A coupled genetic programming Monte Carlo simulation–based model for cost overrun prediction of thermal power plant projects

Muhammad Saiful Islam ; Saeed Reza Mohandes ; Amir Mahdiyar ; Alireza Fallahpour ; Ayokunle Olubunmi Olanipekun

Bibliographic Citation

Islam, M. S., Mohandes, S. R., Mahdiyar, A., Fallahpour, A., & Olanipekun, A. O. (2022). A Coupled Genetic Programming Monte Carlo Simulation–Based Model for Cost Overrun Prediction of Thermal Power Plant Projects. Journal of Construction Engineering and Management, 148(8), 04022073. https://doi.org/10.1061/(ASCE)CO.1943-7862.0002327

A Coupled Genetic Programming-Monte Carlo Simulation-based Model for Cost Overruns Prediction of Thermal Power Plant Projects

Muhammad Saiful Islam[1], Saeed Reza Mohandes[2], Amir Mahdiyar[3*], Alireza Fallahpour[4], Ayokunle Olubunmi Olanipekun[5]

[1]Project Management Discipline, School of Engineering and Technology, Central Queensland University- Melbourne Campus, VIC 3000, Australia. Email: m.islam3@cqu.edu.au; ORCID: 0000-0002-6414-1211

[2]Department of Building and Real Estate (BRE), Faculty of Construction and Environment (FCE), The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong. Email: saeedreza.mohandes@polyu.edu.hk

[3]School of Housing, Building and Planning, Universiti Sains Malaysia, 11800-Penang, Malaysia. Email: amirmahdiyar@usm.my. Tel: +601112132721

[4]School of Mechanical Engineering, Universiti Teknologi Malaysia, Skudai 81310, Malaysia; Email: fallahpour.a@utm.my

[5]School of Architecture and Built Environment, University of Wolverhampton, Wolverhampton, UK. Email: a.olanipekun@wlv.ac.uk

[*] Corresponding author: amirmahdiyar@usm.my

Abstract

Globally, power projects are prone to cost overrun projects. Within the body of knowledge, previous studies have paid less attention to predicting the cost overruns to assist contingency cost planning. Particularly, the thermal power plant projects (TPPPs), the enormous risks involved in their delivery undermine the accuracy of cost overrun prediction. To prevent cost overrun in thermal plant projects, these risks need to be accounted for by employing sophisticated cost overrun prediction techniques. This study aims to develop a Hybrid Predictive-Probabilistic-based Model (HPPM) that integrates a genetic programming technique with a Monte Carlo simulation (MCS). The HPPM was proposed based on the data collected from TPPPs in Bangladesh. Also, the sensitivity of the HPPM was examined to identify the critical risks in cost overruns simulation. The simulation outcomes show that 40.48% of a project's initial estimated budget was the most probable to cost overrun, while the maximum cost overrun will not exceed 75% with 90% of confidence. Practically, the analysis will sensitize project managers to emphasize thermal plants' budget accuracy not only at the initial project delivery phase but throughout the project lifecycle. Theoretically, the HPPM model could be employed for cost overrun prediction in other types of power plant projects.

Keywords: Cost overruns; machine learning; infrastructure; thermal power plant; MCS

1 Introduction

The complex nature of power plant projects (PPPs) leads to unavoidable and significant cost overruns worldwide (Sovacool et al. 2014a). For instance, some PPPs in Europe experienced almost 200% unexpected costs above the budgeted amount of a project (IHS 2014). The cost overrun can effectively be managed by proper risk assessment, management planning, and allocation of contingency costs in the project development stage. Therefore, predicting cost overrun by analyzing critical risks and corresponding cost overruns of similar previous projects is an important step for allocating contingency sum to a project in its initial stage (Islam et al. 2018a, 2019; De Marco et al. 2016). There are different types of power plants, and they are unique in their respective compositions. They include thermal power plants (coal, heavy fuel oil (HFO), combined cycle power plant (CCPP), natural gas, etc.), nuclear, wind, or hydropower plants (Hadikusumo and Tobgay 2015; Kucukali 2016; Yoo et al. 2016). Each type is influenced by changes in project characteristics, ownership, delivery method, fund arrangement, geographic location, etc. (Hashemi et al. 2019). From the literature, the critical risks and contingency costs for predicting hydropower, wind, and nuclear power plants are well documented. Contrastingly, there has been less attention paid to critical risks and contingency costs for predicting TPPPs. An exception is Islam et al. (2021). The study developed a contingency cost prediction model for TPPP. The study did not analyze TPPP risks relationships and their impacts on cost overrun. Therefore, risks relationships that influence TPPP cost overrun are not known and remain a research gap.

Risk assessment contributes to the prediction of cost overrun, and it depends on expert judgment for complex infrastructure projects like power plants. Thus, the models, which can handle semi-quantitative or linguistic data sets, subjectivity, or biases of the data, and are not data-intensive, are appropriate for risk assessment and cost overrun prediction of TPPPs. The fuzzy logic can accommodate subjective risk assessment and cost inputs in the contingency cost prediction model (Islam et al. 2021) with the limitations of appropriately aggregating subjective and linguistic data to assess risks and corresponding cost impacts. Alternatively, MCS is a powerful tool for predicting project cost overrun and it considers various risk scenarios and demonstrated by both academics (Maronati and Petrovic 2019; Shahtaheri et al. 2017) and professional bodies (Government 2017; Nevada DOT 2012; US-DOT 2015). Maronti and Petrovic (2019) developed a distribution-free rank correlation between cost items applying MCS for predicting costs of nuclear PPPs. The study revealed the project characteristics that influenced the cost prediction, but also, differentiates nuclear power projects from TPPPs. The integrated probabilistic risk analysis and MCS approach were proposed by (Shahtaheri et al. 2017) to estimate the risk probability associated with individual activities and assess cost impacts of risks based on the performances of nuclear plant tube replacement projects. Notably, none of them (Maronati and Petrovic (2019) and Shahtaheri et al. (2017)) predicted cost overruns taking critical TPPPs risks and their possible relationships as the inputs for cost overrun prediction using MCS.

On the other hand, Genetic Programming (GP) is a powerful tool that uses the best size tree (that includes functions and arguments having different sizes and shapes in the form of a hierarchy), by applying a nonlinear encoding system to find the best-fit model for different inputs-output relationships (Lin et al. 2020). It can capture the natural phenomena of a complex problem by developing a comprehensive equation through functional evolution during training and testing stages in the prediction model (Z-Flores et al. 2017). Thus, GP overcomes the human bias of subjective data sets in finding real relationships among the input variables to predict the output. Accordingly, the power of GP was demonstrated to handle expert-judgment-based inputs predicting project schedules (Lin et al. 2020). Shahrana et al. (2017) predicted project cost by developing a GP-based model using numerical data sets collected from 210 sewer and water rehabilitation projects. However, no previous study attempted to use GP to predict project cost overruns where expert judgment-based risk assessment would be the sole inputs. Thus, the coupled GP-MCS is a robust integrated approach to overcome the identified research gaps (i.e., lack of a comprehensive model for predicting cost overrun of TPPPs that can capture the complex relationships among the critical risk factors, as well as paucity of a probabilistic-based model for simulating cost overrun of such projects) for proactive risk management and cost control of complex PPPs, TPPPs in particular. Therefore, this study aims to develop a novel Hybrid Predictive-Probabilistic-based Model (HPPM) using the integration of Genetic Programming (GP) and MCS technique for predicting cost overruns of complex TPPs, where subjective judgment-based risks magnitude are the inputs. The unique contributions of this study are as follows:

- Obtaining an inclusive equation for the first time in the literature, based on which the cost overruns of TPPPs can be calculated.

- The model's outcome, i.e., percentage of cost overrun with the probability of occurrence, can directly be applied in practice to make an informed decision about allocating contingency cost for future power plant project delivery.
- The critical factors that are identified and their respective levels of sensitivity levels for predicting cost overruns is a guideline for managing risks and associated uncertainties towards achieving target costs in PPPs.
- The predicting approach of cost overrun can be applied to other complex infrastructure projects, where risk assessment and cost overrun data are largely subjective and correlated with unknown relationships.

The remainder of this study is organized as follows. First, the cost overrun scenario of PPPs is presented to identify the research gap in TPPPs, followed by a critical discussion on existing cost overrun prediction models. Second, the methodology, including step-by-step model development and data collection approach, is discussed. Third, the HPPM is demonstrated for predicting cost overruns in real Bangladeshi TPPPs. A sensitivity analysis was conducted to identify the impacts of critical risk factors (RFs) on cost overrun predictions. Finally, the concluding remarks are presented with practical applications and the limitations leading to further studies.

