## Integrated artificial and deep neural networks with time series to predict the ratio of the low bid to owner estimate

Abdulmohsen S. Almohsen, Naif M. Alsanabani, Abdullah M. Alsugair and Khalid S. Al-Gahtani Department of Civil Engineering, College of Engineering, King Saud University, Riyadh, Saudi Arabia

#### Abstract

**Purpose** – The variance between the winning bid and the owner's estimated cost (OEC) is one of the construction management risks in the pre-tendering phase. The study aims to enhance the quality of the owner's estimation for predicting precisely the contract cost at the pre-tendering phase and avoiding future issues that arise through the construction phase.

**Design/methodology/approach** – This paper integrated artificial neural networks (ANN), deep neural networks (DNN) and time series (TS) techniques to estimate the ratio of a low bid to the OEC (R) for different size contracts and three types of contracts (building, electric and mechanic) accurately based on 94 contracts from King Saud University. The ANN and DNN models were evaluated using mean absolute percentage error (MAPE), mean sum square error (MSSE) and root mean sums square error (RMSSE).

**Findings** – The main finding is that the ANN provides high accuracy with MAPE, MSSE and RMSSE a 2.94%, 0.0015 and 0.039, respectively. The DNN's precision was high, with an RMSSE of 0.15 on average.

**Practical implications** – The owner and consultant are expected to use the study's findings to create more accuracy of the owner's estimate and decrease the difference between the owner's estimate and the lowest submitted offer for better decision-making.

© Abdulmohsen S. Almohsen, Naif M. Alsanabani, Abdullah M. Alsugair and Khalid S. Al-Gahtani. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at http:// creativecommons.org/licences/by/4.0/legalcode

The authors want to thank King Saud University (KSU) for funding this research and providing study data.

*Funding:* The authors thank the Deputyship for Research and Innovation, Ministry of Education in Saudi Arabia, for funding this research work through project no. (IFKSUOR3-380-5).

*Disclosure statement:* No conflicts of interest exist. The submitting author is responsible for the co-author's interests.

Author contributions: Conceptualization, Almohsen, Alsanabani, Alsugair and Al-Gahtani; Data curation, Alsanabani; Formal analysis, Alsanabani and Al-Gahtani; Funding acquisition, Almohsen, Alsugair and Al-Gahtani; Investigation, Almohsen, Alsanabani, Alsugair and Al-Gahtani; Methodology, Alsanabani and Al-Gahtani; Project administration, Almohsen, Alsugair and Al-Gahtani; Resources, Almohsen, Alsugair and Al-Gahtani; Software, Alsanabani; Supervision, Almohsen, Alsugair and Al-Gahtani; Validation, Alsanabani and Al-Gahtani; Visualization, Almohsen, Alsugair and Al-Gahtani; Validation, Alsanabani and Al-Gahtani; Visualization, Almohsen, Alsanabani, Alsugair and Al-Gahtani; Roles/Writing: original draft, Alsanabani and Al-Gahtani; Writing: review and editing, Almohsen, Alsanabani, Alsugair and Al-Gahtani.

*Data availability statement:* The raw data that support the findings of this paper are available on request from the corresponding author.

P

Engineering, Construction and Architectural Management Vol. 31 No. 13, 2024 pp. 79-101 Emerald Publishing Limited 0969-9988 DOI 10.1108/ECAM-05-2023-0454

Received 12 May 2023 Revised 1 October 2023 Accepted 21 October 2023

Artificial and deep neural

networks

79

Originality/value – This study fills the knowledge gap by developing an ANN model to handle missing TS data and forecasting the difference between a low bid and an OEC at the pre-tendering phase.
Keywords Contract, Bid, Cost, Deep, ANN, Time series
Paper type Research paper

#### 1. Introduction

Compared to other industries, the construction sector has one of the highest annual business failure rates associated with adverse effects (Chapman, 2001; Mahamid, 2018). Construction management needs to address many uncertainties and risks. The difference between the winning bid and the owner's estimated cost (OEC) is one of these risks through a pre-tendering phase. The winning bid is often the one with the lowest bid submitted to a tender. In the United States, the low bid method is widely utilized to award construction contracts at the price of the lowest responding bid (Gransberg and Gransberg, 2020). A significant variation between the lowest bid (winning bid) and the OEC is problematic for both parties. Such variances may negatively affect contract delays or cancellations, scope reductions and public distrust (Baek *et al.*, 2019). For example, a high bid much higher than the OEC is likely to result in problems with budget allocation, which might delay or terminate the project, while a low bid significantly lower than the OEC could result in cost overruns (Li *et al.*, 2022; WSDOT, 2011). Cost estimation is also crucial for initiating the project and dramatically impacts performance. Therefore, it is essential to provide full justification for cost variation to prevent the misuse of public funds and ensure the greatest possible economic outcome for all stakeholders (Carr, 2005).

