its performance, sensitivity, and robustness. In the first scenario, where no cost overrun factor was expected, i.e., each factor's severity weight was set at zero, the model produced a PCO of 8.1%, which is reasonable as it assumed no difficulties in the project. In the second scenario, where all factors were considered "very low" risk, the model predicted a cost overrun of 21.1%, which is also logical. When the occurrence probability was set to "low," "medium," "high," and "very high" for different scenarios, the model generated PCOs of 40.3%, 59.7%, 78.9%, and 91.9%, respectively. These outcomes indicate that the model's cost overrun predictions are consistent and logical with the cost overrun factors' probability of occurrence.

For further analysis, the model was tested by changing membership functions and defuzzification methods. Table 6 presents the model's outcome against different compositions (i.e., membership function and defuzzification method) of FIS. It shows a minor variation with the change of a membership function. However, changing the defuzzification method has a significant impact on the model's PCOs. In particular, the centroid method and last of maxima (LOM) provide very similar results compared to other methods. The small of maxima (SOM) provides a lower prediction outcome than the centroid (center of area), and the PCO is constant for the mean of maxima (MOM) and Bisector methods, with no variation in changing membership functions. Overall, the centroid defuzzification is the best selection with any membership functions in terms of PCO of the given data set. In fact, the centroid method is the most frequently used defuzzification method in handling expert judgment-based cost or risk modeling [71,72,65,73,74]. Centroid defuzzification has several advantages: (1) the defuzzification process tends to transition smoothly within the output fuzzy region, (2) the output calculation is relatively easy, and (3) the method is applicable to both fuzzy and singleton output set geometries [71,75,76].

6. Discussion and contributions

This study provides an integrated model to the project's principal contract officer (PCO) with an integrated list of severity-weighted factors influencing cost overruns and a project-specific probability of occurrence for each factor assessed by the project team.

The research adds to the body of knowledge by creating a fuzzy model that ranks the severity of cost overrun factors and uses expert judgment to forecast cost overruns in major construction projects. Because experts have differing degrees of knowledge and experience in unpredictable project environments, their opinions are inherently subjective, ambiguous, and uncertain. Expert opinions, therefore, offer a significant challenge in managing subjectivity and uncertainty. To control the subjectivity and uncertainty surrounding expert opinions when applying fuzzy models, three fundamental strategies were used. These are: (1) adding an expert weight in the risk evaluation; (2) using a variety of factors when selecting experts (experience, position, and academic credentials); and (3) fuzzifying (using the membership function in place of a crisp value) the risk evaluation.

With reference to Fig. 1 in sub-Section 3.2, the experts were chosen based on a range of factors, such as their professional role in projects, their academic background, their experience in risk management, and their involvement in large-scale and other projects [5,24,49]. To minimize subjective biases in an equivalent cohort of experts, these expert selection criteria allow for varying degrees of input.

To reduce subjectivity in expert judgment, a second method is to include the expert's weight in the fuzzy inference system. This entails

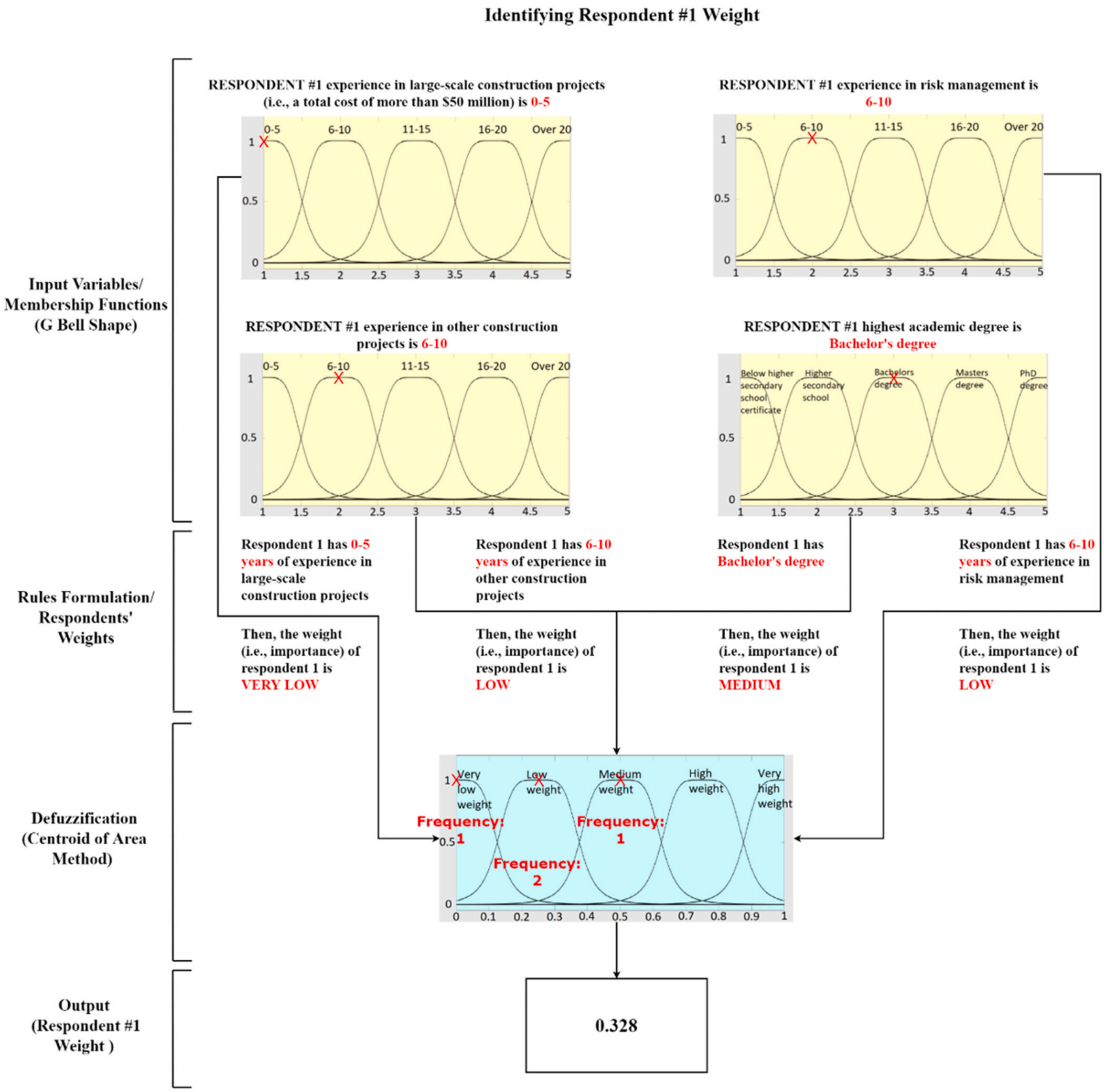

Fig. 6. Summary of the model's computation of the first respondent's weight.

not assigning each expert the same weight because the study's expert selection criteria may impact their assessment. Academic knowledge and experience have a significant impact on an employee's learning curve in engineering judgment [60,61]. Therefore, one limitation of some expert judgment-based studies is that they do not treat experts equally. Following earlier research [47,59,62], the suggested fuzzy model circumvents this restriction by including expert weights in the factor's severity ranking and cost overrun prediction.

When employing fuzzy logic, choosing the membership function is crucial to account for the expert's ambiguous and subjective assessment. Experts assess the degree of risk severity that could result in project cost overruns using their best judgment and frequently disagree on the appropriate level. To account for the subjectivity and biases in expert judgments, the fuzzy membership function fuzzifies the crisp value (0, 1, 2, 3, 4, and 5) or corresponding linguistic (none to very high) evaluation into a three-point (triangular MF), four-point (trapezoidal MF), or continuous value from 0 to 5 (g-bell shape MF) [77-80]. Subsequently, the model employs numerous inference rules (186 in this instance) and an aggregation procedure to summate all the expert responses. A defuzzification technique is then used to generate an output or risk severity level [81,80,82].

Furthermore, the field's decision makers are aware that the result is a contextual scenario of a project rather than a precise value. A high-risk severity level, for example, of 0.549/F31, indicates that the risk's most likely outcome is high, with medium being the most favorable scenario and very high being the worst. The project manager or risk management team can then manage project risk and use their cost contingency budget to control cost overruns by keeping this result in mind.