2 Literature Review

2.1 TPPPs in cost overrun literature

Table S1 shows a summary of the studies on the risk assessment of PPPs. It was observed that the previous studies mainly focused on assessing the risks in hydropower, wind power, and nuclear PPPs, while TPPPs were neglected. Meanwhile, it is notable that thermal power plants are the major sources of energy globally, especially the ones based on natural gas, oil, and coal (Hans-Wilhelm 2016). Also, increasing demand for energy supply from thermal plants for the next twenty-five years has been projected to guarantee energy access (Singer and Peterson 2017). Despite the commonness of thermal power plants, which have enhanced energy access globally, the successful provision of this infrastructure was affected by cost overrun (Sovacool et al. 2014a). Therefore, there is a need to understand the RFs that lead to cost overrun in TPPPs. This will enable better control of thermal power plants' budgeting and, by extension, lead to more thermal power projects procured at optimum budgets to increase energy access globally. The previous studies on the risks and cost overruns in different PPPs such as hydro, nuclear, and wind power plants (Alsharif and Karatas 2016; Filho et al. 2017; Hashemi et al. 2019; Hossen et al. 2015; Kim et al. 2017; Maronati and Petrovic 2019; Sudirman and Hardjomuljadi 2011a; Wang and Tiong 2000) showed that they differ significantly in their economic, technical, stakeholder policies, environmental, and spatial characteristics (Gilbert et al. 2017; Sovacool et al. 2014a; b). Also, the consequences and likelihood of cost overruns vary significantly according to project size, type, and contracting system (Sovacool et al. 2014b; a). Although Islam et al. (2021) proposed a contingency cost prediction model for TPPPs, the study was limited to addressing risk impacts on allocating contingency cost during preliminary budgeting instead of directly predicting cost overruns using RFs as input variables. Thus, the real risk scenarios and exposure of TPPPs to cost overruns are not reflected in the existing literature. Also, the existing cost or contingency prediction models are not yet justified to predict cost overruns of TPPPs. To conceptualize the risk terminology, it refers to the uncertainties that may cause either negative or positive project impacts. However, in this study, it refers to uncertainties that may cause negative project impacts, (similar to many research in the literature (e.g., (Tabatabaee et al. 2021))

2.2 Towards potential cost overrun prediction models for TPPPs

There are several tools and techniques for predicting project costs and contingencies. These are log-logistic probability distribution (Love et al. 2013, 2014, 2015, 2016), Monte Carlo simulation (MCS) (Bouayed 2016; Chang and Ko 2017; Traynor and Mahmoodian 2019a), artificial neural networks (ANN) (Attalla and Hegazy 2003; Elmousalami 2020a; b; Matel et al. 2019; Tijanić et al. 2020), multiple linear regression (Thal et al. 2010), step-wise regression (Diab et al. 2017), fuzzy logic (Islam et al. 2021; Salah and Moselhi 2015), and some integrated approach like probabilistic risk analysis and MSC (Shahtaheri et al. 2016). The following section discusses the applications and limitations of these models to find their potentiality in the cost overrun modeling in TPPPs.

The regression analysis, ANN, and probabilistic cost prediction models are data-intensive and depend on numerical cost data from similar previous projects. Besides, ANN has some particular limitations, such as selecting the optimal number of hidden layers, and uncontrolled black-box in the training of the neurons significantly influence prediction outcomes. Probabilistic models (Love et al. 2013, 2014, 2015, 2016) look for the best-fit probability distribution of the historical and numerical cost or cost overrun data, limiting their application to PPPs in general as they vary significantly from project to project (Hashemi et al. 2019). Besides, some studies critiqued Love et al.'s (2013, 2014, 2015) probabilistic models that they do not accommodate behavioral or conscious bias although most of the megaprojects are intentionally under budget in development stages and inflated in execution phases (Ansar et al. 2014; Flyvbjerg et al. 2002, 2018). The fuzzy-based models can accommodate subjective bias in risk assessment and cost prediction models (Salah and Moselhi 2014). Accordingly, Islam et al. (2021) proposed a fuzzy-Bayesian contingency cost prediction model for TPPPs. Their study was limited to addressing risk impacts on contingency cost modeling during preliminary budgeting instead of predicting cost overruns. However, the fuzzy models lack an appropriate method for aggregating subjective and linguistic evaluation, and their transfer of linguistic data to numerical values was also critiqued. Besides, the Bayesian belief networks can only handle the causal relationships among the risks developed by the experts, which are subjected to different biases. The MCS is a widely accepted tool for cost or contingency cost prediction under uncertain project environments (Afzal et al. 2020; Bouayed 2016; Maronati and Petrovic 2019; Touran and Lopez 2006; Traynor and Mahmoodian 2019b). Overall, MCS was applied effectively when huge quantitative data sets from similar previous projects are available, while its integration with other methods can improve its performance when such data is not available.

Some machine learning models (regression neural network (RNN), genetic algorithm (GA), genetic programming) are the potential to handle natural relationships among the variable for predicting expected outcomes. The RNN model can be used for nonlinear relationships among the numerical variables for the project's cost prediction (Car-Pusic et al. 2020; Yip et al. 2014). While the model performs better even for limited data sets, numerical cost data from past similar projects should be used. A coupled ANN-GA model (Hashemi et al. 2019) shows better performance than using only ANN for conceptual cost estimation of PPPs, where GA assists in selecting the best neural architecture. However, like other ANN models, it also requires sufficient numerical data sets from previous similar projects (Hashemi et al. 2019).

The PPPs are uncommon, and characteristically dissimilar due to geological features, demand-based design (i.e., fuel sources and supply, electricity demand in the locality (MW), variations in plant equipment), financial arrangement (single source, or multiple sources), etc. (Hashemi et al.

2019; Islam et al. 2019). Thus, access to the quality cost data sets and objective risk assessment for PPPs is unrealistic. This leads to export-dependent subjective risk assessment and cost data sets. Therefore, further study must find a suitable tool that can accommodate subjective data sets and natural relationships among the variables to predict project cost and/or contingency.

The GP is a powerful programming tool that uses a set of genes based on a Genetic Algorithm to imitate human's way of solving critical problems (Ahvanooey et al. 2019). Thus, it can accommodate subjective bias PPPs (Flyvbjerg et al. 2018). It also determines relationships among input variables but does not reveal the output prediction process (Ahvanooey et al. 2019). Thus, unlike Bayesian networks used in the study conducted by Islam et al. (2021), GP can be used to establish unbiased risk relationships for predicting project cost overruns. Accordingly, Shahrara et al. (2017) developed a project cost prediction model using a GP algorithm and found a higher accuracy level (84.67% correlation accuracy) for estimating the cost of water and sewerage rehabilitation projects. They did not consider risks as input variables for cost overrun prediction and did not demonstrate the model for TPPPs. Besides, adding MCS to capitalize the power of simulation with the developed GP model is undiscovered, and project cost overrun prediction through coupling GP and MCS is particularly unattended.

From the above discussion, two specific research gaps are clear: (1) cost overrun prediction model that employs critical RFs as input variables for TPPPs is not yet developed or demonstrated. This is of paramount importance as the related risk assessment is imbued with subjectivity; thus, there is a need to have a comprehensive predictive model for taking note of the complex relationship among the risk factors affecting the cost overrun of such projects, and (2) the probabilistic failure of TPPPs in terms of their budgeting, based on the existing critical risk factors impacting such projects, has not been uncovered yet. All these gaps are systematically addressed throughout the paper by demonstrating a novel GP-MCS model for predicting cost overruns of TPPPs. Such cost overruns prediction model as couple GP-MCS can significantly assist the risk managers or decision-makers for controlling potential and critical RFs and subsequent contingency cost management throughout the execution phases of the power plant and similar infrastructure projects. The methodology underlying the GP-MCS coupling is presented in the next section.

3 Methods

Following the identification of the most important RFs causing cost overrun in TPPPs, the hybrid GP-MCS model was developed to predict the cost overrun based on the most significant risk factors. In addition, cost overruns of such projects were simulated using a large dataset gathered from TPPPs in Bangladesh. To this end, first, the most significant risk factors were identified by refining a pool of risk factors discussed in the literature based on experts' opinions. Second, GP was employed to develop a cost overrun predictive equation based on the experts' feedback regarding the significance of risk factors and real data of TPPPs cost overruns. The GP was used due to its multiple offered benefits compared to other methods for equation development. Third, MC simulation — as a powerful and commonly-used quantitative risk assessment approach — was used to simulate the cost overruns in TPPPs projects using thousands of trials. It is worth mentioning that for simulation purposes, the gathered data in this first step (i.e., experts' opinions and real data related to TPPPS cost overrun) and the generated equation in the second step were used. The research flowchart and integration of techniques are described in this section and illustrated in Fig. 1.

3.1 Questionnaire development and data collection

The geographical scope where data was collected is Bangladesh, a fast-growing developing country in TPPPs in Aisa (Gilbert et al. 2017; Harrison et al. 2014; Islam et al. 2019; Maronati and Petrovic 2019). The cost overrun RFs were compiled from the existing literature (Islam et al. 2018b, 2019) and verified for thermal power plants by experts in Bangladesh. The list of RFs contributing to cost overruns in TPPPs is tabulated in Table S2. The experts occupy executive positions in the Bangladesh Power Development Board (BPDB). They evaluated the cost overrun RFs in terms of probability of occurrence, and severity in a structured questionnaire. The linguistic scale from 0 (none) to 6 (extremely high) was presented to the experts for evaluation. Also, the experts provided the percentage cost overruns, project ownership, and contract type for the projects that formed the basis of their evaluation. The questionnaire was randomly distributed to 100 experts. The experts' profiles are presented in Fig. 2. Among them, 64 completed the questionnaires and were used for analysis. Academically, the experts completed a bachelor's degree level of education. Also, they have varying years of work experience in power plant project delivery. The experts' profiles confirm their validity to provide more objective information about power plant project delivery performance (Islam et al. 2019; Li and Wang 2016). Furthermore, all 64 experts revealed that their projects adopted EPC (Engineering, Procurement, and Construction) or Turnkey contract, and are owned by the government (i.e., BPDB). Thus, no further analysis was made considering contract type or project ownership.