The Federal Highway Agency (FHWA 2004) recommended that the difference ratio between the low bid and OEC be within  $\pm 10\%$  (Li *et al.*, 2022). Moreover, the California Department of Transportation (Caltrans) devised a performance metric comparing the OEC to low bids (i.e. low bids within 10% of the OEC) to assess the precision of cost estimation (Caltrans, Planning Cost Estimate, 2006). However, several agencies have needed support achieving and maintaining an acceptable range of low bids to OEC for their highway projects (Li *et al.*, 2021).

Only some studies (Li *et al.*, 2022) attempted to improve the quality of the OEC and enhance estimation processing by performing model forecasting of the ratio of a low bid to OEC (R) using time series (TS) techniques. However, this model was only used for highway construction, and it needed to consider the amount of the contract cost and different contract types.

For this paper, the R data for 94 contracts was collected from 2010 to 2021 for the building, electric and mechanic contracts. However, there is a significant shortage of information on the TS of the R for the different OECs of the three types of contracts. This shortage causes the deep neural networks (DNN) to be unable to predict future TS using long-short-term memory (LSTM), which has merit in dealing with the non-stationary and nonlinear characteristics of a TS (Bala and Singh, 2019). The integrated artificial neural networks (ANN) model is developed to address this lack of information issues. The techniques of maximizing data (Zaved, 2001) and improving accuracy (Pasini, 2015) were integrated into the developed ANN model. The input layer of the developed ANN model has a specific pattern or configuration to allow the generation of the TS for different amounts of OEC for the three contract types from 2010 to 2021. The improvement of input layer methodology has not been studied before and contributes to this paper. The DNN model was used to forecast the future TS using this generated data. The configuration that followed in the input layer with the techniques used in the developed ANN model will assist researchers and engineers in establishing a reasonable TS for different applications. Moreover, the paper's results will improve the accuracy of the OEC for assessing the contract cost at the pre-tendering phase and preventing future problems that develop during the construction phase (such as owner financial difficulties or cost overruns for contractors). The significance of the study is enhanced cost estimation, improved decision-making, early identification of cost deviations and optimization of bidding strategies.

ECAM

31.13

#### 2. Literature review

Several studies dealt with the difference between the low bid and OEC through the pretendering phase was investigated by examining and investigating the causes that increase these differences (Alsugair, 2022; Flyvbjerg *et al.*, 2003; Saqer *et al.*, 2020). On the other hand, several studies were considered in the forecasted model of cost deviation; regression analysis was used by Li *et al.* (2021) to measure the effects of influential factors on the cost deviation and identify factors impacting it. The explanatory model was developed using previously collected costs between the years 2011 and 2015 for Louisiana highway construction contracts. They stated that the level of bidding competition significantly influences cost deviation, the scope of the contract, the number of activities, the crude oil price and the value of the paving projects. Li *et al.* (2022) analyzed and investigated the identifying risk variables affecting the accuracy of the client's estimate for highway projects to predict the ratio of a low bid to the OEC using the TS model.

TS forecasting can be performed in various ways, commonly grouped into traditional statistical and nonlinear models. The first categories represent a linear analysis of the previous observations and include average, exponential smoothing and autoregressive integrated moving averages (ARIMA). Table 1 shows the studies of the TS for traditional statistics. The nonlinear method aims to overcome the linear limitation of TS. Khashei and Bijari (2011) summarized the methods as the bilinear model, the threshold autoregressive (TAR) model, the autoregressive conditional heteroscedastic (ARCH) model, general autoregressive conditional heteroscedastic dynamics and ANN. Table 2

References	Method	Application
Hwang (2011) Corrêa <i>et al.</i> (2016)	ARMA Auto-Regressive Integrated Moving Average with eXogenous variables and Generalized Auto-Regressive Conditional Heteroscedasticity (WARIMAX-GARCH)	Construction cost Information technology
Zhao <i>et al.</i> (2020) Zhao <i>et al.</i> (2019) Rubio <i>et al.</i> (2016)	Casual method + Seasonal ARIMA (SARIMA) Exponential smoothing models (ESM) + SARIMA Fuzzy Time Series (FTS)	Building cost index Building cost index Economic applications
Naim <i>et al.</i> (2018)	BATS (Exponential smoothing state space with Box-Cox transformation, ARMA errors, Trend and Seasonal components) + TBATS (Trigonometric Exponential smoothing state space with Box-Cox transformation, ARMA errors, Trend and Seasonal components)	Natural gas consumption
Source(s): Author	s' own work	