Referring to Table 4, the ten most severe factors in large-scale Saudi construction projects are 'frequent changes in design', 'delays in

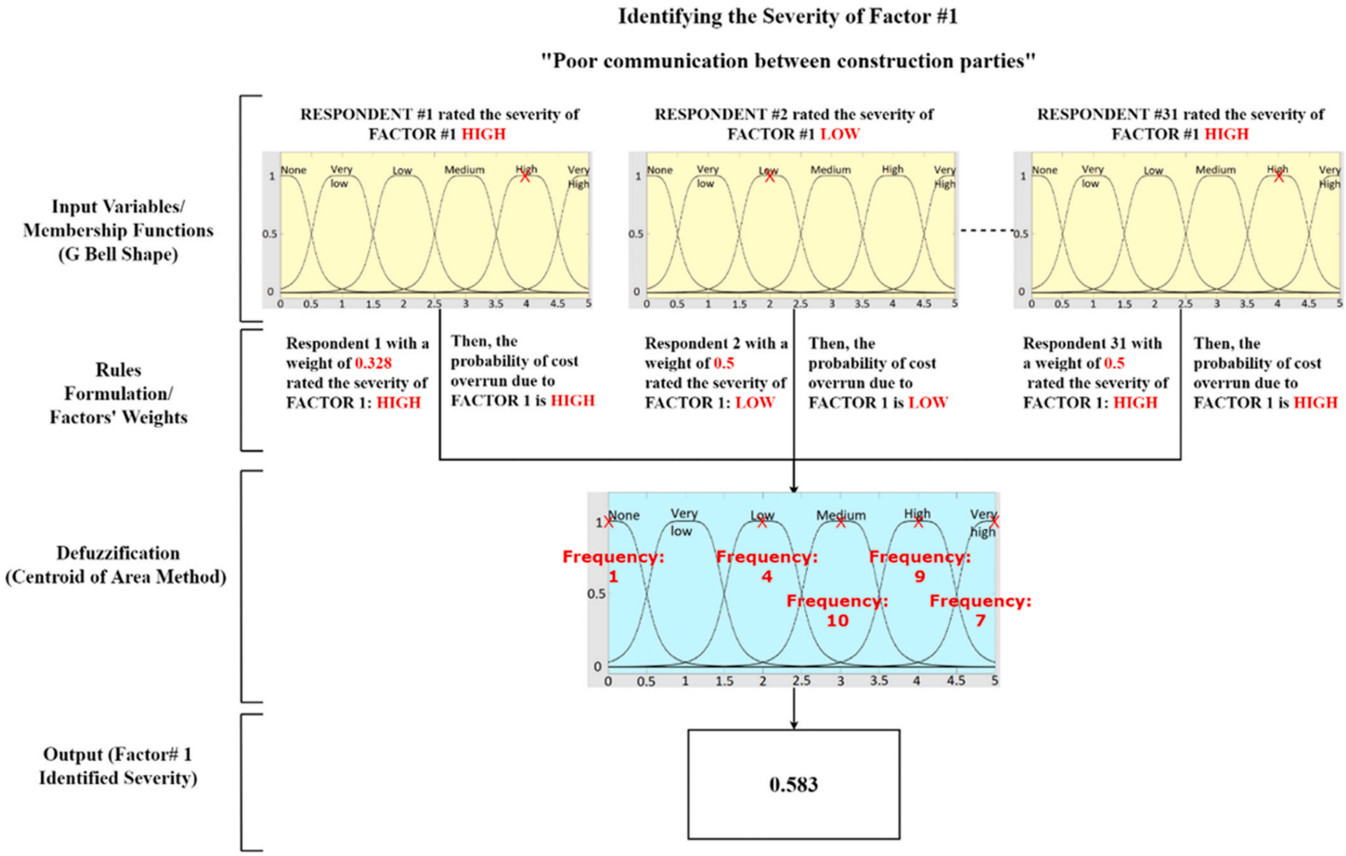

Fig. 7. The model's process for obtaining the factor 1 severity weight.

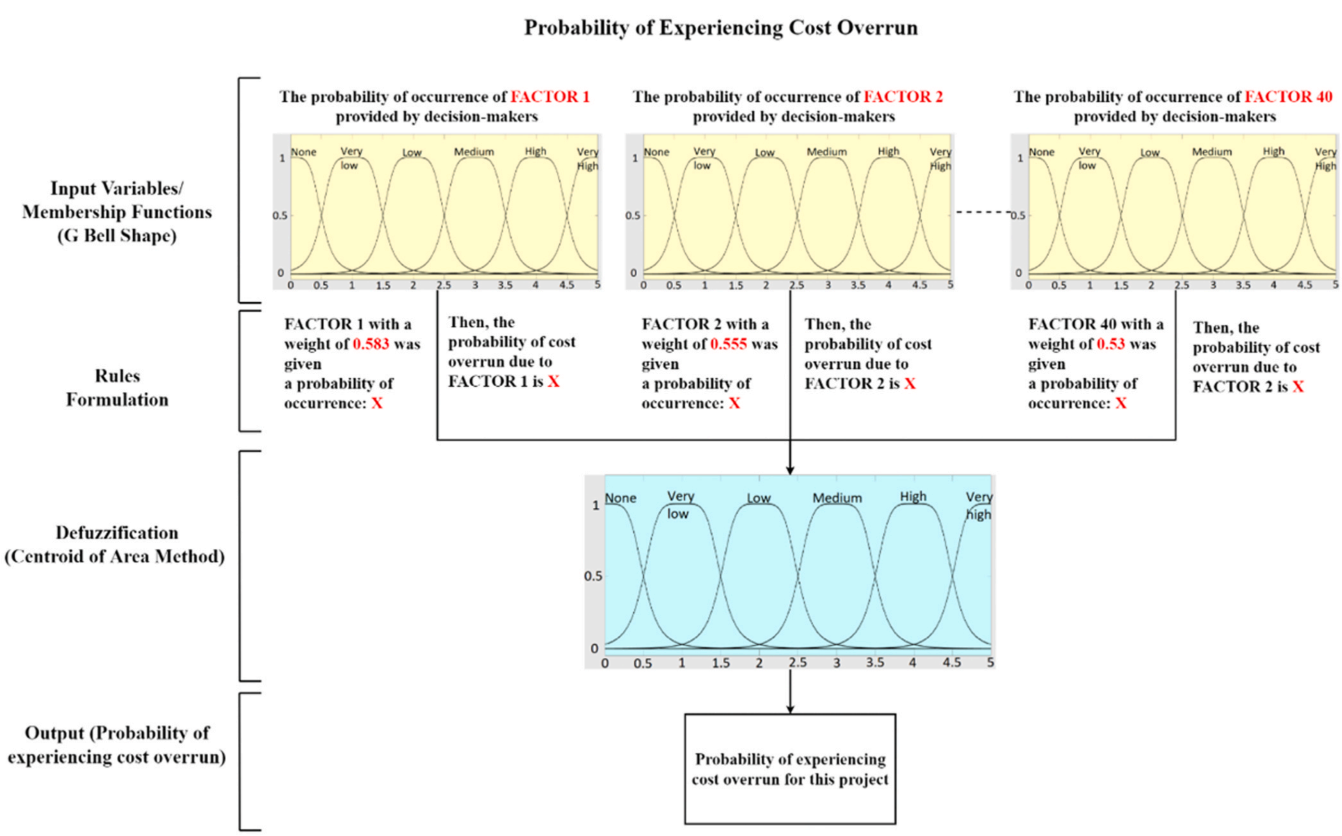

Fig. 8. The flow diagram of the final model to predict the cost overrun.