3.2 Refinement of critical RFs

Initially, potential RFs affecting the cost overrun of the PPPs were selected from the previous literature (as shown in Table S2). Subsequently, the domain experts were requested to assess the risk based on a quantitative measurement scale. Only the highly influential critical RFs were selected based on experts' judgments, so as to ensure an accurate and useful predictive model has been developed. The influential ones were selected using a defined risk magnitude threshold. A RF with a higher risk magnitude than the threshold was considered *highly important* and selected for predictive model development. The risk magnitude threshold by Mahdiyar et al. (2018) was followed. On a scale of 1–5, "3.5" is defined as the threshold. This threshold value was then adjusted based on the calculated risk magnitudes as explained in subsection 4.1.

3.3 Genetic programming

Koza (1994) authored the GP in 1994, based on the concept of Genetic Algorithm (GA) to produce a non-linear mathematical model output for given input values. The GA and GP were adopted for solving and generating equations, respectively. The initial population of individuals in the GP tree was created randomly (Sadrossadat et al. 2018). These individuals consist of three main parts known as root nodes, function nodes, and terminals nodes. Fig. 3 depicts a GP tree($\sqrt{(T/Q+2)}$), in which the leaves that were located at the end of nodes are called terminal nodes (such as T and Q).

In construction engineering and management research, there are several advantages associated with the utilization of GP against the other types of machine learning-based algorithms or statistical-based techniques, such as linear/non-linear techniques (Yun et al. 2019). For instance, the optimized solutions are produced in the GP algorithm without limiting the length of solutions. Where variables are limited, the GP-based algorithms can unravel the relationships between inputs and output, and accordingly, they are used during the development stage (Emigdio et al. 2017). Additionally, GP-based models can capture the complexity and sophisticated nature of a problem by achieving a comprehensive equation(s) through the evolution of functions during the training and testing stages (Garg and Lam 2015). This means, in comparison with others, when complex interrelationships or nonlinear transformations exist in a problem, the utilization of GP-based models leads to a more reflective solution (i.e., a comprehensive predictive equation for dealing with a particular problem at hand). Notably, the supremacy of the adoption of GP-based models against the other types of machine-learning- or statistical-based techniques can be found in diverse studies (Castelli et al. 2015; Chavoya et al. 2012; Olague and Trujillo 2011; Yamashita et al. 2022). Considering the aforesaid advantages of GP-based models, this study utilized such an algorithm to develop a reflective equation for predicting the cost overrun in the TPPPs.

With the above said, to develop the GP model, an initial population in a GP tree was selected in a non-systematic manner. Subsequently, the first equation was produced, and the accuracy of the generated equation was computed. If the result is suitable, the process is stopped; if not, a new population was then re-produced to obtain the best individuals with the most suitable accuracy by mutation (Gandomi et al. 2015). As illustrated in Fig. 4, the specific steps for developing an accurate GP model are stated as follows:

Step 1. Collecting the required data set. In this study, the dataset was the risk magnitudes of the critical RFs contributing to cost overruns in TPPPs. The data gathered from the experts were used as the inputs. The size of data was considerably prudent when compared to the size of data in other related studies (Chan et al. 2021; Dahiru et al. 2021; Fallahpour et al. 2021; Quintero-Domínguez et al. 2021).

Step 2. Determining the training and testing dataset. In this study, 75% of the data obtained was used as a training dataset, while the rest (25%) was used as a testing dataset. The datasets include the critical RFs determined in the previous step and the cost overrun of the projects.

Step 3. Applying the GP using the training data set. At the end of this stage, a comprehensive equation, which enables a precise calculation of the cost overrun associated with TPPPs based on different variables, was derived.

Step 4. Evaluating the accuracy of the equation derived in the prediction of cost overrun percentage using the testing dataset (Alavi and Gandomi 2011; Mostafavi et al. 2013).

3.4 Monte Carlo simulation

3.4.1 Background

The MCS is a quantitative risk assessment technique for evaluating the deterministic and probabilistic impacts of risks holistically (Mahdiyar et al. 2021). As a result, the MCS has been extensively utilized for evaluating deterministic and probabilistic risk impacts in the engineering field (Arnold and Yildiz 2015; Chang and Ko 2017b; Hollmann 2007b; Maronati and Petrovic 2019; Sadeghi et al. 2010). Considering the model requirements, random sampling techniques were taken into consideration in the MC risk analysis model. Additionally, to obtain accurate results, statistical analyses were undertaken. Specifically, the statistical analysis of historical data must be performed because the MCS is a strong probabilistic technique that is useful when the variables to be analyzed are not deterministic. Considering this, the MC model comprises multiple variables with diverse ranges and distribution functions, which can be seen from Fig. 5 (i.e., illustration of the operation principle of MCS) (Mahdiyar et al. 2021). At every iteration of the MCS model, a random number for each input is selected, albeit the distribution functions of the variables. Before the analysis, the exact number of the required iterations needs to be set out. On the other hand, the range of outputs can statistically be analyzed once all the iterations have been obtained. Therefore, the inputs and outputs in the MCS-based technique are assigned myriad values within a specified range (Mahdiyar et al. 2021). Another strong quality of MCS is that it can prudently incorporate both independent and dependent variables into one model, which leads to more accurate and inclusive results that take into account the interrelationships among the identified variables.

3.4.2 Application of MCS in predicting the cost overruns of PPPs

This section is concerned with elaborating the steps involved in exploiting the MCS technique towards the risk analysis of the TPPPs. To do this, this study followed five steps, as illustrated in Fig. 6, and is stated subsequently.

Step 1. The influential variables that have already been refined together with their corresponding risk magnitudes (that were collected from all the experts involved in the study) values were compiled.

Step 2. This step is concerned with the consideration of each variable's risk magnitude distribution function. Notably, the Risk Simulator was used to calculate the distribution functions of all variables (which is an MS-Excel plug-in feature) (Mahdiyar et al. 2021). The 'best-fit' function in the platform of Risk Simulator was used to apply the distribution function of the variables. The continuous probability distribution (CPD) for all the inputs was taken into account during the use of the mentioned function.

Step 3. In this step, based on the inclusive equation derived from the exploitation of GP, the percentage of cost overrun was simulated using MCS. For this purpose, the Latin hypercube sampling was considered since it privileges simple random sampling (Mahdiyar et al. 2021). It is noteworthy that 10,000 trials were executed in the current study to warrant the analysis of all the combinations of variable values.

Step 4. At this stage, conducting sensitivity analysis was taken into account. Sensitivity analysis is an important aspect of testing the variability of the model's outcomes. In the cost overrun simulation model, the input parameters were uncertain and varied. The probabilities of the RFs depended on expert judgments, and the outcomes (i.e., cost overruns) were simulated based on the probability of RFs. Thus, sensitivity analysis had significance in identifying the critical input parameters (i.e., RFs) that may substantially impact the final estimate of cost overruns. To figure out how much the variation in the cost overrun was explained by the variations in each RF in a dynamic simulation environment, a sensitivity analysis was conducted.

Step 5. As the output of the simulation, a range of cost overrun percentages was obtained—including 10,000 values—so that the minimum, maximum, and most probable cost overrun were interpreted.

3.5 Verification and validation

In order to verify and validate the results obtained; different indices were taken into account as follows:

- As regards the reliability of the results obtained from the pool of experts, the Cronbach reliability test (α) was considered (Mohandes and Zhang 2021). In doing so, α for each of the risk parameters (i.e., probability and severity) was separately calculated. If the calculated α of a developed survey related to each of the considered risk parameters crossed the value of 0.7, then the responses provided by the pool of experts would be considered reliable. Otherwise, the corresponding survey needed to be redistributed to the experts.

- In order to check the validation of the results produced by GP, several statistical tests were considered in this study, as suggested by Yong et. al (2021). The considered tests include mean square error (MSE) and coefficient of determination (R^2).
- The validation of the results produced by MCS is checked by Risk Simulator. As opined by Mahdiyar et al. (2021), once the results have been attained using MCS, there was a need to conduct sensitivity analysis to check the extent to which the outputs were sensitive to the inputs. If the output was seen to be very sensitive to any of the inputs considered, then the simulation procedure was imbued with a problem, thus all the related steps need to be checked.

4 Results

The results are presented in three sections. The first section shows the RFs refinement. The second section presents the developed predictive equation, and the third section shows the findings of cost overruns simulation and the sensitivity analysis.