References	Method	Application	
Almonacid et al. (2013)	ANN	Energy science	
Camelo <i>et al.</i> (2018)	ANN	Wing generation	
Iwana and Uchida (2021)	(DNN)	Data augmentation of TS	
Kardakos et al. (2013)	ANN	Power generation	
Khashei and Bijari (2011)	ARIMA + ANN	Information technology	
Withington et al. (2021)	ANN	Expert system application	
Pai and Lin (2005)	ARIMA + Support vector machines model (SVM)	Stock price	Table 2
Chen and Wang (2007)	SARIMA + SVM	Industry application	Research that applied
Khashei et al. (2009)	ANN + ARIMA + Fuzzy logic	Information technology	the nonlinear method
Source(s): Authors' own	work		to TS

Artificial and deep neural networks

Table 1.

in TS

Studies dealt with the traditional method

summarizes the research of the nonlinear methods on TS. In addition, traditional TS approaches like ARIMA, SARIMA (Seasonal ARIMA) and ETS (error, trend, seasonality model) are made to manage a TS with a single seasonality; however, when several seasons occur, these techniques do not perform as well (Naim et al., 2018).

Rashid and Louis (2019) illustrated the merits of using DNN. They stated that recent developments in DNN, specifically recurrent neural networks (RNN), present new opportunities to classify sequential TS data with recurrent lateral connections. In addition, Bala and Singh (2019) stated that a TS's non-stationary and nonlinear characteristics could be learned by a LSTM network, reducing predicting error. The LSTM is utilized as one layer in DNN. Therefore, the DNN was utilized to forecast TS.

TS often suffer from missing data and are thus difficult to use in forecasting. There are conventional and modern techniques to deal with the missing data. Table 3 shows these methods. However, the methods may not handle a relative amount of missing data, affecting the analysis's accuracy. In addition, some of the methods, such as DNN, require big data for missing data treatment.

In this paper, the ANN was developed to deal with missing data and ensure the handling of missing data without affecting the data quality. On the other hand, the limited and insufficient availability of the required data represents a challenge in its usage of ANN. Maweu et al. (2021) stated that data scarcity and class imbalance are common occurrences in healthcare datasets and undermine the classification performance of machine learning models. The maximize data technique (Pasini, 2015) and the improved data quality method (Zaved, 2001) were used to overcome the minor data issues. However, the two techniques may not be adequate for TS data. The new contribution in this paper, the time variable, which was changed from 2010 to 2021, was decomposed into 12 variables (one for each year). These variables were changed to zero or one depending on the used data. Therefore, the ANN was utilized to compute TS from 2010 to 2021 based on the 94 collected data after implementing the maximize data innovation (Pasini, 2015) and evaluated using mean absolute percentage error (MAPE), mean sum square error (MSSE) and root mean sum square error (RMSSE). The reason ANNs are used to generate TS is that one key benefit of ANN models over other types of nonlinear models is that they are universal approximators that can estimate a broad class of functions with high accuracy. Their strength derives from the information in the data being processed in parallel (Khashei and Bijari, 2011).

	Reference	Method
	<i>Conventional techniques</i> Andridge and Little (2010) Strike <i>et al.</i> (2001), Dhevi (2014) Little and Rubin (2019)	Hot and Cold Deck Imputation Mean Imputation Multiple Imputation
	Modern techniques Lobato et al. (2015), Aydilek et al. (2013), Azadeh et al. (2013) Wu et al. (2015) Shao et al. (2014) Banbura et al. (2014) Amiri et al. (2016) Sitarom et al. (2015)	Genetic Algorithm Optimization Based Support Vector Machine Interpolation Maximum Likelihood Fuzzy-Rough Set Similarity Measure
Table 3.         Handling techniques of time series	Shara et al. (2013)           Zhang et al. (2022)           Torres et al. (2021)	Bayesian Dynamic Regression DNN
missing data	Source(s): Authors own work	

82

ECAM

31.13

#### 3. Methodology

The methodology utilized ANN and DNN to predict a contract cost more accurately. Therefore, it mainly consists of five steps: (1) collect data to create the database; (2) implement size and normality test to ensure the collected sample represents the actual sample; (3) develop an ANN to compute additional current TS; (4) generate TS; (5) establish a DNN to forecast the future TS. The flow chart of the methodology is shown in Figure 1.

#### 3.1 Collect data

The OEC and contract value difference is affected by many factors and causes that differ from one country to another and from one work environment to another. Therefore, limiting the study to a specific environment was necessary to facilitate the study by taking information and neutralizing some known factors (including administrative procedures) and unknown factors. Thus, KSU's (King Saud University) projects in Riyadh, KSA (Kingdom of Saudi Arabia) determined the field of study.