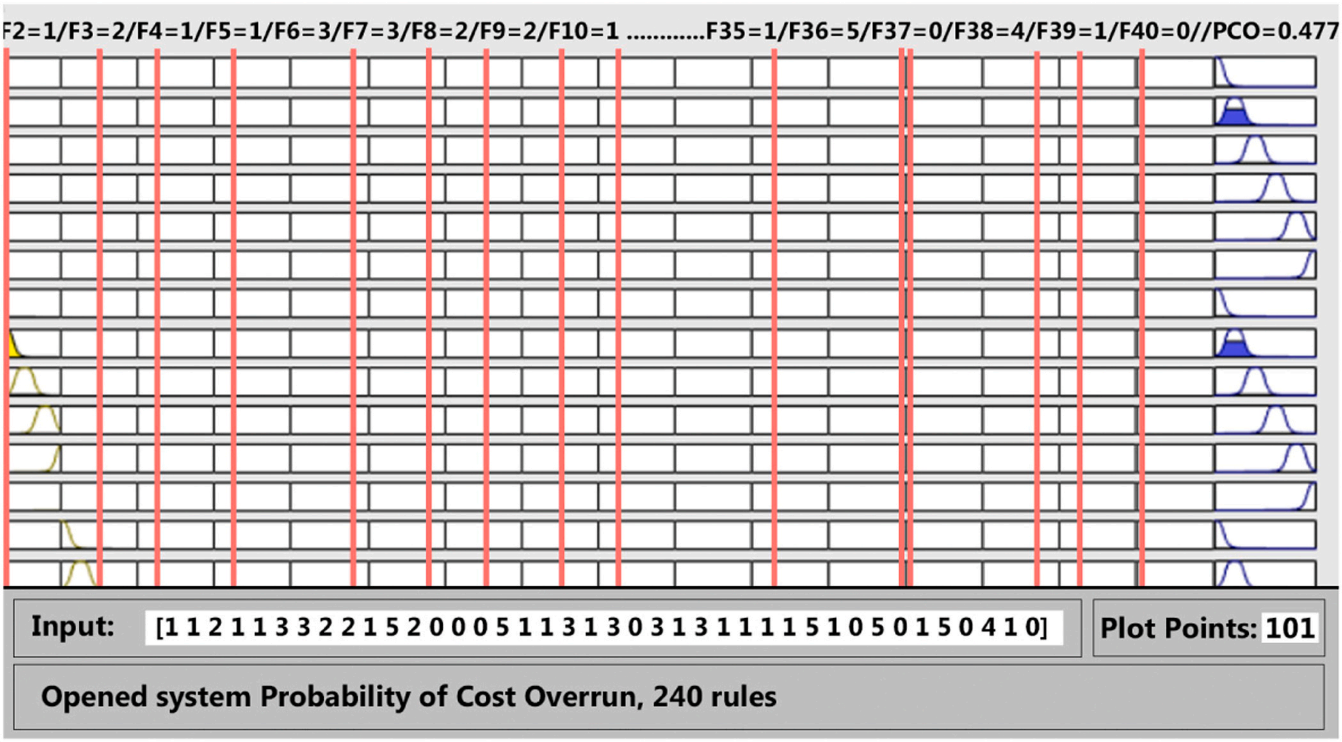

Fig. 9. The model's estimation for predicting the cost overrun of the "slaughterhouse" project.

progress payments', 'undefined or change in project scope', 'inexperienced project manager for owner', 'poor quality control', 'delays in supplying and approving drawings', 'contractor's poor site management', 'delays in decision making', 'financial status of contractor', and 'poor communication between the parties'. Given that these factors have been mentioned in several earlier Saudi Arabian studies, this result is to be expected. The most severe factor, "frequent changes in design," for example, is associated with schedule overruns and raises the project's ultimate cost. It typically arises from the owner's lack of early involvement or design errors [42]. Similarly, protracted approval and payment delays of some government agencies to associated parties account for the second most serious factor, "delays in progress payments" [9]. The third and fifth significant factors are "undefined or change in the project scope" and "poor quality control/assurance," which arise from owners' tendency to underinvest time and funds in choosing the best consultants in Saudi Arabia [31]. In contrast, inexperienced project managers are common due to a lack of workshops and proper training [32], which is a major contributor to inadequate time and cost monitoring and management [31].

To evaluate the suggested model's efficacy and applicability for different uses, it can be contrasted with other fuzzy models in the literature. Few offer cost overrun prediction models; most end with identifying important cost overrun factors. Of those, although the model developed by Islam et al. (2022) that combines a genetic algorithm with Monte Carlo simulation is highly focused on predicting power plant project cost overruns, it neglects to address the subjectivity and uncertainty that come with expert judgment. Similarly, the accuracy of [16] data mining algorithm predicting cost overrun from quantitative cost data sets is poor, and it does not assess risks.

In contrast, Plebankiewicz's (2018) model only involves three factors and 27 inference rules to predict cost overruns using a fuzzy-Mamdani inference system without the need for quantitative data sets from previous projects. As an illustration, our model generates 240 rules for cost overrun prediction, 186 inference rules for severity weighting, and 20 for expert weighting. As a result, our model helps project management teams make informed decisions about early risk planning and management, allocating contingency costs, and controlling cost overruns by providing deeper insights into the relevant and potentially significant factors. Table 7 compares our model and related models from the literature.

In addition, this study provides specific contributions to both the current state of knowledge and professionals in the industry, including:

- Potential and significant contributing factors to large-scale Saudi Arabian construction project cost overruns have been identified. In order to prevent cost overruns, this makes it easier for important project stakeholders—owners, consultants, and contractors—to understand their responsibilities and create risk management plans early on.
- Expert knowledge/experience weights are used to generate the PCO to minimize subjective biases, factor severity weights, and probability of occurrence. When determining appropriate contingency amounts and management reserves for a project, the PCO works with cost estimators and management teams.
- A novel fuzzy-Mamdani model is created for evaluating the risk of cost overruns, and creating the PCO using the MATLAB fuzzy toolbox is a vital resource that Saudi Arabia and other Middle Eastern nations with comparable characteristics can utilize immediately. Furthermore, the developed fuzzy approach can be tailored to other global construction industries by modifying the MATLAB coding to incorporate additional risk factors and different assessments from various expert groups.

The model can also provide a realistic solution to the field professionals in the following ways:

- They can directly anticipate cost overruns and set aside enough money as a contingency for future projects.
- Since the frequency of any risk usually becomes more apparent as the project moves forward, the model can serve as a dynamic tool for modifying the cost contingency. This can provide an early warning of the project's cost performance.

Table 5
Rank and severity weight of cost overrun factors.

Rank	Factor (code)	Severity weight	Severity level
1	Frequent changes in design (F12)	0.652	High
2	Delays in progress payments (F13)	0.647	High
3	Undefined or change in the scope of the project (F15)	0.64	High
4	Inexperienced project manager for owner (F4)	0.635	High
5	Poor quality control/assurance (F28)	0.625	High
6	Delays in supplying and approving drawings (F26)	0.624	High
7	Contractor's poor site management and supervision skills (F3)	0.585	High
8	Delays in decision making (F14)	0.584	High
9	The financial status of contractors or sub-contractors (F22)	0.584	High
10	Poor communication between construction parties (F1)	0.583	High
11	Deficiencies in cost predictions (F30)	0.577	High
12	Lack of knowledge and experience for laborers (F7)	0.576	High
13	Poor planning and scheduling (F19)	0.576	High
14	Lack of knowledge and experience for consultants/designers (F7)	0.575	High
15	Bid award for the lowest price (F11)	0.57	High
16	Delays in subcontractor's work (F23)	0.562	High
17	Poor productivity (F6)	0.561	High
18	Inadequate experience and comprehension of the scope of work and site condition (F21)	0.56	High
19	Design errors (F25)	0.559	High
20	Disputes between parties (F2)	0.555	High
21	Market conditions (availability and cost of materials, equipment, and labor) (F8)	0.552	High
22	Obstacles from government (F31)	0.549	High
23	Unrealistic contract duration and requirements imposed (F16)	0.544	High
24	Long period between design and time of implementation (F20)	0.543	High
25	Poor consultant's management skills (F5)	0.533	High
26	Delays in material delivery (F9)	0.531	High
27	Project size and complexity (F32)	0.531	High
28	High and inconsistent interest rates charged by bankers on loans (F39)	0.531	High
29	Number of projects the contractor works on at the same time (F24)	0.53	High
30	Level and number of competitors (F40)	0.53	High
31	Inflation and taxes (F37)	0.524	High
32	Adoption of a fast-track project delivery strategy (F17)	0.523	High
33	Project location and terrain condition (F35)	0.522	High
34	Social and cultural impacts (F36)	0.516	High
35	Many stakeholders (F18)	0.504	High
36	Inconvenient site access (F33)	0.502	High
37	Inadequate or changes in material specifications and type (F27)	0.501	High
38	Limited construction area (F34)	0.486	Medium
39	Weather condition (F38)	0.478	Medium
40	Laborers' insurance, work security, or health problems (F10)	0.44	Medium

- If the project scenario changes, risks can be added or removed from the model; in this case, MATLAB's model can be modified slightly.
- Assume a group of experts has a unique model. If the experts' characteristics and risk severity assessment are provided, the proposed fuzzy-Mamdani model can be used or adjusted to account for this. The changes will then be incorporated into the final MATLAB cost overrun prediction model. The MATLAB codes attached as an additional file greatly aid in the model's adoption of a particular project-based model.