4.1 Risk factors refinement

Initially, 49 RFs were observed to affect the cost performance of the TPPPs—the risk codes were shown in Table 1 and their descriptions were presented in Table S2. Then, the probability and severity of each RF were computed by averaging the scores provided by 64 domain experts. The risk magnitude was calculated as a product of the probability and severity of the risk. As shown in Table 1, fourteen RFs—those whose risk magnitudes are equal to zero—were considered not influential; so they were immediately removed from the list. The magnitudes of the other remaining 35 RFs range from 2–10. Notably, the threshold risk magnitude value was defined as 7.00 (as explained in subsection 3.2). Among 35 remaining RFs, only eight factors were selected for model development meaning that they were important to predict cost overruns in TPPPs. Table 1 shows the results of this step regarding the selected RFs. Notably, the calculated α was equivalent to 0.7659 and 0.8472 for probability and severity parameters respectively, which indicated good consistency among the pool of experts' responses involved in the study.

4.2 Implementing the GP-based mathematical model

At this stage, the GP was employed to develop a predictive equation (Eq. (1)) for cost overrun. To this end, as the inputs for developing such a predictive equation, a dataset was required including the magnitude of eight selected RFs — defined in the previous step — and the reported actual cost overruns for each case. To make it more explicit, the magnitude of RFs in a particular project and the corresponding actual cost overrun of the respective project constituted one dataset point during the GP model development. Then, having the magnitude of each RF for each case (which were considered as the inputs) and the corresponding actual cost overrun reported by the experts (which was the output), GP model development was undertaken using GeneXproTools 4.0 software, and accordingly, Eq. (1) was derived at the final stage of the GP utilization (as explained in sub-section 3.3). As shown in Eq. (1), the cost overrun was predicted based on the dependent variables (which were the magnitude of selected RFs).

$$Cost\ Overrun = \left[\left(\cos\left(\sqrt[3]{x_1 x_5}\right)^2\right)(x_8)\right] + \left[\left(\left(\sqrt{Atan(x_2)} + \sin(x_6)\right)\left(x_8 + 4 - \left(\frac{x_1}{3.37}\right)\right)\right) + 4\right] + \left[\left[\sqrt{x_7 \sqrt[3]{x_7 x_6}}\right]\right] + \left[\cos\left(\sqrt[3]{x_5}\right)\right] + \left[\left[\left(\frac{\sin(3.53 x_3)}{x_2 - 1.93}\right)(x_4)\right]\right] \quad (1)$$

where $x_1, x_2, x_3, x_4, x_5, x_6, x_7$, and x_8 are Mng_Cntr-MW, Ctr_LKE, Owner_DPTP, Mng_CMW, Ctr_PPS, Cslt_LKE, Owner_GCPC, and Mng_Poor FS, respectively.

As stated in sub-section 3.3, 75% of the collected dataset was applied for finding the best structure. Table 2 shows the most suitable structure—which was the most precise one for training and testing—that was used for the model development.

Fig. 7 shows the performance of the training dataset for cost overrun prediction with the percentage of the baseline cost (x-axis) as compared to the percentage cost overruns (y-axis) experienced in real-life PPPs. The mean square error (MSE) was 3.76 and the R^2 value of the correlation model for prediction and actual cost overruns data was 0.9562. This high R^2 value indicated that almost 95% of the predicted cost overruns variance can be explained by the variance of the actual cost overruns of studied PPPs. The figure also demonstrated that the prediction trend line follows a gentle but linear slope. It was notable to mention that the difference between predicted and real value of cost overrun was very small in the lower tail, i.e., up to 30%, but it was higher in the trend line's upper tail. For instance, if the predicted cost overrun was 20%, the real cost overrun was also the same as 20%; for a further 10% increment of predicted cost overrun, the corresponding real figure was only 2% (i.e., 30% to 22%).

Fig. 8 presents predicted cost overrun (%) performance against reality for testing the data set with an acceptable MSE of 3.65. The R^2 value of the correlation model for prediction and actual cost overruns data was also high, i.e., 0.9609. This high R^2 value indicated that almost 96% of the predicted cost overruns variance can be explained by the variance of the actual cost overruns of studied PPPs. Besides, the line graph showed a gentle but a bit steep slope than the prediction of the training data set shown in Fig. 7. Thus, the differences between real cost overruns and predicted cost overruns were minimized. For example, if the predicted cost overrun was 40% of the baseline cost, the corresponding real figure was 38%; and for 80% predicted cost overrun, it was 77% in reality. This showed a reduction of the gap between predicted and real cost overruns in the upper tail of the graph compared to that of the training data set. The lower tail performances were almost the same for both graphs.

4.3 Cost overrun simulation

In the process of cost overrun simulation, the proposed predictive cost overrun equation, together with the data obtained from 64 experts (refer to Table 3), related to the refined RFs were used. Table 3 shows the expert judgment-based risk magnitude of the most significant factors leading to cost overruns of PPPs. This table also shows the nature of distribution functions of the RFs considering all responses gathered from the experts.

Fig. 9 shows the distribution model of cost overrun simulation. The simulation results showed that the cost overruns can vary from 0 to 182% with the minimum, most probably, and maximum amounts 0.029%, 40.48%, 181.90% successively. Fig. 9 also depicted that the cost overruns in PPPs will not exceed 75% with a 90% confidence level under usually experienced RFs. Also, there was a 50% probability that the cost overruns will not exceed 25%. It indicated that the initial cost estimate with the addition of a 25% extra cost as a contingency to the baseline estimate may result in no cost overrun of a project in 50% cases.

Fig. 10 demonstrates the confidence levels of probability in terms of cumulative frequency (or probability) of experiencing a percentage of cost overrun of each project. The simulated percentage of cost overrun is indicated in the x-axis, and the cumulative frequency of the specific percentage of cost overrun is shown in the y-axis. Fig. 10 also compares actual, predicted, and simulated cost overruns. The cumulative frequency of cost overrun for every estimate (actual, predicted, and simulated) comes together at 40% of cost overruns experienced by the PPPs. Around 90% confidence level can be achieved for all cases by considering a 70% cost overrun for every project. Moreover, the percentages of actual cases, predicted cost overruns (based on Eq. (1)), and 10,000 simulated trials are shown in 10 %-ranges. As it can be seen, 20.31% of actual projects have experienced between 10% and 20% cost overruns. On the other hand, considering the most influential risk factors (that have been used in developing Eq. (1)), — this percentage was predicted to be 17.19%. Likewise, using the magnitude of risk factors and the predictive equation (Eq. (1)), the cost overrun was simulated with 10,000 trials and the outputs show that the cost overrun of 24.3% of trials was within that range.

4.3.1 Sensitivity Analysis

Table 4 shows the extent to which the variation in cost overrun depends on the variations in RFs in descending order of dependence. It shows the sensitivity analysis of the RFs on simulating cost overruns; poor feasibility study risks are the most sensitive factor for producing cost overrun, followed by the contractor's lack of experience and poor planning and scheduling. While the contractor's managerial weakness, consultant's lack of experience, and the owner's delay in the project tendering process have the highest level of risk magnitudes, these factors were comparatively less sensitive for a minor change of input variables to model cost overruns. These outcomes confirmed the reliability of the simulation as it showed an acceptable level of sensitivity of the cost overrun to the variations in variables' values.

5 Discussion

The discussion of the study's findings was structured in different dimensions such as critical RFs considered in this study and other similar studies, distribution patterns of cost overruns data, and outcome of the error analysis of this integrated model and the outcomes of other models proposed in previous studies in the same domain.

In this study, the most important RFs, which were considered for cost overruns modeling, was the contractor's managerial weakness (Mng_Cntr-MW), contractor's lack of knowledge and experiences (Cntr_LKE), and owner's delay in the project tendering process (Owner_DPTP). Others are consultant's managerial weakness (Mng_CMW), contractor's poor planning and scheduling (Ctr_PPS), consultant lack of knowledge and experiences (Cslt_LKE), government's customs policy, and complexity (Owner_GCPC), and poor feasibility study (Mng_Poor FS). It is important to note that the RFs were assessed (i.e., probability and severity) based on the average score of the respondents. However, in a past similar study (Islam et al. 2018b) conducted in Bangladesh, the owner-related issues, e.g., land acquisition delay, delay in project tendering, and decision-making were the top-ranked factors. Besides, contractor-related factors like contractor's lack of knowledge, procurement delay, decision-making delay, and poor planning and scheduling were found to be high-impact factors to cause cost overruns, and the consultant's lack of knowledge and experiences was also paid attention in this regard. Poor feasibility study, contractor's managerial weakness, equipment unavailable in the local market, site constraints, and project complexity were also critical to producing cost overruns of PPPs in Bangladesh (Islam et al. 2019).