The cost data of cost estimation accuracy covers 94 projects completed at KSU between 2010 and 2021. The projects are classified as building, highway, electric and mechanic. The data contained the initial estimated cost, year of the award, contract amount and project type. Table 4 presents 94 project data. The initial estimated cost is difficult to acquire. The absolute cost deviation can be estimated as Eq. (1), which is shown in the sixth column.

$$R = \frac{Low \, bid}{Owner'estimate \, cost} \tag{1}$$



End



Collect data Implement size and normality test Develop ANN Establish ANN model' structure Maximize data First analysis Mun ANN model Assess the ANN model Improve the training data Second analysis Mun ANN model Assess the ANN model Concerne TS from2010 to 2021 using ANN model Artificial and deep neural networks



ECAM 31,13	No.	Project	Time	CC (M SAR)	OEC (M SAR)	R	No.	Project	Time	CC (MSAR)	OEC (M SAR)	R
	1	Building	2013	0 207	0 250	0.830	48	Building	2019	0 498	0 499	0 999
	2	Dununig	2013	0.285	0.285	1 000	49	Dunung	2018	0.285	0.290	0.983
	3		2014	0.241	0.250	0.965	50		2018	0.200	0.297	1 000
	4		2014	140727	200.000	0.300	51		2010	0.297	0.297	1,000
84	5		2014	0.404	0.990	0.408	52		2017	0.237	0.300	0.950
04	- 6		2014	0.321	0.380	0.400	53		2017	0.200	0.239	1.037
	7		2014	0.155	0.240	0.644	54		2018	0.230	0.230	1,000
	8		2017	0.075	0.075	0.996	55		2017	0.200	0.450	0.550
	9		2017	32 142	35,000	0.918	56		2017	0.469	0.490	0.957
	10		2017	0.288	0.288	1 000	57		2018	0.499	0.499	1 000
	11		2018	0.018	0.018	1.000	58		2018	0.478	0.478	1,000
	12		2018	0.168	0.180	0.932	59		2018	0.451	0.451	1,000
	12		2018	0.100	0.100	0.885	60	Flectric	2018	2 849	3 500	0.814
	14		2018	0.087	0.610	0.143	61	Liccure	2018	0.227	0.260	0.872
	15		2018	0.480	0.510	0.960	62		2018	0.105	0.135	0.778
	16		2010	0.400	0.300	0.988	63		2010	0.105	0.100	0.078
	10		2018	0.207	0.200	0.986	64		2013	0.290	0.400	0.967
	18		2018	0.050	0.700	0.997	65		2018	0.295	0.300	0.983
	10		2010	0.200	0.180	1 000	66		2010	5 782	6,000	0.964
	20		2010	0.100	0.100	0.998	67		2013	0.464	0.000	0.928
	20		2013	0.403	0.495	0.000	68		2020	0.404	0.500	0.920
	21		2013	0.451	0.435	0.000/	69		2020	0.402	0.300	1,000
	23		2019	0.402	0.400	0.989	70		2017	0.300	0.300	1,000
	20		2019	11 835	10,000	1 184	71		2017	0.006	0.120	0.802
	25		2019	0.232	0.299	0.776	72	Mechanical	2010	8.850	10.00	0.885
	26 26		2013	0.497	0.498	0.998	73	Wieenamear	2010	21 49	24.00	0.895
	27		2014	0.210	0.230	0.912	74		2011	13.96	10.00	1.396
	28		2014	0.464	0.200	0.927	75		2011	0.353	0.353	1,000
	29		2015	0.479	0.490	0.978	76		2017	0.260	0.270	0.961
	30		2015	0.320	0.350	0.915	77		2017	0.200	0.299	0.973
	31		2015	0.769	0.800	0.961	78		2017	0.260	0.270	0.961
	32		2015	0.493	0.498	0.989	79		2017	0.593	0.593	1 000
	33		2015	0.498	0.499	0.998	80		2017	0.259	0.285	0.910
	34		2015	0.700	0.800	0.875	81		2017	0.296	0.298	0.992
	35		2016	0.492	0.498	0.987	82		2018	0.072	0.075	0.955
	36		2017	0.260	0.275	0.945	83		2018	0.036	0.040	0.888
	37		2018	0.221	0.240	0.922	84		2018	0.593	0.600	0.989
	38		2018	0.186	0.200	0.929	85		2018	0.296	0.300	0.987
	39		2018	0.223	0.235	0.947	86		2018	0.042	0.061	0.683
	40		2018	0.097	0.099	0.979	87		2018	0.036	0.038	0.934
	41		2018	0.040	0.042	0.950	88		2018	0.090	0.090	1 000
	42		2019	0.475	0.480	0.989	89		2018	0.072	0.120	0.597
	43		2019	0.375	0.390	0.962	90		2019	0.229	0.320	0.717
	44		2019	0.067	0.070	0.963	91		2019	0.375	0.375	1.000
	45		2021	34.669	35.000	0.991	92		2019	0.229	0.320	0.717
	46		2021	59.139	65.000	0.910	93		2018	0.460	0.490	0.938
Table /	47		2019	0.469	0.469	1.000	94		2019	0.116	0.135	0.853
Collected data	Sou	<b>rce(s):</b> Au	thors' o	own work								

*3.2 Implement size and normality test* As the sample space (construction projects) is ample and unknown, the sample size can be computed using Eq. (2) (Badawy *et al.*, 2022).