7. Conclusion

Predicting cost overruns is an important step in taking proactive actions to control cost overruns during the execution phases of a project. However, for large projects, this usually depends on expert judgment-based linguistic data sets, which are uncertain, subjective, and vague in nature. Accordingly, fuzzy logic-based models are the most suitable for addressing such issues in expert judgments. Of the many alternatives available, the Mamdani-FIS is used to develop an integrated risk assessment and cost overrun prediction model. It is the most frequently used method due to its simplicity and capability to aggregate linguistic assessments of many factors with unclear interrelationships to produce realistic outcomes. Considering the number of factors and nature of linguistic data sets for PCO, the Mamdani-FIS theory was employed to construct the model as it can produce the most precise and accurate outcomes.

This study develops and demonstrates a three-step fuzzy-Mamdani model. Step 1 involves assigning weights to each expert, Step 2 computes and ranks the severity weight for each factor, and Step 3 predicts the cost overrun using inputs from Steps 1 and 2, as well as the probability of each factor's occurrence for a specific project. The model comprises 20 rules for expert weighting, 186 rules for severity ranking, and 240 rules for predicting the cost overrun of any given project, ensuring its comprehensiveness compared to previous cost overrun prediction models.

In model building and demonstration, this study initially identified 40 potential and critical cost overrun factors from the literature and developed a structured questionnaire. Fifty-two experts from the Saudi construction industry were invited to participate in the questionnaire survey, and 31 returned completed questionnaires. These experts assessed the severity of each factor on a six-point scale, ranging from 'none' (0) to 'very high' (5). The model calculated the expert weights for Step 1, considering their various characteristics, such as professional position in the project, experience in large and other projects, risk management experience, and academic level. In Step 2, the severity weights of the 40 identified factors were measured and ranked using the inputs of each factor's severity level (ranging from 'none' to 'very high') and the calculated expert weights from Step 1. In the third step, three inputs (i.e., expert weights, severity weights, and occurrence probabilities of each factor in each case study project) were used to predict cost overruns.

The model was demonstrated using two case study projects from the Saudi Construction Industry (SCI). For each case project, a group interview involving the project manager (contractor's side), senior cost estimator (contractor), and senior design consultant was conducted to rate the probability of occurrence of each cost overrun factor using a similar six-point scale to that used for severity evaluation.

A sensitivity analysis was conducted in two steps. Varying frequency levels were input to test the consistency in the model's PCO outcomes, which yielded highly consistent results. Additionally, different combinations of membership functions and defuzzification methods were developed to demonstrate the model's performance in PCO. The results show that selecting a defuzzification method significantly impacts the PCO, while a minor effect was observed when selecting the type of membership function. Ultimately, the centroid method for defuzzification and the generalized bell-shaped membership function proved to be an ideal combination for predicting realistic cost overruns in the case study projects. The severity analysis in Step 2 of the model reveals 37 high-level factors contributing to cost overruns in the SCI. The top-ranked critical factors include frequent changes in design, delays in progress payments, undefined or changing project scope, inexperienced (owner's side) project managers, and poor quality control during the project's execution phase.

The model's performance was demonstrated in two actual construction projects with cost overruns of 27.7% and 21.4%, where project teams' PCO probabilities were 35% to 50% and 30% to 45%, respectively. The model's PCOs for these projects were 47.4% and 42.2%, respectively, closely aligned with their PCOs. The project teams consider it a reliable tool as it allows them to estimate the expected range of

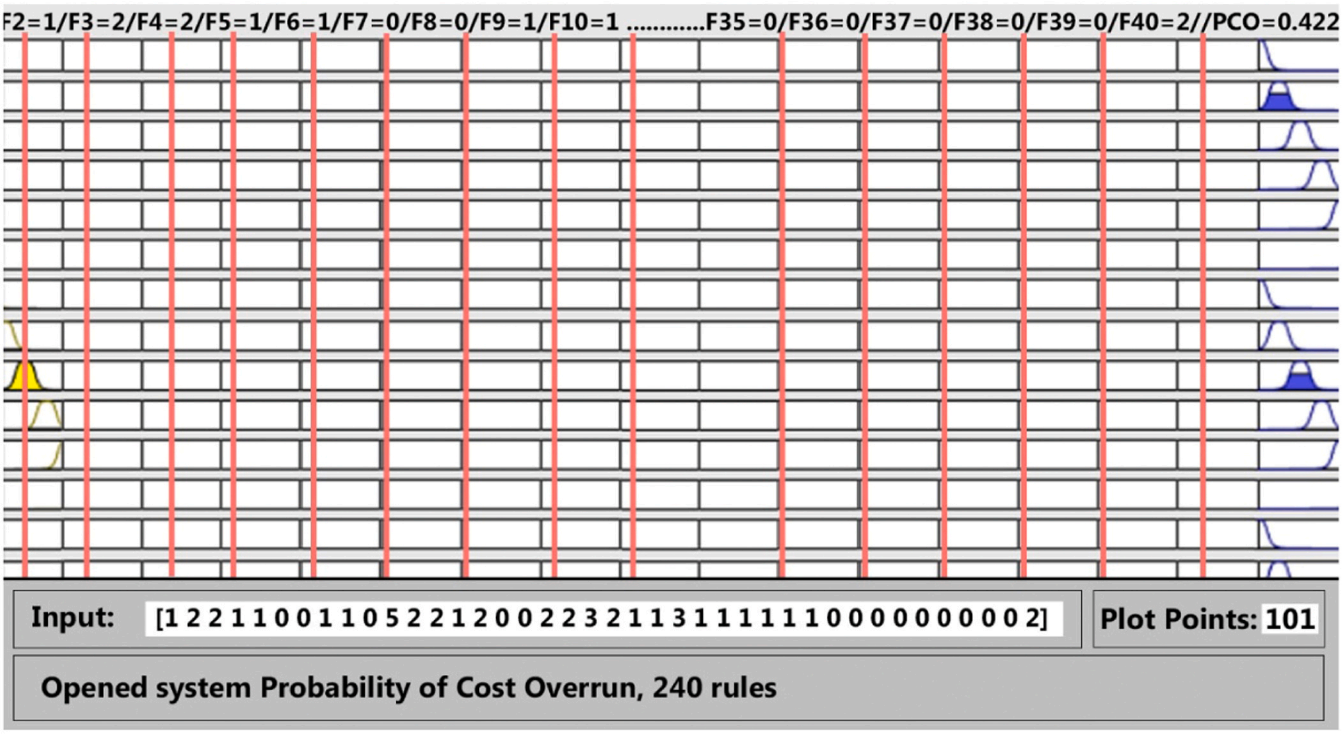

Fig. 10. The model's prediction of the cost overrun in the "Site development in faculty housing area" project.

Table 6
The model's performance for different FIS compositions.

Membership function	Defuzzification method				
	COA	MOM	LOM	SOM	Bisector
Gbmf	0.477	0.4	0.48	0.32	0.46
Gaussian	0.478	0.4	0.47	0.33	0.46
Triangle	0.479	0.4	0.46	0.34	0.46
Trapezoidal	0.477	0.4	0.48	0.32	0.46

profitability in these projects. With the model's outcomes, contractors can make more accurate bid estimates and allocate realistic contingency costs within their bid prices before submitting bids.

The developed Mamdani-FIS model is superior to other models for several reasons. Firstly, it encompasses a maximum of 240 rules within the inference system, considers expert weights, utilizes a single input (specifically, the frequency of each factor) for PCO, achieves high prediction accuracy, and accommodates many factors. Additionally, the model serves two distinct purposes: risk assessment and ranking, as well as cost overrun prediction, which sets it apart from many previous studies. Within the developed model, only a single input, i.e., the occurrence probability of cost overrun factors in a particular project, is necessary to predict its cost overrun. Therefore, it provides an invaluable tool for contractors involved in large-scale projects during the bid estimation phase of project procurement.