Some of the factors from the previously mentioned studies were disregarded because of not identified as significant to cost overruns in PPPs. The main reason for the differences between the results produced from this study and those of others lies in the adoption of different methodological approaches. For instance, Islam et al. (2018b) applied a modified fuzzy group decision-making approach and considered the effects of project phases in risk evaluation. This method can be added with the cost overrun modeling approach for selecting important RFs contributing to cost overruns in PPPs. Islam et al. (2021) considered only four factors, i.e., inflation, construction delay, change order, and improper soil investigation for contingency modeling. The construction delay and change orders were not the primary causes rather, they could be produced by some other factors considered in this study and cost overrun modeling. For example, construction delay in PPPs can be occurred due to the land acquisition delay, owner's delay in project approval decision, fund shortage, unavailability of equipment in the local market, complexity in lifting heavy equipment, etc. and change order can be the reason of design error or changes in design and specification, owner's additional requirement, owner/consultant's lack of knowledge and experience, etc. (Islam et al. 2019). Poor site management by the parties involved, changes in geological conditions, decision-making delays, and restricted site access were top-ranked critical factors for cost overruns of other types of PPPs in Indonesia (Sudirman and Hardjomuljadi 2011b). Delay in site handover, financial difficulties, delay in design and equipment procurement, etc. were some major RFs of Iranian PPPs (Zegordi et al. 2012). Thus, identifying critical RFs to cost overruns was an issue for applying this cost overrun prediction model. However, the model can be applied commonly in power plants and other infrastructure projects as infrastructure projects have some common characteristics of encountering cost overruns.

Previous studies claimed that the cost data of real-world projects follow a variety of distribution patterns. For example, Traynor and Mahmoodian (2019a) used a combination of Triangular, LogNormal, and Normal distribution, Asmar et al. (2011) applied beta distribution, Ayub et al. (2019) used the normalized discrete distribution for modeling costs and contingencies of different construction projects. Besides, the cost data of transit projects (Gurgun et al. 2013) and industrial projects (Barraza et al. 2007) were fitted with normal distribution, while the cost data of nuclear energy projects were assumed to be *distribution-free* (Maronati and Petrovic 2019). From these findings, it was clear that the cost data of different construction projects do not follow a consistent distribution pattern. Our study found that the cost overrun risk data mostly follows *Weibull and Gumble maximum distributions*. It is noted that the domain experts provided the cost overrun risk data using their subjective knowledge. Thus, this is a knowledge area for further exploration as subjective data set follows such distributions as mostly Weibull and Gumble maximum.

The error analysis shows that the cost overrun predictions of both training dataset and testing dataset performance were at a satisfactory level with an MSE of 3.76 and 3.65, respectively. The error analyses of similar previous studies clearly indicated that the percentage errors of different construction cost predictions range from -4% to 50% in some cases. For example, Williams and Gong (2014) found that their models (data mining classification algorithms including radial basis function neural network) had an average of 43.72% prediction accuracy of construction cost overruns. Elmousalami (2020b) proposed a fuzzy-based machine learning approach for the project's cost prediction and found 9% MSE with a 92.9% adjusted R^2 value. In a similar study, Idrus et al. (2011) proposed a fuzzy expert system for infrastructure project's contingency cost prediction with the highest 18% error. Thus, the amount of error in our proposed model was reasonable and acceptable compared to the findings of similar models in previous studies.

Having compared Tables 1 and 4, it can be seen that there may not be a direct relationship between the RFs influence and their explained impacts on cost overrun variations. For instance, Owner_DPTP was ranked as the third most influential RF, while based on the outputs, its variation has almost the lowest possible impact on cost overrun. This interesting finding was justifiable considering the undeniable existence of interdependencies among the RFs. According to these outputs, considering the *cause and effect* relationships, it was believed that although Owner_DPTP is a significant RF, it was an *effect* and its significance stems from its indirect influences on the cost overrun—which were received from *causal* RFs. Thus, the most significant RFs (based on their significance levels) may not necessarily lead to drastic variations in the cost overrun percentage of TPPPs, even though they were of high importance for such projects. The results suggested that the project managers need to reduce the risk magnitude of the factors explaining the most variation in the cost overrun values. Moreover, the concerned stakeholders ought to tilt their attention towards any circumstances leading to the variations of the most influential ones (according to the sensitivity analysis reported in Table 4), rather than the RFs of the highest magnitude (i.e., cost overrun percentage of TPPPs).

6 Conclusion and Recommendation

Cost overrun in complex infrastructure projects such as PPPs is a common feature worldwide. Whilst there have been some studies investigating cost overrun risks of the TPPPs, the current body of literature lacks a systematic approach for predicting and simulating the cost overruns of TPPPs based on the significant RFs and their interrelationships. To fill this gap, this study proposed a novel HPMM—which was based on the integration of machine learning- and probabilistic-based techniques—to meticulously analyze the cost overruns of TPPPs for the first time throughout the current body of literature. Based on the application of the proposed HPMM to the selected TPPPs in Bangladesh, the following conclusions were observed:

1) Based on the responses of the experts involved in the study, eight critical RFs playing a major role in the cost overruns of PPPs were determined. These are contractor's managerial weakness, contractor's lack of knowledge and experiences, owner's delay in project tendering process, consultant's managerial weakness, contractor's poor planning and scheduling, consultant lack of knowledge and experiences, government's customs policy and complexity, and poor feasibility study.

2) Using the proposed GP-based algorithm, an inclusive equation for predicting the cost overruns of related projects was put forward, based on which the cost overrun of related projects can be predicted.
3) Based on the proposed derivative produced from GP, the probabilistic cost overrun of related projects was simulated using MCS technique by considering 10,000 trials. The findings show that the most probable cost overrun was 40.48%, with a maximum of 182%. Besides, the Bangladeshi PPPs will experience 25% cost overruns with a 50% probability, and this will not exceed 75% with a 90% confidence level under usually experienced RFs.
4) The sensitivity analysis showed that cost overrun of PPPs was highly sensitive to a poor feasibility study, contractor's lack of experience, and contractor's poor planning and schedule.

6.1 Research implications

In this research, the most influential RFs have been selected from a comprehensive list based on their impacts on many real-life projects. Moreover, since the cost overrun was simulated based on actual data and rigorous mathematical approaches, the outcomes gave clear overall insight to the managers on the probability of cost overrun in their projects. The developed mathematical model is a practical prediction model, by which managers can predict the cost overrun of their projects based on the actual risks' magnitudes associated with their projects with a high level of accuracy. The developed model can directly be used in the PPPs in Bangladesh and economically and culturally similar countries for their further project budgeting, contingency cost allocation, and risk management-based contingency disbursement. The developed cost overruns prediction approach can be used for any real-life infrastructure projects regardless of geo-political boundaries as the process itself is robust for cost overruns prediction.

6.2 Limitations and suggestions for future studies

The limitations of the study are twofold. Firstly, although the developed HPPM is replicable for other PPPs around the globe, the findings are region-specific and may not be fully adopted in PPPs in other countries. It is worth mention that the cost overrun can be predicted and simulated for projects in other regions by reconsidering the most critical RFs and their magnitudes based on the available records in that specific region. Secondly, the model considered only eight top-ranked factors for modeling cost overruns of complex PPPs. As a result, future studies can be conducted to consider more RFs in predicting the cost overrun of PPPs and compare the accuracy of such prediction models with HPPM. In addition, the findings of the research revealed that there are some relationships between the RFs. Although a highly accurate model is developed in this research, such relationships should be investigated in future studies for possible cost overrun predictive model development. The proposed model can also be implemented or demonstrated in cost overrun modeling of other infrastructure projects of its broader applications. This study also found by a literature review (please refer to subsection 2.2) that the regression neural network (RNN) is a potential tool to consider critical risks as input variables to cost overrun prediction of PPPs, and can work with small data set, which overcomes the limitation of data-intensive models. Thus, the RNN is recommended for further research justifying its capability to cost overrun prediction of TPPPs or similar infrastructure projects. While risks can be defined as threats and opportunities, this study only focuses on the risks, which have threats or negative impacts to project cost and produced cost overruns. Thus, future research is recommended to address risks creating cost-saving opportunities in developing a cost overrun prediction model for TPPPs.

Data Availability Statement

Some or all data, models, or code that support the findings of this study are available from the corresponding author upon reasonable request.

References

Afzal, F., S. Yunfei, D. Junaid, and M. S. Hanif. 2020. "Cost-risk contingency framework for managing cost overrun in metropolitan projects: using fuzzy-AHP and simulation." *International Journal of Managing Projects in Business*, 13 (5): 1121–1139. https://doi.org/10.1108/IJMPB-07-2019-0175.

Ahvanooey, M. T., Q. Li, M. Wu, and S. Wang. 2019. "A survey of genetic programming and its applications." *KSII Transactions on Internet and Information Systems*, 13 (4): 1765–1794. https://doi.org/10.3837/tiis.2019.04.002.

Alavi, A. H., and A. H. Gandomi. 2011. "A robust data mining approach for formulation of geotechnical engineering systems." *Engineering Computations: Int J for Computer-Aided Engineering*, 28 (3): 242–274. Emerald Group Publishing Limited.

Alsharif, S., and A. Karatas. 2016. "A Framework for Identifying Causal Factors of Delay in Nuclear Power Plant Projects." *Procedia Engineering*, 145 (248): 1486–1492. Elsevier B.V. https://doi.org/10.1016/j.proeng.2016.04.187.

Ansar, A., B. Flyvbjerg, A. Budzier, and D. Lunn. 2014. "Should we build more large dams? The actual costs of hydropower megaproject development." *Energy Policy*, 69: 43–56. Elsevier. https://doi.org/10.1016/j.enpol.2013.10.069.

Arnold, U., and Ö. Yildiz. 2015. "Economic risk analysis of decentralized renewable energy infrastructures – A Monte Carlo Simulation approach." *Renewable Energy*, 77: 227–239. https://doi.org/10.1016/j.renene.2014.11.059.