Sample size = 
$$\frac{Z^2 p(1-p)}{C^2}$$
 (2) Artificial a deep neu network

where Z is a value corresponding to a 95% confidence level and is equal to 1.96 and p represents the probability choice of 0.5. C is the confidence interval, which should be less than 0.2 (Badawy et al., 2022). Therefore, the minimum sample size for a confidence level of 95% was 44, which was less than the number of data (97 data). The low bid to OEC data ratio should be a normal distribution regarding project classifications. Hence, the data were tested using Kolmogorovand Shapiro tests in SPSS. The results revealed that the significant value of the two tests for the building, electric and mechanic projects was less than 0.05, as shown in Table 5. Therefore, the cost deviation for the three project types followed a normal distribution. However, the two tests cannot apply to highway projects due to the limited number of highway projects.

#### 3.3 Develop ANN model

The primary purpose of the developed ANN model is to compute the TS of the ratio of the lower bid to OEC in several OEC and project types. The ANN model generally consists of three layers: input, hidden and output layer. The hidden layer also comprises one or two layers with different nodes (neurons). ANN is a method that computes the output by learning an algorithm from any function (Loy, 2019). The benefits of applying the ANN are simplicity and enables dealing with the nominal data such as the project type and time in this paper. The SPSS IBM software can sketch an ANN model with bias values and weight connections between neurons. Also, the SPSS IBM software, which offers relative errors at the two data with anticipated result values, simplifies choosing the training and testing data percentage.

3.3.1 Establish the ANN model' structure. Data are split into training and testing datasets. Data from January 2010 to December 2017 are used for fitting bid distributions and training the neural network. The constructed network forecasts low bids in 2018 (out-of-sample predictability). No unique relation or rule controls the number of neurons at hidden layers. Zaved (2001) suggested a formula to determine the neurons of hidden layers as 2m + 1, where m is the number of neurons at the input layer. The data utilized in the input layer was OEC, time (year of award) and project time, while the R was considered the output. The time ranged from 2010 to 2021 and was considered nominal data in the ANN model as factors whose value changed to zero or one. Therefore, the time can be represented in the ANN as 12 neurons (neurons for each year). In addition, the project types were considered as the nominal data and considered in the input layer of the ANN model as three neurons: B (building), E (electric) and M (Mechanic). Hence, the number of neurons (m) was 16 (1 for OEC, 12 for time and 3 for project types).

	Kolmo	gorov-Smirr	10V <sup>a</sup>	S	hapiro–Wilk	
Project type	Statistic	df	Sig.	Statistic	df	Sig.
Building	0.296	59	0.000	0.560	59	0.000
Electric	0.282	12	0.009	0.829	12	0.020
Mechanic	0.227	24	0.002	0.782	24	0.000
Note(s): a. Lillie Source(s): Auth	fors significance co nors' own work	prrection				

nd ral networks

85

Regarding the hidden layer, two layers with thirty-three neurons of each layer (2m + 1) were considered in the ANN model. The ANN model's structures are shown in Figure 2. For example, when using data from the project, No. 1, as shown in Table 4, was used in the ANN model, the neurons of the OEC, B, and 2013 were set as 0.207, 1, and 1, respectively, while the other neurons of the input layers were set as zero. On the other hand, the R' neuron was set as 0.830 in the output layer.

*3.3.2 Maximize data.* Because the data were respectively small, consisting of 94 data sets, the data was augmented using the method introduced by (Pasini, 2015). The data was divided, in this paper, into 10 subgroups, and one of them was considered as test data while the other was train data and inserted as fully train data. Based on the location of the test data subgroup, the 10 train data groups were generated, as shown in Figure 3. The first and second analyses were implemented, as illustrated in the following section.

3.3.3 First analysis. The first analysis consists of three steps: running the ANN model, assessing the ANN model and enhancing training data.

3.3.3.1 Run ANN model. The ANN model was run several times for each train data group, and the ANN model for each group was then taken as the average computed of the R ( $R_{com}$ ). Notably, the structure of the ANN model for the 10 train groups is the same. However, the weight values of the connection among the layers were different. Therefore, the ANN models for the 10 groups differed; there were 10 ANN models (ANN model per train data group).