Furthermore, this model can function as a dynamic risk management tool, aiding in the monitoring and managing of cost overrun risks and contingency costs during the project execution phases. The risk management team has the flexibility to adjust the severity level of each factor as the project progresses in Step 2 (Severity Ranking Model). This allows them to assess the evolving risk scenarios more clearly during the execution phases and anticipate the impacts on PCO, providing early warnings for effective project cost management.

The study is limited in relying heavily on expert judgment from project personnel to evaluate each cost overrun factor's probability of occurrence, making the model's predictions dependent on their expertise. Additionally, the number of participating experts was limited due to the study being conducted during the COVID-19 restrictions from March 2021 to May 2021, which restricted access to domain experts. Furthermore, the model was developed and demonstrated on only two case projects within the Saudi construction industry. To generalize its applications, it is recommended to validate the model further using more case studies from various construction projects in Saudi Arabia. The model's validation could also include data from diverse infrastructure projects within different construction industry cultures worldwide. Finally, the Mamdani-FIS model is used without comparing the outcomes with other fuzzy models. Therefore, it would be beneficial for future research to explore such alternative fuzzy methods as Sugeno and Tsukamoto to identify the most suitable prediction model for construction cost overruns.

Data sharing

The data used for this study can be obtained from the corresponding author for further academic purposes.

CRediT authorship contribution statement

Yaman Saeid Al-Nahhas: Conception and design of study, Acquisition of data, Analysis and/or interpretation of data, Writing – original draft, Writing – review & editing. **Laith A. Hadidi:** Conception and design of study, Analysis and/or interpretation of data, Writing – original draft, Writing – review & editing. **Muhammad Saiful Islam**: Conception and design of study, Analysis and/or interpretation of data, Writing – original draft, Writing – review & editing. **Martin Skitmore:** Analysis and/or interpretation of data, Writing – review & editing. **Ziyad Abunada:** Writing - a part of introduction and model validation sections.

Declaration of Competing Interest

The authors declare that they have no known competing financial

Table 7
Comparison of the proposed model with other models in the literature.

Comparison areas	Our model	Plebankiewicz[29]	Islam et al.[24]	Williams and Gong[16]	El-Kholy[47]	Leu et al.[18]	Adoko et al.[83]	Knight et al.[78]	Bhargava et al. [84]	Idrus et al. [75]
Method/algorithm	Mamdani-type FIS	Mamdani-type FIS	Genetic Programming and MCS	Data mining model- Ridor, K-Star, Radial Basis neural network, Support Vector decomposition	Random forest regression, support vector regression	Dynamic Bayesian network (BN) and Markov method	Logistic regression	Fuzzy Set Theory	Monte Carlo Simulation	Mamdani-type FIS
No. of inference rules developed	240 rules	Only three rules	Not relevant	Not relevant	Not relevant	No rule but 16 risk-propagation networks	Not relevant	10 scenarios were developed	Not relevant	25 rules
Handling uncertainty in subjective judgment	Yes	Yes	No, risks were assessed using the averaging technique	No subjective risk assessment data is used.	No, Likert scale-based data were used and not fuzzified.	No, quantitative data was used for model validation	No	Partially handle uncertainty, no membership function	No	Yes
Prediction accuracy	Very good (expert's actual prediction 35-50%, model's prediction 44.2%)	80%	Excellent (90% confidence level in cost overrun prediction)	Poor (43.72% prediction accuracy)	Poor (75%)	86%	60% (reliability)	Varies from 85% to 91%	88% (average)	80%
Considered risks as inputs	41 risk factors	Only three factors	Eight factors	Many factors were considered but not specified as cost overrun risks	12 factors	Nine factors	Only five complexity factors	Eight risk factors	Five cost variation factors	Yes- 14 risk factors
General applications of the developed model	Yes, there is an option for updating the risk probabilities or adding new risks	No	No	No- only for power plant projects	No, specific to building and engineering services costs	Adding extra risks and finding a huge probability data is a challenge	Only for complex projects	Yes, project characteristics and risk factors can be updated	No, only for specific projects	Yes, option for updating risk factors
Data required for the model development	31 experts participated in the survey, but a small group of experts' inputs can predict PCO	Very detailed activity-based cost data required	Large data sets are required for training and testing the model (67 projects' data were used)	Large data sets are required. Bootstrapping is needed if small data sets are available.	95 similar building projects' data were used to train and test the model, a data-intensive model	Four experts were interviewed to develop a Bayesian network, a huge amount of probability data was required	66 projects' data were given as inputs, a data-intensive model	Seven experts were interviewed, 18 projects data were given as inputs	15 projects' data were used for model development	Eight experts participated

(*continued on next page*)

Table 7 (*continued*)

Comparison areas	Our model	Plebankiewicz[29]	Islam et al.[24]	Williams and Gong[16]	El-Kholy[47]	Leu et al.[18]	Adoko et al.[83]	Knight et al.[78]	Bhargava et al.[84]	Idrus et al.[75]
Applications in project phases	Throughout the project period-cost estimation and budgeting to execution and monitoring	It can be updated in execution phases based on available cost data	Good for preliminary budgeting	Good for preliminary budgeting	Good for preliminary budgeting	Possible. Need to develop phase-based BN	Good for preliminary budgeting or developing cost baseline	Only applicable for a project proposal development stage	Planning phase, good for preliminary budgeting	Good for preliminary budgeting
User preference	Easy- only frequency of each risk factor is the input	Moderate- activity cost data is input	Moderate- simulation knowledge is required	High tech- Machine learning skills and detailed project reports are required	High tech- Machine learning skills are required, and models are also data intensive	Factors' probabilities and dependence relationships are inputs	Only five complexity levels are the inputs, easy to use	User rates each characteristic of a project and risk factor from 0 to 1 with 0.2 interval, easy to use	A software package was developed, easy to use	Both risk likelihood and risk severity are inputs
Limitation	Two projects were used for the model validation	The model requires lengthy and complex calculations for large-size projects. A limited number (three only) of inference rules limit its application in an uncertain project environment.	The model cannot handle subjective judgment. Risks were scored and ranked using a simple Probability x Impact method. Only eight risks were given as inputs.	Data-driven model. Only eight critical factors were considered for cost overrun prediction, cannot handle subjective bias	Large data sets are required, and selected top factors can be varied with the project context, which can significantly change the model's performance	Six projects were used for model validation	Not very comprehensive to address cost overrun factors, data-driven model	Only design fee cost overruns were predicted, three project characteristics were inputs for validation, could not properly handle uncertainty	Only cost data of a single project in different phases was used for model validation	Three projects were used for the model validation

interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Appendix 1. Structured questionnaire to collect data

General statement: As a construction management professional, you are cordially invited to participate in this questionnaire survey assisting my research on "identifying potential and critical factors affecting cost overruns in large construction projects" as a part of a postgraduation thesis called "Predicting Cost Overruns in Large Construction Projects – A Fuzzy-Mamdani Approach". You are allowed to accept or reject your participation in this survey. If you are agreed, please answer the following:

Part A: Please provide information (√) about yourself based on the following three criteria:

Experience in large construction projects (years)	Experience in other construction projects (years)	Experience in risk management (years)	Academic qualification
0 to 5	0 to 5	0 to 5	Ph.D.
5 to 10	5 to 10	5 to 10	M.Sc.
10 to 15	10 to 15	10 to 15	Bachelor
15 to 20	15 to 20	15 to 20	Higher secondary
over 20	over 20	over 20	Below higher secondary

Part B: Please evaluate (√) the severity of the listed factors affecting the cost overruns in large construction projects. Please evaluate each factor based on the 6-point scale where 0 =none, 1 = very low, 2 = low, 3 = medium, 4 = high, and 5 = very high.