Asmar, M. El, A. S. Hanna, and G. C. Whited. 2011. "New Approach to Developing Conceptual Cost Estimates for Highway Projects." *Journal of Construction Engineering and Management*, 137 (November): 942–949. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000355.

Attalla, M., and T. Hegazy. 2003. "Predicting Cost Deviation in Reconstruction Projects: Artificial Neural Networks versus Regression." *Journal of Construction Engineering and Management*, 129 (4): 405–411. https://doi.org/10.1061/(ASCE)0733-9364(2003)129:4(405).

Ayub, B., M. J. Thaheem, and F. Ullah. 2019. "Contingency Release During Project Execution : The Contractor ' s Decision-Making Dilemma." *Project Management Journal*, 50 (6): 734–748. https://doi.org/10.1177/8756972819848250.

Barraza, G. A., M. Asce, and R. A. Bueno. 2007. "Cost Contingency Management." *Journal of Management in Engineering*, 23 (3): 140–146.

Bouayed, Z. 2016. "Using Monte Carlo simulation to mitigate the risk of project cost overruns." *International Journal of Safety and Security Engineering*, 6 (2): 293–300. WIT Press.

Car-Pusic, D., S. Petruseva, V. Zileska Pancovska, and Z. Zafirovski. 2020. "Neural Network-Based Model for Predicting Preliminary Construction Cost as Part of Cost Predicting System." *Advances in Civil Engineering*, 2020: 5–16. https://doi.org/10.1155/2020/8886170.

Castelli, M., L. Trujillo, and L. Vanneschi. 2015. "Energy consumption forecasting using semantic-based genetic programming with local search optimizer." *Computational intelligence and neuroscience*, 2015. Hindawi.

Chan, K. Y., C. K. Kwong, and H. Jiang. 2021. "Analyzing imbalanced online consumer review data in product design using geometric semantic genetic programming." *Engineering Applications of Artificial Intelligence*, 105: 104442. Elsevier.

Chang, C. Y., and J. W. Ko. 2017. "New Approach to Estimating the Standard Deviations of Lognormal Cost Variables in the Monte Carlo Analysis of Construction Risks." *Journal of Construction Engineering and Management*, 143 (1): 1–7. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001207.

Chavoya, A., C. Lopez-Martin, I. R. Andalon-Garcia, and M. E. Meda-Campaña. 2012. "Genetic programming as alternative for predicting development effort of individual software projects." *PloS one*, 7 (11): e50531. Public Library of Science San Francisco, USA.

Dahiru, A. T., C. W. Tan, S. Salisu, and K. Y. Lau. 2021. "Multi-configurational sizing and analysis in a nanogrid using nested integer linear programming." *Journal of Cleaner Production*, 323: 129159. Elsevier.

Diab, M. F., A. Varma, and K. Panthi. 2017. "Modeling the Construction Risk Ratings to Estimate the Contingency in Highway Projects." *Journal of Construction Engineering and Management*, 143 (8): 1–9. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001334.

Elmousalami, H. H. 2020a. "Artificial Intelligence and Parametric Construction Cost Estimate Modeling: State-of-the-Art Review." *Journal of Construction Engineering and Management*, 146 (1): 03119008. https://doi.org/10.1061/(asce)co.1943-7862.0001678.

Elmousalami, H. H. 2020b. "Comparison of Artificial Intelligence Techniques for Project Conceptual Cost Prediction: A Case Study and Comparative Analysis." *IEEE Transactions on Engineering Management*, 1–14. IEEE. https://doi.org/10.1109/TEM.2020.2972078.

Emigdio, Z., M. Abatal, A. Bassam, L. Trujillo, P. Juárez-Smith, and Y. el Hamzaoui. 2017. "Modeling the adsorption of phenols and nitrophenols by activated carbon using genetic programming." *Journal of cleaner production*, 161: 860–870. Elsevier.

Fallahpour, A., K. Y. Wong, S. Rajoo, and G. Tian. 2021. "An evolutionary-based predictive soft computing model for the prediction of electricity consumption using multi expression programming." *Journal of Cleaner Production*, 283: 125287. Elsevier.

Filho, G. L. T., I. F. S. dos Santos, and R. M. Barros. 2017. "Cost estimate of small hydroelectric power plants based on the aspect factor." *Renewable and Sustainable Energy Reviews*, 77 (April 2016): 229–238. Elsevier Ltd. https://doi.org/10.1016/j.rser.2017.03.134.

Flyvbjerg, B., A. Ansar, A. Budzier, S. Buhl, C. Cantarelli, M. Garbuio, C. Glenting, M. S. Holm, D. Lovallo, D. Lunn, E. Molin, A. Rønnest, A. Stewart, and B. van Wee. 2018. "Five things you should know about cost overrun." *Transportation Research Part A: Policy and Practice*, 118 (December): 174–190. https://doi.org/10.1016/j.tra.2018.07.013.

Flyvbjerg, B., M. S. Holm, and S. Buhl. 2002. "Underestimating costs in public works, error or lie?" *American Planning Association Journal*, 68 (3): 279–295. https://doi.org/10.1080/01944360208976273.

Gandomi, A. H., A. H. Alavi, and C. Ryan. 2015. *Handbook of genetic programming applications*. Springer.

Garg, A., and J. S. L. Lam. 2015. "Improving environmental sustainability by formulation of generalized power consumption models using an ensemble based multi-gene genetic programming approach." *Journal of Cleaner Production*, 102: 246–263. Elsevier.

Gilbert, A., B. K. Sovacool, P. Johnstone, and A. Stirling. 2017. "Cost overruns and financial risk in the construction of nuclear power reactors: A critical appraisal." *Energy Policy*, 102: 644–649. https://doi.org/10.1016/j.enpol.2016.04.001.

Government, Q. 2017. "Project Cost Estimating Manual Seventh Edition." *Queensland Government- Department of Transport and Main Road*, (July).

Gurgun, A. P., Y. Zhang, and A. Touran. 2013. "Schedule contingency analysis for transit projects using a simulation approach." *Journal of Civil Engineering and Management*, 19 (4): 465–475. Taylor & Francis. https://doi.org/10.3846/13923730.2013.768542.

Hadikusumo, B. H. W., and S. Tobgay. 2015. "Construction Claim Types and Causes for a Large-Scale Hydropower Project in Bhutan." *Journal of Construction in Developing Countries*, 20 (1): 49–63.

Hans-Wilhelm. 2016. *World Energy Resources | 2016*.

Harrison, T. J., R. J. Moses, and G. F. Flanagan. 2014. "Cost Overrun risk assessment of power plant project." (October).

Hashemi, S. T., O. M. Ebadati, and H. Kaur. 2019. "A hybrid conceptual cost estimating model using ANN and GA for power plant projects." *Neural Computing and Applications*, 31 (7): 2143–2154. Springer London. https://doi.org/10.1007/s00521-017-3175-5.

Hollmann, J. K. 2007. "The Monte-Carlo Challenge: A Better Approach." *AACE International Transactions*, 03: 1–7.

Hossen, M. M., S. Kang, and J. Kim. 2015. "Construction schedule delay risk assessment by using combined AHP-RII methodology for an international NPP project." *Nuclear Engineering and Technology*, 47 (3): 362–379. Elsevier B.V. https://doi.org/10.1016/j.net.2014.12.019.

Idrus, A., M. F. Nuruddin, and M. A. Rohman. 2011. "Development of project cost contingency estimation model using risk analysis and fuzzy expert system." *Expert Systems with Applications*, 38 (3): 1501–1508. Elsevier Ltd. https://doi.org/10.1016/j.eswa.2010.07.061.

IHS. 2014. "IHS Costs and Strategic Sourcing. Power Capital Costs Index and European Power Capital Costs Index. Available at https://www.ihs.com/info/cera/ihsindexes/index.html."

Islam, M. S., M. P. Nepal, and M. Skitmore. 2018a. "Modified Fuzzy Group Decision Making Approach to the cost overrun risk assessment of power plant projects." *Journal of Construction Engineering and Management*, 145 (2): 04018126-1–15. https://doi.org/https://doi.org/10.1061/(ASCE)CO.1943-7862.0001593.

Islam, M. S., M. P. Nepal, and M. Skitmore. 2018b. "Modified Fuzzy Group Decision Making Approach to the cost overrun risk assessment of power plant projects." *Journal of Construction Engineering and Management*, 145 (2): 04018126-1–15. https://doi.org/https://doi.org/10.1061/(ASCE)CO.1943-7862.0001593.

Islam, M. S., M. P. Nepal, M. Skitmore, and R. Drogemuller. 2021. “Risk induced contingency cost modeling for power plant projects.” *Automation in Construction*, 123 (February 2020): 103519. Elsevier B.V. https://doi.org/10.1016/j.autcon.2020.103519.

Islam, M. S., M. Nepal, M. Skitmore, and G. Kabir. 2019. “A knowledge-based expert system to assess power plant project cost overrun risks.” *Expert Systems With Applications*, 136: 12–32. Elsevier Ltd. https://doi.org/10.1016/j.eswa.2019.06.030.