3.3.3.2 Assess the ANN models. The ANN model was evaluated using three statistic indicators: MAPE, MSSE and RMSSE. The formula of the three indicators was shown in Eq. (3)–(5), respectively, as:



86

ECAM

31.13

Figure 2. ANN model' structures



$$MAPE = \frac{1}{n} \sum_{i=1}^{n} \frac{|R_{obs} - R_{com}|}{R_{obs}} \times 100$$
(3)

$$MSSE = \frac{\sum\limits_{i=1}^{n} (R_{obs} - R_{com})^2}{n}$$
(4)

$$RMSSE = \sqrt{\frac{\sum_{i=1}^{n} (R_{obs} - R_{com})^2}{n}}$$
(5)

where  $R_{obs}$  and  $R_{com}$  are the observed and computed of the R, respectively, and n is the data set number in the train data group.

3.3.3.3 Improve the training data. For each training data group, the training data was improved by deleting the abnormal data, which provided a significant error. The purpose of deleting the data with a critical residual error. Hence, the absolute percentage error (APE) was utilized to identify the abnormal data, and the APE can be computed using Eq. (6) as:

$$APE_i = \frac{|R_{obs-i} - R_{com-i}|}{R_{obs} - i} \tag{6}$$

where  $R_{obs-i}$  is the observed R of ith contact, and  $R_{com-i}$  is the computed R of the ith case (contract) by the ANN model. Badawy (2020) considers 0.2 as a threshold value for the allowable and not allowable relative error. In this paper, the data with an  $APE_i$  value of more than 0.2 is considered abnormal and deleted from the train data group. After the deletion of the abnormal data, the modified train data group was established and utilized in the second analysis.

E	CAN	I
31	.13	

88

3.3.4 Second analysis. The ANN model was trained through the ten modified train data groups and performed the 10 ANN models. The output of the models was evaluated using MAPE, MSSE and RMSSE. One of the 10 ANN models was used to generate the TS.

#### 3.4 Generate TS from 2010 to 2021 using the ANN model

The appropriate ANN model was utilized to generate the TS of the R from 2010 to 2021 for the three types of projects (building, electric and mechanic) and several OECs (10,000 SAR, 100,000 SAR, 1,000,000 SAR, 10,000,000 and 100,000,000SAR).

#### 3.5 Forecast TS using DNN

DNN was utilized for each TS using the ANN model to predict the future TS. The DNN structure consists of the five-layer input, LSTM, drop, full connection function and regression, as shown in Figure 4 (MATLAB, 2021). The input layer represents the TS generated by the ANN model. The data type is a sequence due to the nature of the TS. Regarding the LSTM layer, the RNN has the issue of vanishing gradient learning; gradient learning represents the primary component of principle learning. Hochreiter and Schmidhuber (1997) designed the LSTM to overcome the vanishing gradient issue. An LSTM layer learns the long-term relationships between the sequence data and the time step in the series. The layer's addition function can improve gradient flow over extended training sequences. In addition, the DNN has a large number of layers. However, overfitting is a serious concern in such networks since merging the prediction of many outcomes is difficult. Thus, Drop is considered the problem-solving technique in such cases. The fundamental idea is to remove weight values randomly and of the connections from the neural network while it is being trained. As a result, units are prevented from over-co-adapting (Srivastava et al., 2014).

The TS was divided into training and testing TS at 65 and 35%, respectively. The training and testing data were standardized based on the mean of the ratio  $(R_{mean})$  and standardized deviation (std), as shown in Eq. (7):

$$R_{standradized-i} = \frac{R_i - R_{mean}}{std} \tag{7}$$

The DNN model was evaluated by computing the RMSSE between  $R_{obs}$  and  $R_{for}$  throughout the testing process using Eq. (3). It should be noted that the output data of the DNN should be unstandardized before they implement the evaluation processes. Unstandardized output data can be performed using Eq. (8) as

$$R_{for} = \left(R_{standardized-i}^{for}\right)(std) + R_{mean} \tag{8}$$



Figure 4. DNN processes used in the TS within MATLAB



#### 4. Results and discussions

There is a question about using TS to predict missing data instead of the ANN. The impact of missing data on TS analysis increases as the proportion of missing values grows. The TS can handle low missing data (5%) accurately. At the same time, accuracy degenerates when the proportion of missing values exceeds 10% (Junger and Leonm, 2015). In this paper, the missing data for TS after taking the average R per year is shown in Table 6. Therefore, the TS data used in this paper has a significant amount of missing data, reaching at least 30%. Therefore, TS cannot handle the extensive missing data (more than 10%).