No.	Factor (code)	Severity					
1	Poor communication between construction parties (F1)	0	1	2	3	4	5
2	Disputes between parties (F2)						
3	Contractor's poor site management and supervision skills (F3)						
4	Inexperienced project manager for the owner (F4)						
5	Poor consultant's management skills (F5)						
6	Poor productivity (F6)						
7	Lack of knowledge and experience of laborers (F7)						
8	Market conditions (availability and cost of materials, equipment, and labor) (F8)						
9	Delays in material delivery (F9)						
10	Labors' insurance, work security, or health problems (F10)						
11	Bid award for the lowest price (F11)						
12	Frequent changes in design (F12)						
13	Delays in progress payments (F13)						
14	Delays in decision making (F14)						
15	Undefined or change in the scope of the project (F15)						
16	Unrealistic contract duration and requirements imposed (F16)						
17	Adoption of a fast-track project delivery strategy (F17)						
18	Many stakeholders (F18)						
19	Poor planning and scheduling (F19)						
20	Long period between design and time of implementation (F20)						
21	Inadequate experience and comprehension of the scope of work and site condition (F21)						
22	The financial status of contractors or sub-contractors (F22)						
23	Delays in subcontractor's work (F23)						
24	Number of projects the contractor works on at the same time (F24)						
25	Design errors (F25)						
26	Delays in supplying and approving drawings (F26)						
27	Inadequate or changes in material specifications and type (F27)						
28	Poor quality control/assurance (F28)						

(*continued on next page*)

(*continued*)

No.	Factor (code)	Severity
29	Lack of knowledge and experience for consultants/designers (F29)	
30	Deficiencies in cost predictions (F30)	
31	Obstacles from the government (F31)	
32	Project size and complexity (F32)	
33	Inconvenient site access (F33)	
34	Limited construction area (F34)	
35	Project location and terrain condition (F35)	
36	Social and cultural impacts (F36)	
37	Inflation and taxes (F37)	
38	Weather condition (F38)	
39	High and inconsistent interest rates charged by bankers on loans (F39)	
40	Level and number of competitors (F40)	

Appendix A. Chai-square test result for the severity analysis of 40 factors

Factor	p-value	Comment	Factor	p-value	Comment
F1	0.0013	Rejected	F21	0.0017	Rejected
F2	0.0002	Rejected	F22	0.0004	Rejected
F3	0.0009	Rejected	F23	0	Rejected
F4	0.0011	Rejected	F24	0.0074	Rejected
F5	0.0011	Rejected	F25	0.0024	Rejected
F6	0.0133	Rejected	F26	0	Rejected
F7	0.001	Rejected	F27	0.2085	Rejected
F8	0.0001	Rejected	F28	0	Rejected
F9	0.002	Rejected	F29	0.0001	Rejected
F10	**0.1166**	**Retained**	F30	0.0001	Rejected
F11	0.0004	Rejected	F31	0.045	Rejected
F12	0	Rejected	F32	0	Rejected
F13	0	Rejected	F33	**0.072**	**Retained**
F14	0	Rejected	F34	0.0442	Rejected
F15	0	Rejected	F35	0.028	Rejected
F16	0	Rejected	F36	**0.0515**	**Retained**
F17	0.0014	Rejected	F37	0.02	Rejected
F18	0.0157	Rejected	F38	0.0366	Rejected
F19	0.0003	Rejected	F39	0.0292	Rejected
F20	0.0001	Rejected	F40	0.01	Rejected

References

[1] R. Foti, Managing large construction projects, PM, Netw., PMI 15 (2001) 24–31.
[2] M. Siemiatycki, The making and impacts of a classic text in megaproject management: the case of cost overrun research, Int. J. Proj. Manag. (2016), https://doi.org/10.1016/j.ijproman.2016.07.003.
[3] H. Lind, F. Brunes, Explaining cost overruns in infrastructure projects: a new framework with applications to Sweden, Constr. Manag. Econ. 33 (2015), https://doi.org/10.1080/01446193.2015.1064983.
[4] A. Gilbert, B.K. Sovacool, P. Johnstone, A. Stirling, Cost overruns and financial risk in the construction of nuclear power reactors: a critical appraisal, Energy Policy 102 (2017) 644–649, https://doi.org/10.1016/j.enpol.2016.04.001.
[5] M.S. Islam, M.P. Nepal, M. Skitmore, R. Drogemuller, Risk induced contingency cost modeling for power plant projects, Autom. Constr. 123 (2021), 103519, https://doi.org/10.1016/j.autcon.2020.103519.
[6] B.K. Sovacool, A. Gilbert, D. Nugent, Risk, innovation, electricity infrastructure and construction cost overruns: Testing six hypotheses, Energy 74 (2014) 906–917, https://doi.org/10.1016/j.energy.2014.07.070.
[7] R.F. Herrera, O. Sánchez, K. Castañeda, H. Porras, Cost overrun causative factors in road infrastructure projects: A frequency and importance analysis, Appl. Sci. 10 (2020), https://doi.org/10.3390/app10165506.
[8] IHS, IHS Costs and Strategic Sourcing. Power Capital Costs Index and European Power Capital Costs, (2014). https://doi.org/10.1016/0263–7863(95)00025-L.
[9] F.S. Allahaim, L. Liu, Causes of cost overruns on infrastructure projects in Saudi Arabia, Int. J. Collab. Enterp. 5 (2015) 32, https://doi.org/10.1504/ijcent.2015.073176.
[10] A. Senouci, A. Ismail, N. Eldin, Time Delay and Cost Overrun in Qatari Public Construction Projects, : Procedia Eng. (2016), https://doi.org/10.1016/j.proeng.2016.11.632.
[11] E. Hayat, Risk attitude and risk controllability: their implications on the subjective quantification of risk in international construction projects, Int. J. Adv. Eng. Res. Sci. 4 (2017) 33–45, https://doi.org/10.22161/ijaers.4.10.6.
[12] J.A. Hackett, C.M. Hicks, Estimating as a profession in UK construction, Eng. Constr. Archit. Manag. 14 (2007) 26–51, https://doi.org/10.1108/09699980710716963.
[13] M.S. Islam, S.R. Mohandes, A. Mahdiyar, A. Fallahpour, A.O. Olanipekun, A coupled genetic programming Monte Carlo simulation–based model for cost