Kim, M., I. Lee, and Y. Jung. 2017. “International project risk management for nuclear power plant (NPP) construction: Featuring comparative analysis with fossil and gas power plants.” *Sustainability (Switzerland)*, 9 (3). https://doi.org/10.3390/su9030469.

Koza, J. R. 1994. “Genetic programming II: Automatic discovery of reusable subprograms.” *Cambridge, MA, USA*, 13 (8): 32.

Kucukali, S. 2016. “Risk scorecard concept in wind energy projects: An integrated approach.” *Renewable and Sustainable Energy Reviews*, 56: 975–987. Elsevier. https://doi.org/10.1016/j.rser.2015.12.017.

Li, Y., and X. Wang. 2016. “Risk assessment for public–private partnership projects: using a fuzzy analytic hierarchical process method and expert opinion in China.” *Journal of Risk Research*, 9877 (April): 1–22. Routledge. https://doi.org/10.1080/13669877.2016.1264451.

Lin, J., L. Zhu, and K. Gao. 2020. “A genetic programming hyper-heuristic approach for the multi-skill resource constrained project scheduling problem.” *Expert Systems with Applications*, 140: 112915. Elsevier Ltd. https://doi.org/10.1016/j.eswa.2019.112915.

Love, P. E. D., D. D. Ahiaga-Dagbui, and Z. Irani. 2016. “Cost overruns in transportation infrastructure projects: Sowing the seeds for a probabilistic theory of causation.” *Transportation Research Part A: Policy and Practice*, 92: 184–194. Elsevier Ltd. https://doi.org/10.1016/j.tra.2016.08.007.

Love, P. E. D., C.-P. Sing, B. Carey, and J. T. Kim. 2015. “Estimating Construction Contingency: Accommodating the Potential for Cost Overruns in Road Construction Projects.” *Journal of Infrastructure Systems*, 21 (2): 04014035. https://doi.org/10.1061/(asce)is.1943-555x.0000221.

Love, P. E. D., C.-P. Sing, X. Wang, Z. Irani, and D. W. Thwala. 2014. “Overruns in transportation infrastructure projects.” *Structure and Infrastructure Engineering*, 10 (2): 141–159. https://doi.org/10.1080/15732479.2012.715173.

Love, P. E. D., X. Wang, C. Sing, and R. L. K. Tiong. 2013. “Determining the Probability of Project Cost Overruns.” *Journal of Construction Engineering and Management*, 139 (3): 321–330. https://doi.org/10.1061/(asce)co.1943-7862.0000575.

Mahdiyar, A., S. Tabatabaee, A. Abdullah, and A. Marto. 2018. “Identifying and Assessing the Critical Criteria Affecting Decision-Making for Green Roof Type Selection.” *Sustainable*

Cities and Society, 39 (March): 772–783. Elsevier. https://doi.org/10.1016/j.scs.2018.03.007.

Mahdiyar, A., S. Tabatabaee, K. Yahya, and S. R. Mohandes. 2021. "A probabilistic financial feasibility study on green roof installation from the private and social perspectives." *Urban Forestry and Urban Greening*, 58. Elsevier GmbH. https://doi.org/10.1016/j.ufug.2020.126893.

De Marco, A., C. Rafele, and M. J. Thaheem. 2016. "Dynamic management of risk contingency in complex design-build projects." *Journal of Construction Engineering and Management*, 142 (2). https://doi.org/10.1061/(ASCE)CO.1943-7862.0001052.

Maronati, G., and B. Petrovic. 2019. "Estimating cost uncertainties in nuclear power plant construction through Monte Carlo sampled correlated random variables." *Progress in Nuclear Energy*, 111 (November 2018): 211–222. Elsevier. https://doi.org/10.1016/j.pnucene.2018.11.011.

Matel, E., F. Vahdatikhaki, S. Hosseinyalamdary, T. Evers, and H. Voordijk. 2019. "An artificial neural network approach for cost estimation of engineering services." *International Journal of Construction Management*, 0 (0): 1–14. Taylor & Francis. https://doi.org/10.1080/15623599.2019.1692400.

Mohandes, S. R., and X. Zhang. 2021. "Developing a Holistic Occupational Health and Safety risk assessment model: An application to a case of sustainable construction project." *Journal of Cleaner Production*, 291: 125934. Elsevier.

Mostafavi, E. S., S. I. Mostafavi, A. Jaafari, and F. Hosseinpour. 2013. "A novel machine learning approach for estimation of electricity demand: An empirical evidence from Thailand." *Energy conversion and management*, 74: 548–555. Elsevier.

Nevada DOT. 2012. *Risk management and risk-based cost estimation guidelines*.

Olague, G., and L. Trujillo. 2011. "Evolutionary-computer-assisted design of image operators that detect interest points using genetic programming." *Image and Vision Computing*, 29 (7): 484–498. Elsevier.

Quintero-Domínguez, L. A., C. Morell, and S. Ventura. 2021. "A propositionalization method of multi-relational data based on Grammar-Guided Genetic Programming." *Expert Systems with Applications*, 168: 114263. Elsevier.

Sadeghi, N., a. R. Fayek, and W. Pedrycz. 2010. "Fuzzy Monte Carlo Simulation and Risk Assessment in Construction." *Computer-Aided Civil and Infrastructure Engineering*, 25 (4): 238–252. https://doi.org/10.1111/j.1467-8667.2009.00632.x.

Sadrossadat, E., B. Ghorbani, R. Oskooei, and M. Kaboutari. 2018. "Use of adaptive neuro-fuzzy inference system and gene expression programming methods for estimation of the bearing capacity of rock foundations." *Engineering Computations*. Emerald Publishing Limited.

Salah, A., and O. Moselhi. 2014. "Estimating post- and pre-mitigation contingency in construction." *WIT Transactions on Information and Communication Technologies*, 47: 243–250. https://doi.org/10.2495/RISK140211.

Salah, A., and O. Moselhi. 2015. "Contingency modelling for construction projects using fuzzy-set theory." *Engineering, Construction and Architectural Management*, 22 (2): 214–241. https://doi.org/10.1108/ECAM-03-2014-0039.

Shahrara, N., T. Çelik, and A. H. Gandomi. 2017. "Gene expression programming approach to cost estimation formulation for utility projects." *Journal of Civil Engineering and Management*, 23 (1): 85–95. https://doi.org/10.3846/13923730.2016.1210214.

Shahtaheri, M., C. T. Haas, and R. Rashedi. 2017. "Applying Very Large Scale Integration Reliability Theory for Understanding the Impacts of Type II Risks on Megaprojects." *Journal of Management in Engineering*, 33 (4): 04017003. https://doi.org/10.1061/(ASCE)ME.1943-5479.0000504.

Shahtaheri, M., C. T. Haas, and T. Salimi. 2016. "A Stochastic Simulation Approach for the Integration of Risk and Uncertainty into Megaproject Cost and Schedule Estimates." *Construction Research Congress*, 1669–1679.

Singer, L. E., and D. Peterson. 2017. *International Energy Outlook 2017*.

Sovacool, B. K., A. Gilbert, and D. Nugent. 2014a. "An international comparative assessment of construction cost overruns for electricity infrastructure." *Energy Research & Social Science*, 3: 152–160. Elsevier Ltd. https://doi.org/10.1016/j.erss.2014.07.016.

Sovacool, B. K., A. Gilbert, and D. Nugent. 2014b. "Risk, innovation, electricity infrastructure and construction cost overruns: Testing six hypotheses." *Energy*, 74: 906–917. Elsevier Ltd. https://doi.org/10.1016/j.energy.2014.07.070.

Sudirman, W. B., and S. Hardjomuljadi. 2011a. "Project Risk Management in Hydropower Plant Projects: A Case Study from the State-owned Electricity Company of Indonesia." *Journal of Infrastructure Development*, 3 (2): 171–186. https://doi.org/10.1177/097493061100300205.

Sudirman, W. B., and S. Hardjomuljadi. 2011b. "Project Risk Management in Hydropower Plant Projects: A Case Study from the State-owned Electricity Company of Indonesia." *Journal of Infrastructure Development*, 3 (2): 171–186. https://doi.org/10.1177/097493061100300205.

Tabatabaee, S., A. Mahdiyar, and S. Ismail. 2021. "Towards the success of Building Information Modelling implementation: A fuzzy-based MCDM risk assessment tool." *Journal of Building Engineering*, 43. Elsevier Ltd. https://doi.org/10.1016/j.jobe.2021.103117.

Thal, A. E., J. J. Cook, and E. D. W. Iii. 2010. "Estimation of Cost Contingency for Air Force Construction Projects." *Journal of Construction Engineering and Management*, 136 (11): 1181–1188.

Tijanić, K., D. Car-Pušić, and M. Šperac. 2020. "Cost estimation in road construction using artificial neural network." *Neural Computing and Applications*, 32 (13): 9343–9355. https://doi.org/10.1007/s00521-019-04443-y.

Touran, A., and R. Lopez. 2006. "Modeling Cost Escalation in Large Infrastructure Projects." *Journal of Construction Engineering and Management*, 132 (8): 853–860. https://doi.org/10.1061/(ASCE)0733-9364(2006)132:8(853).