Figure 5 shows the results assessments of the ten ANN models for the first and second analyses in terms of MAPE; its values in the first analysis range from 4% to 14.485. On the other hand, the MAPE values of the second analysis do not exceed 5.2%, which is close to 5%. It is included that the ANN models' accuracy is very accurate based on the accuracy classification. In addition, Figure 8 also shows the significant reduction of the MAPE of the models between the first and second analyses. In other words, deleting the abnormal data in the second analysis increases the accuracy of the ANN models. The ANN3 model provides the minimum value of MAPE for the two analyses. On the other hand, the ANN8 and ANN2 provided the maximum MAPE value for the first and second analyses, respectively.

For evaluating the ANN models in terms of MSSE and RMSSE, Figure 6 displays the MSSE of the ten ANN models. The two analyses' values were generally close to zero, indicating that the observed and computed R values were identical. For the first analysis, the MSSE was maximum at ANN8 (0.0154) and minimum value at the ANN1 model. Furthermore, for the second analysis, the MSSE of the ANN models was less than 0.002, except for the ANN2 model, whose value was 0.005. Figure 7 shows the RMSSE of the ten models for the two analyses. The minimum and maximum of the RMSSE for the first

Project type	Building	Electric	Mechanic	Table 6.
Proportion of missing data	33%	41.67%	58.33%	Proportion of missing data for different
Source(s): Authors' own work				project type



**Figure 5.** MAPE of different

ANN models for first and second analysis

Artificial and deep neural networks

89

# ECAM 31,13

90

Figure 6. MSSE of different ANN models for first and second analysis







**Figure 7.** RMAPE of different ANN models for first and second analysis

analysis were 0.007 (ANN1) and 0.12 (ANN8), respectively. Moreover, the ANN1 and ANN2 provide the minimum (0.024) and maximum values (0.071) for the second analysis.

The average MAPE for the first and second analyses was 11.11 and 2.94%, respectively, as shown in Figure 8. Deleting the abnormal data reduces the MAPE by 8.17% on average. In terms of MSSE, the average value of the MSSE and RMSSE was 0.109 and 0.104, respectively, for the first analysis. In addition, the value for the second analysis was 0.0016 and 0.039. Deleting the abnormal data in the second analysis decreases the MSSE and RMSSE by 0.108 and 0.065 on average, respectively. Based on the high accuracy of the ANN models, especially in the second analysis, the ANN models can be utilized to compute the TS of the R from 2010 to 2021 for different OECs.

The ANN2 model was considered to generate TS of the R due to several reasons: (1) the data training contains all years (2010–2021), the other ANN models have lost in one or more years, (2) the MAPE, MSSE and RMSSE of the ANN2 models provide a maximum value



compared the other model. Hence, the computed R values are a conservative and upper estimate. Using the SPSS and setting the OEC, the R TS was generated for Building, Electric, and Mechanic contracts, as shown in Figures 9–11, respectively. In general, the R' TS leads to a slight increase.

For building contracts, the R TS for OEC of 10,000, 1,000,000 and 100,000,000 SAR has a trend to stable, and the R-value was less than one (contract cost is lower than the owner estimate cost). On the other hand, the TS for the OEC of 100,000 and 10,000,000 SAR has notable variance with time, and the R-value exceeds the unit in several years, as shown in Figure 9.

Regarding the electric contracts, all R TS has a value less than the unit except the OEC of 1,000,000 SAR. In addition, the OEC 10,000, 100,000 and 100,000,000 SAR TS slightly trend toward increasing with a value less than a unit. However, the TS with the OEC of 1,000,000 and 10,000,000 SAR exhibit evident variance, as shown in Figure 10.



Figure 9. Generated TS of R using ANN for building contracts



The TS of the mechanic project shown in Figure 11 has a remarkable change in value throughout time except for the TS of the OEC 1,000,000 SAR.

Figure 12 shows a typical observed and forecast TS of the ratio lower bid to the OEC through test data for mechanic contracts with an OEC of 100,000 SAR. The differences between the two TS were too slight, with an RMSSE of 0.065. Therefore, the DNN model provides reasonable accuracy for predicting a TS of a Mechanic contract with an OEC of 100,000 SAR. For evaluating the accuracy of the DNN model for different contract types and sizes through the testing process, Figure 13 shows the RMSSE for building, electric and mechanic, and the OEC ranges from 10,000 SAR to 100,000 SAR. In general, the DNN model provides reasonable accuracy for predicting the R-value, where the maximum value of the RMSSE does not exceed 0.3, representing a small value. The RMSSE increases with increasing the OEC's value for the mechanic contract. In other words, the accuracy of the



**Figure 11.** Generated TS of R using ANN for mechanic contracts



Artificial and deep neural networks

### 93

Figure 12. Typical observed and forecast of R of testing data (mechanic contract with OEC = 100,000 SAR)



Figure 13. RMSSE of observed and forecast of R in the testing process for different projects and OEC

DNN model decreases with increasing the OEC. The RMSSE's performance follows as bell form for building and electric contracts, with their maximum values at the OEC's value of 10,000,000 SAR.