overrun prediction of thermal power plant projects, J. Constr. Eng. Manag. 148 (2022) 1–14, https://doi.org/10.1061/(asce)co.1943-7862.0002327.
[14] A.E. Yildiz, I. Dikmen, M.T. Birgonul, K. Ercoskun, S. Alten, A knowledge-based risk mapping tool for cost estimation of international construction projects, Autom. Constr. 43 (2014) 144–155, https://doi.org/10.1016/j.autcon.2014.03.010.
[15] S. Shoar, N. Chileshe, J.D. Edwards, Machine learning-aided engineering services' cost overruns prediction in high-rise residential building projects: Application of random forest regression, J. Build. Eng. 50 (2022), 104102, https://doi.org/10.1016/j.jobe.2022.104102.
[16] T.P. Williams, J. Gong, Predicting construction cost overruns using text mining, numerical data and ensemble classifiers, Autom. Constr. 43 (2014) 23–29, https://doi.org/10.1016/j.autcon.2014.02.014.
[17] M.S. Islam, S.R. Mohandes, A. Mahdiyar, A. Fallahpour, A.O. Olanipekun, A. Coupled Genetic Programming-Monte Carlo Simulation-based Model for Cost Overruns Prediction of Power Plant Projects, 2022. ⟨http://www.tjyybjb.ac.cn/CN/article/downloadArticleFile.do?attachType=PDF&id=9987⟩.
[18] S. Leu, C. Lu, P. Wu, Dynamic-Bayesian-Network-Based Project Cost Overrun Prediction Model, Sustainability 15 (2023) 1–14.
[19] A. Makui, S.M. Mojtahedi, S.M. Mousavi, Project risk identification and analysis based on group decision making methodology in a fuzzy environment, Int. J. Manag. Sci. Eng. Manag. 5 (2010) 108–118, https://doi.org/10.1080/17509653.2010.10671098.
[20] H.M. Gharaibeh, Cost control in mega projects using the delphi method, J. Manag. Eng. 30 (2013) 1–7, https://doi.org/10.1061/(ASCE)ME.1943-5479.0000218.
[21] I. Dikmen, M.T. Birgonul, S. Han, Using fuzzy risk assessment to rate cost overrun risk in international construction projects, Int. J. Proj. Manag. 25 (2007) 494–505, https://doi.org/10.1016/j.ijproman.2006.12.002.
[22] C. Samantra, S. Datta, S.S. Mahapatra, Fuzzy based risk assessment module for metropolitan construction project: an empirical study, Eng. Appl. Artif. Intell. 65 (2017) 449–464, https://doi.org/10.1016/j.engappai.2017.04.019.
[23] M. Monjezi, M. Rezaei, Developing a new fuzzy model to predict burden from rock geomechanical properties, Expert Syst. Appl. 38 (2011) 9266–9273, https://doi.org/10.1016/j.eswa.2011.01.029.
[24] I.S. Zaqout, M.S. Islam, L.A. Hadidi, M. Skitmore, Modeling bidding decisions and bid markup size for construction projects: a fuzzy approach, Eng. Appl. Artif. Intell. 113 (2022), 104982, https://doi.org/10.1016/j.engappai.2022.104982.
[25] F. Camastra, A. Ciaramella, V. Giovannelli, M. Lener, V. Rastelli, A. Staiano, G. Staiano, A. Starace, A fuzzy decision system for genetically modified plant environmental risk assessment using Mamdani inference, Expert Syst. Appl. 42 (2015) 1710–1716, https://doi.org/10.1016/j.eswa.2014.09.041.
[26] R. Mohammad, A. Mostafa, M. Abbas, H. Mohammad Farouq, Prediction of representative deformation modulus of longwall panel roof rock strata using Mamdani fuzzy system, Int. J. Min. Sci. Technol. 25 (2015) 23–30, https://doi.org/10.1016/j.ijmst.2014.11.007.
[27] A. Daftaribesheli, M. Ataei, F. Sereshki, Assessment of rock slope stability using the fuzzy slope mass rating (FSMR) system, Appl. Soft Comput. J. 11 (2011) 4465–4473, https://doi.org/10.1016/j.asoc.2011.08.032.
[28] E. Plebankiewicz, D. Wieczorek, Adaptation of a cost overrun risk prediction model to the type of construction facility, Symmetry (Basel) 12 (2020) 1–17, https://doi.org/10.3390/sym12101739.
[29] E. Plebankiewicz, Model of predicting cost overrun in construction projects, Sustain 10 (2018), https://doi.org/10.3390/su10124387.
[30] S.A. Assaf, S. Al-Hejji, Causes of delay in large construction projects, Int. J. Proj. Manag. 24 (2006) 349–357, https://doi.org/10.1016/j.ijproman.2005.11.010.
[31] A.Bin Seddeeq, S. Assaf, A. Abdallah, M.A. Hassanain, Time and cost overrun in the Saudi Arabian oil and gas construction industry, Buildings 9 (2019) 1–17, https://doi.org/10.3390/buildings9020041.
[32] A. Alhomidan, Factors Affecting Cost Overrun in Construction Projects, Int. J. Eng. Adv. Technol. 9 (2019) 4831–4833, https://doi.org/10.35940/ijeat.b4415.129219.
[33] H.A.M. Abusafiya, S.M.A. Suliman, Causes and Effects of Cost Overrun on Construction Project in Bahrain: Part I (Ranking of Cost Overrun Factors and Risk Mapping), Mod. Appl. Sci. 11 (2017) 20, https://doi.org/10.5539/mas.v11n7p20.
[34] N. Al-Hazim, Z. Abusalem, Delay and cost overrun in road construction projects in Jordan, Int. J. Eng. Technol. 4 (2015) 288, https://doi.org/10.14419/ijet.v4i2.4409.
[35] I. Mahamid, N. Dmaidi, Risks leading to cost overrun in building construction from consultants' perspective, Organ. Technol. Manag. Constr. Int. J. 5 (2013) 860–873, https://doi.org/10.5592/otmcj.2013.2.5.
[36] R.M. Johnson, R.I.I. Babu, Time and cost overruns in the UAE construction industry: a critical analysis, Int. J. Constr. Manag. 20 (2020) 402–411, https://doi.org/10.1080/15623599.2018.1484864.
[37] T.Q. Vu, C.P. Pham, T.A. Nguyen, P.T. Nguyen, P.T. Phan, Q.Le. Hoang Thuy To Nguyen, Factors influencing cost overruns in construction projects of international contractors in Vietnam, J. Asian Financ. Econ. Bus. 7 (2020) 389–400, https://doi.org/10.13106/JAFEB.2020.VOL7.NO9.389.
[38] M.S.B.A.A. El-Karim, O.A. Mosa El Nawawy, A.M. Abdel-Alim, Identification and assessment of risk factors affecting construction projects, HBRC J. 13 (2017) 202–216, https://doi.org/10.1016/j.hbrcj.2015.05.001.
[39] A.M. Kamaruddeen, C.F. Sung, W. Wahi, A study on factors causing cost overrun of construction projects in Sarawak, Malaysia, Civ. Eng. Archit. 8 (2020) 191–199, https://doi.org/10.13189/cea.2020.080301.
[40] G.D. Creedy, M. Skitmore, J.K.W. Wong, Evaluation of risk factors leading to cost overrun in delivery of highway construction projects, J. Constr. Eng. Manag. 136 (2010) 528–537, https://doi.org/10.1061/(asce)co.1943-7862.0000160.
[41] E. Forcael, H. Morales, D. Agdas, C. Rodríguez, C. León, Risk Identification in the Chilean Tunneling Industry, EMJ - Eng. Manag. J. 30 (2018) 203–215, https://doi.org/10.1080/10429247.2018.1484266.
[42] A. Alghonamy, Cost overrun in construction projects in Saudi Arabia: Contractors' perspective, Int. J. Mech. Mechatron. Eng. 15 (2015) 35–42.
[43] G.D. Creedy, M. Skitmore, J.K.W. Wong, Evaluation of Risk Factors Leading to Cost Overrun in Delivery of Highway Construction Projects, J. Constr. Eng. Manag. 136 (2010) 528–537.
[44] R.O. Asiedu, N.K. Frempong, H.W. Alfen, Predicting likelihood of cost overrun in educational projects, Eng. Constr. Archit. Manag. 24 (2017) 21–39, https://doi.org/10.1108/ECAM-06-2015-0103.
[45] P.E.D. Love, C.-P. Sing, B. Carey, J.T. Kim, Estimating construction contingency: accommodating the potential for cost overruns in road construction projects, J. Infrastruct. Syst. 21 (2015), 04014035, https://doi.org/10.1061/(asce)is.1943-555x.0000221.
[46] B.A. Traynor, M. Mahmoodian, Time and cost contingency management using Monte Carlo simulation, Aust. J. Civ. Eng. 17 (2019) 11–18, https://doi.org/10.1080/14488353.2019.1606499.
[47] A.M. El-Kholy, Exploring the best ANN model based on four paradigms to predict delay and cost overrun percentages of highway projects, Int. J. Constr. Manag. 21 (2021) 694–712, https://doi.org/10.1080/15623599.2019.1580001.
[48] M.M. Ghazal, A. Hammad, Application of knowledge discovery in database (KDD) techniques in cost overrun of construction projects, Int. J. Constr. Manag. 0 (2020) 1–15, https://doi.org/10.1080/15623599.2020.1738205.
[49] M.S. Islam, M.M. Islam, H.R.T. Imran, M. Skitmore, M.P. Nepal, Grasping the Nonconformities in Building Construction Supply Chains, J. Perform. Constr. Facil. 35 (2021), 04020141, https://doi.org/10.1061/(asce)cf.1943-5509.0001546.
[50] A.A. Abu Hammad, S.M. Alhaj Ali, G.J. Sweis, A. Bashir, Prediction model for construction cost and duration in Jordan, Jordan J. Civ, Eng 2 (2008) 250–266.
[51] A. Touran, Probabilistic model for cost contingency, J. Constr. Eng. Manag. 129 (2003) 280–284, https://doi.org/10.1061/(asce)0733-9364(2003)129:3(280).
[52] A. Firouzi, W. Yang, C.-Q. Li, Prediction of total cost of construction project with dependent cost items, J. Constr. Eng. Manag. 142 (2016), 04016072, https://doi.org/10.1061/(asce)co.1943-7862.0001194.
[53] P. Loizou, N. French, Risk and uncertainty in development: a critical evaluation of using the Monte Carlo simulation method as a decision tool in real estate development projects, J. Prop. Invest. Financ. 30 (2012) 198–210, https://doi.org/10.1016/j.autcon.2010.02.008.
[54] D.D. Ahiaga-dagbui, S.D. Smith, Rethinking construction cost overruns, Cogn., Learn. Estim., J. Financ. Manag. Prop. 19 (2014) 38–54, https://doi.org/10.1108/JFMPC-06-2013-0027.
[55] L.M.P. Beltrão, M.T.M. Carvalho, Prioritizing construction risks using fuzzy AHP in Brazilian public enterprises, J. Constr. Eng. Manag. 145 (2019) 05018018, https://doi.org/10.1061/(asce)co.1943-7862.0001606.
[56] M.S. Islam, M.P. Nepal, M. Skitmore, Modified Fuzzy Group Decision-Making Approach to Cost Overrun Risk Assessment of Power Plant Projects, 04018126-1–15, J. Constr. Eng. Manag. 145 (2019), https://doi.org/10.1061/(ASCE)CO.1943-7862.0001593.
[57] M.S. Islam, M.P. Nepal, M. Skitmore, G. Kabir, A Knowledge-based Expert System to Assess Power Plant Project Cost Overrun Risks, Expert Syst. Appl. 136 (2019) 12–32.
[58] SCA, Saudi Contractor Authority, (n.d.). ⟨https://muqawil.org/public/publications/b0c7e107eff513f6ed7c6aa0b1d82d1adc9c4fd0.pdf⟩.
[59] SCE, Saudi Council of Engineers, (n.d.). ⟨www.saudieng.sa⟩.
[60] Y.-C. Kuo, S.-T. Lu, Using fuzzy multiple criteria decision making approach to enhance risk assessment for metropolitan construction projects, Int. J. Proj. Manag. 31 (2013) 602–614, https://doi.org/10.1016/j.ijproman.2012.10.003.
[61] M. Abdelgawad, A.R. Fayek, Risk Management in the Construction Industry Using Combined Fuzzy FMEA and Fuzzy AHP, J. Constr. Eng. Manag. 136 (2010) 1028–1036, https://doi.org/10.1061/(ASCE)CO.1943-7862.0000210.
[62] A. Jamshidi, A. Yazdani-Chamzini, S.H. Yakhchali, S. Khaleghi, Developing a new fuzzy inference system for pipeline risk assessment, J. Loss Prev. Process Ind. 26 (2013) 197–208, https://doi.org/10.1016/j.jlp.2012.10.010.
[63] M.M.G. Elbarkouky, A.R. Fayek, N.B. Siraj, N. Sadeghi, Fuzzy Arithmetic Risk Analysis Approach to Determine Construction Project Contingency, J. Constr. Eng. Manag. 142 (2016), 04016070, https://doi.org/10.1061/(ASCE)CO.1943-7862.0001191.
[64] A. Hajian, P. Styles, Applications of Fuzzy Logic in Geophysics 8 (2018) 301–371, https://doi.org/10.1007/978-3-319-66532-0_4.
[65] A. Nieto-Morote, F. Ruz-Vila, A fuzzy approach to construction project risk assessment, Int. J. Proj. Manag. 29 (2011) 220–231, https://doi.org/10.1016/j.ijproman.2010.02.002.
[66] E. Plebankiewicz, D. Wieczorek, Prediction of cost overrun risk in construction projects, Sustain 12 (2020) 1–15, https://doi.org/10.3390/su12229341.
[67] M. Fasanghari, G.A. Montazer, Design and implementation of fuzzy expert system for Tehran Stock Exchange portfolio recommendation, Expert Syst. Appl. 37 (2010) 6138–6147, https://doi.org/10.1016/j.eswa.2010.02.114.
[68] H. Mittal, P. Bhatia, Optimization Criteria for Effort Estimation using Fuzzy Technique, CLEI Electron. J. 10 (2007), https://doi.org/10.19153/cleiej.10.1.2.
[69] S.A. Gamel, A.I. Saleh, H.A. Ali, A fog-based Traffic Light Management Strategy (TLMS) based on fuzzy inference engine, Neural Comput. Appl. 34 (2022) 2187–2205, https://doi.org/10.1007/s00521-021-06525-2.
[70] N. Gerami Seresht, A.R. Fayek, Computational method for fuzzy arithmetic operations on triangular fuzzy numbers by extension principle, Int. J. Approx. Reason. 106 (2019) 172–193, https://doi.org/10.1016/j.ijar.2019.01.005.