Traynor, B. A., and M. Mahmoodian. 2019a. "Time and cost contingency management using Monte Carlo simulation." *Australian Journal of Civil Engineering*, 17 (1): 11–18. Taylor & Francis. https://doi.org/10.1080/14488353.2019.1606499.

Traynor, B. A., and M. Mahmoodian. 2019b. "Time and cost contingency management using Monte Carlo simulation." *Australian Journal of Civil Engineering*, 17 (1): 11–18. Taylor & Francis. https://doi.org/10.1080/14488353.2019.1606499.

Wang, S. Q., and L. K. Tiong. 2000. "Case study of government initiatives for PRC's BOT power plant project." *International Journal of Project Management*, 18 (1): 69–78. https://doi.org/10.1016/S0263-7863(98)00072-6.

Williams, T. P., and J. Gong. 2014. "Predicting construction cost overruns using text mining, numerical data and ensemble classifiers." *Automation in Construction*, 43: 23–29. Elsevier B.V. https://doi.org/10.1016/j.autcon.2014.02.014.

Yamashita, G. H., F. S. Fogliatto, M. J. Anzanello, and G. L. Tortorella. 2022. "Customized prediction of attendance to soccer matches based on symbolic regression and genetic programming." *Expert Systems with Applications*, 187: 115912. Elsevier.

Yip, H. L., H. Fan, and Y. H. Chiang. 2014. "Predicting the maintenance cost of construction equipment: Comparison between general regression neural network and Box-Jenkins time series models." *Automation in Construction*, 38: 30–38. Elsevier B.V. https://doi.org/10.1016/j.autcon.2013.10.024.

Yong, W., J. Zhou, D. Jahed Armaghani, M. M. Tahir, R. Tarinejad, B. T. Pham, and V. van Huynh. 2021. "A new hybrid simulated annealing-based genetic programming technique to predict the ultimate bearing capacity of piles." *Engineering with Computers*, 37 (3): 2111–2127. Springer.

Yoo, W. S., J. Yang, S. Kang, and S. Lee. 2016. "Development of a computerized risk management system for international NPP EPC projects." *KSCE Journal of Civil Engineering*, 21: 1–16. https://doi.org/10.1007/s12205-016-0784-y.

Yun, L., W. Li, A. Garg, S. Maddila, L. Gao, Z. Fan, P. Buragohain, and C.-T. Wang. 2019. "Maximization of extraction of Cadmium and Zinc during recycling of spent battery mix: An application of combined genetic programming and simulated annealing approach." *Journal of Cleaner Production*, 218: 130–140. Elsevier.

Zegordi, S. H., E. Rezaee Nik, and A. Nazari. 2012. “Power plant project risk assessment using a fuzzy-ANP and fuzzy-TOPSIS method.” *International Journal of Engineering, Transactions B: Applications*, 25 (2): 107–120. https://doi.org/10.5829/idosi.ije.2012.25.02b.04.

Z-Flores, E., M. Abatal, A. Bassam, L. Trujillo, P. Juárez-Smith, and Y. El Hamzaoui. 2017. “Modeling the adsorption of phenols and nitrophenols by activated carbon using genetic programming.” *Journal of Cleaner Production*, 161: 860–870. https://doi.org/10.1016/j.jclepro.2017.05.192.

List of figures

Table 1: RF refinement based on their importance to cost overrun

RF code	Risk Factor	Probability*	Severity*	Magnitude	Decision
Mng_Cntr-MW	Contractor's managerial weakness	2.94	3.13	9.18	Select
Ctr_LKE	Contractor's lack of experience	2.83	3.23	9.15	Select
Owner_DPTP	Delay in project tendering process	2.97	3.03	9.00	Select
Mng_CMW	Consultant's managerial weakness	2.77	2.98	8.25	Select
Ctr_PPS	Contractor's poor planning and scheduling	2.67	3.02	8.06	Select
Cslt_LKE	Consultant's lack of experience	2.56	3.00	7.69	Select
Owner_GCPC	Government's customs policy and complexity	2.72	2.77	7.52	Select
Mng_Poor FS	Poor feasibility study	2.48	3.00	7.45	Select
Owner_DDM	Owner's delay in decision making	2.19	3.14	6.89	Reject
Mng_Own-IPM	Incapable project manager for owner	2.12	3.22	6.83	Reject
Owner_CODC	Change order during construction	2.58	2.62	6.77	Reject
Mng_LCBP	Lack of communication between the parties	2.26	2.98	6.74	Reject
Ctr_DDM	Contractor's delay in decision making	2.38	2.81	6.71	Reject
F_CFS	Contractor's fund shortage	2.45	2.73	6.71	Reject
P_SC	Site constraints	2.59	2.58	6.69	Reject
ME_TD	Transportation difficulties	2.39	2.63	6.28	Reject
P_PC	Project complexity	2.39	2.63	6.28	Reject
F_Contr-LFP	Contractor's lack of financial plan	2.41	2.58	6.20	Reject
Mng_CBP	Conflict between the parties	2.38	2.55	6.05	Reject
ME_GIPC	Government's import policy and complexity	2.34	2.53	5.93	Reject
Cslt_ED	Error in design	2.22	2.64	5.86	Reject
Mnp_LE	Lack of experienced mapower	2.23	2.59	5.80	Reject
P_land acquisition	Land acquisition delay	2.45	2.25	5.52	Reject
Own_GIP	Government's interference in procurement	2.25	2.42	5.45	Reject
F_Infl	Inflation	2.28	2.36	5.38	Reject
F_BIH	Bank interest rate high	2.22	2.42	5.37	Reject
Owner_DPGF	Delay to provide government fund	2.14	2.25	4.82	Reject
F_MSF	Multiple sources of fund	2.23	2.09	4.68	Reject
ME_DDPE	Delay to deliver plant equipment at site	1.98	2.11	4.19	Reject
F_OwnerPFM	Owner's poor financial management	1.80	2.11	3.79	Reject
ME_ANER	Additional new equipment required in construction phase	1.75	2.16	3.77	Reject

P_GovtP	Government's policy (e.g., tax and incentive)	1.80	2.00	3.59	Reject
Owner_LBS	Lowest bidder selection	1.94	1.75	3.39	Reject
Cslt_CES	Change of equipment and/or specifications	1.64	1.97	3.23	Reject
ME_IMDS	International market (demand, supply, and price escalation of long-lead items)	1.55	1.78	2.76	Reject
Mnp_LSPS	Lack of skilled personnel at site	0.00	0.00	0.00	Reject
Mnp_PP	Poor productivity of the manpower	0.00	0.00	0.00	Reject
Mnp_LS	Labour shortage	0.00	0.00	0.00	Reject
ME_SME	Shortage of materials and equipment	0.00	0.00	0.00	Reject
ME_FPE	Failure of plant equipment during setup	0.00	0.00	0.00	Reject
ME_UNE	Unfamiliar with new equipment	0.00	0.00	0.00	Reject
Owner_CBS	Complex bureaucratic system	0.00	0.00	0.00	Reject
Ctr_PD	Contractor's procurement delay	0.00	0.00	0.00	Reject
Cslt_DSWD	Consultant's delay to supply working drawing	0.00	0.00	0.00	Reject
Cslt_DIAWE	Consultant's delay to inspect and approve work or equipment	0.00	0.00	0.00	Reject
Env_BW	Bad weather	0.00	0.00	0.00	Reject
Env_UF	Unusual flood	0.00	0.00	0.00	Reject
Env_UC	Unexpected casualty	0.00	0.00	0.00	Reject
Env_EPL	Environmental protection law	0.00	0.00	0.00	Reject
Threshold value		7.00			
	*average value obtained from the responses of all 64 experts.				

Table 2 The most suitable structure of the GP algorithm

Parameters	Value
Number of generation	50-500
Maximum program size	70
Function set	×,/,+, power (x, y*) ,Atan, Cos, Exp,
Initial program size	22
Crossover rate	0.5, 0.80
Homologous crossover	0.83
Mutation rate	0.055
Number of demes	40
Instruction mutation rate	0.4
Data mutation rate	0.2

Table 3. The inputs used for cost overrun simulation

RF	Experts' opinions on risk magnitude		Distributing function*
	Minimum	Maximum	
Mng_Cntr-MW	1	36	Gamma
Ctr_LKE	0	36	Weibull
Owner_DPTP	0	36	Gumble maximum
Mng_CMW	0	25	Gumble maximum
Ctr_PPS	0	36	Weibull
Cslt_LKE	0	36	Weibull 3
Owner_GCPC	0	36	Weibull 3
Mng_Poor FS	0	20	Gumble maximum

* obtained from the best-fit function of Risk Simulator

Table 4. Sensitivity analysis

RF	Per cent variation explained
Mng_Poor FS (Poor feasibility study)	14.41
Ctr_LKE (Contractor's lack of experience)	3.88
Ctr_PPS (Contractor's poor planning and scheduling)	1.44
Owner_GCPC (Government's customs policy and complexity)	0.97
Mng_CMW (Consultant's managerial weakness)	0.33
Mng_Cntr-MW (Contractor's managerial weakness)	0.18
Cslt_LKE (Consultant's lack of experience)	0.07
Owner_DPTP (Delay in project tendering process)	0.02