[71] X. Zhao, B. Hwang, A.M. Asce, S.P. Low, Developing Fuzzy Enterprise Risk Management Maturity Model for Construction Firms, J. Constr. Eng. Manag. (9) (2013) 1179–1189, https://doi.org/10.1061/(ASCE)CO.1943-7862.0000712.
[72] N. Sadeghi, a R. Fayek, W. Pedrycz, Fuzzy Monte Carlo simulation and risk assessment in construction, Comput. Civ. Infrastruct. Eng. 25 (2010) 238–252, https://doi.org/10.1111/j.1467-8667.2009.00632.x.
[73] A. Nieto-Morote, F. Ruz-Vila, A fuzzy approach to construction project risk assessment, Int. J. Proj. Manag. 29 (2011) 220–231, https://doi.org/10.1016/j.ijproman.2010.02.002.
[74] M. Khanzadi, F. Nasirzadeh, M. Alipour, Integrating system dynamics and fuzzy logic modeling to determine concession period in BOT projects, Autom. Constr. 22 (2012) 368–376, https://doi.org/10.1016/j.autcon.2011.09.015.
[75] A. Idrus, M.F. Nuruddin, M.A. Rohman, Development of project cost contingency estimation model using risk analysis and fuzzy expert system, Expert Syst. Appl. 38 (2011) 1501–1508, https://doi.org/10.1016/j.eswa.2010.07.061.
[76] E. Cox, M. O'Hagan. The Fuzzy Systems Handbook, Second edi.,, Morgan Kaufmann,, 1998.
[77] F. Afzal, S. Yunfei, D. Junaid, M.S. Hanif, Cost-risk contingency framework for managing cost overrun in metropolitan projects: using fuzzy-AHP and simulation, Int. J. Manag. Proj. Bus. 13 (2020) 1121–1139, https://doi.org/10.1108/IJMPB-07-2019-0175.
[78] K. Knight, A.R. Fayek, A.M. Asce, Use of Fuzzy Logic for Predicting Design Cost Overruns on Building Projects, J. Constr. Eng. Manag. 128 (2002) 503–512.
[79] M. Abdelgawad, A.R. Fayek, Comprehensive Hybrid Framework for Risk Analysis in the Construction Industry Using Combined Failure Mode and Effect Analysis, Fault Trees, Event Trees, and Fuzzy Logic, J. Constr. Eng. Manag. 138 (2012) 642–651, https://doi.org/10.1061/(ASCE)CO.1943-7862.0000471.
[80] A.R. Fayek, Fuzzy Logic and Fuzzy Hybrid Techniques for Construction Engineering and Management, J. Constr. Eng. Manag. 146 (2020), 04020064, https://doi.org/10.1061/(asce)co.1943-7862.0001854.
[81] J. Zeng, M. An, N.J. Smith, Application of a fuzzy based decision making methodology to construction project risk assessment, Int. J. Proj. Manag. 25 (2007) 589–600, https://doi.org/10.1016/j.ijproman.2007.02.006.
[82] V. Carr, J.H. Tah, A fuzzy approach to construction project risk assessment and analysis: construction project risk management system, Adv. Eng. Softw. 32 (2001) 847–857, https://doi.org/10.1016/S0965-9978(01)00036-9.
[83] M.T. Adoko, T.A. Mazzuchi, S. Sarkani, Developing a cost overrun predictive model for complex systems development projects, Proj. Manag. J. 46 (2016) 111–125, https://doi.org/10.1002/pmj.21545.
[84] A. Bhargava, S. Labi, S. Chen, T.U. Saeed, K.C. Sinha, Predicting cost escalation pathways and deviation severities of infrastructure projects using risk-based econometric models and Monte Carlo simulation, Comput. Civ. Infrastruct. Eng. 32 (2017) 620–640, https://doi.org/10.1111/mice.12279.