Owner cost estimate (SAR)

Source(s): Authors own work

For building contracts, the forecasting TS of the R can be categorized into three classifications: (1) periodic, (2) semi-periodic (slightly increasing of the R with time) and (3) attenuation series (decreasing of the R with time), as shown in Figure 14. The TS, for OEC equal to 100,000 SAR, has periodic performance with significant changes in the R-value. The TS of the OEC is equal to 10,000 SAR, and 10,000,000 SAR has a semi-periodic performance. The curve periodic time is 13 years and eight years, respectively, varying the R-value range from 0.88 to 1.04 for the OEC of 10,000 SAR and ranges from 0.44 to 1.3 for the OEC of 10,000,000 SAR. On the other hand, the performance TS of the OEC equal to 1,000,000 SAR and 100,000,000 SAR decayed to 0.82 and 0.75, respectively. Therefore, the R-value tends to equal 1 unit for the owner's estimate of 10,000 SAR and 10,000,000 SAR. The study's results are consistent with the Li *et al.* (2022) study on road projects, as the percentage increases with time, with the ratio of the lower bid to the OEC ranging from 0.8 to 1.1.



Regarding the electric contract, the forecasting TS of the R has only a periodic curve with different periods (T). It is defined as the time that it takes for two successive crests. The T value for the five-TS (from the OEC = 10,000 SAR to 100,000,000 SAR) was 2, 13.5, 7, 7 and 15 years, respectively, as shown in Figure 15. The R-value ranges from 0.6 to 1.1. The TS of the mechanic contract is similar to the Electric contracts except for the OEC of 100,000 SAR; it is decayed to the R-value of 0.89. In addition, the TS repeats itself in 15, 2, 16 and 8 years for the OEC of 10,000 SAR, 1,000,000 SAR, 10,000,000 SAR and 100,000,000 SAR, respectively, as shown in Figure 16. The TS of the OEC of 10,000 SAR and 100,000,000 SAR suffer a significant change in the R-value, which varies from 0.5 to 1.1 and from 0.6 to 1.0,



respectively. While the range of the R-value for the OEC of 10,000 SAR and 1,000,000 SAR is narrow, starting from 0.83 to 1.03 and from 0.93 to 1.09.

According to the above information, the TS of the electric and mechanic contract are stationary with time, which repeats itself for several periods. However, the TS of the building contracts sometimes suffers from non-periodic performance, and the R increases with time, especially in a building contract with an OEC of 10,000,000 SAR. The study's results are consistent with the study of (Li *et al.*, 2022) on highway projects, as the percentage increases with time, with the ratio of the lower bid to the OEC ranging from 0.8 to 1.1. No studies



performed TS on the relation between OEC and contract cost to implement more discussion with the study results. For three contracts and the OEC of 100,000,000 SAR, the contract (project) has an R-value of less than 1.0, where the low bid is less than the owner's estimate. As a result, the contract is expensive for both the owner and the contractor. It may result in disagreements, change orders, financial constraints on the contractor and project cost overruns (Jahren and Ashe, 1990; Li *et al.*, 2021).

On the other hand, in the three contracts with the OEC of 10,000 SAR, the R-value is close to 1.0. The contract is close to stable through the construction stage and does not suffer a heavy burden on the owner or fund difficulties for the contractor. Regarding the remaining contracts, the R-value remarkably fluctuates around the 1.0 value. The contracts suffer either

because the contract places a significant responsibility load on the owner to complete the planned projects on time in cases where R is greater than 1.0 or because the contract is cumbersome since it may lead to change orders, disagreements, financial pressure on the contractor and project cost overruns when the R-value is less than 1.0.

5. Conclusion

The paper aims to estimate the ratio of a low bid to an OEC using TS, ANN and DNN to enhance cost estimation processing in the future. Data from ninety-four contracts were collected from KSU in Rivadh, KSA, for three contract types (building, electric and mechanic). After performing the size and normality test, the data were classified into underestimated, optimum and overestimated data depending on the R-value (ratio of the low bid to OEC). The underestimated data were considered to develop the ANN model after using the maximize techniques to overcome the minor data issues. Then, the evaluation of the ANN models was implemented using MAPE, MSSE and RMSSE indicators to check the accuracy of the models. After that, the appropriate ANN model was selected and utilized to generate TS of the R from 2010 to 2021 for the three contract types and different amounts of the OEC. The generated TS was inserted into the DNN, which divided the TS data into training data (65%) and testing data (35%). Finally, the forecasting TS was estimated using the DNN for the three contracts and the different OECs. The finding revealed that the percentage of underestimated, optimum and over-estimated data was 4.2%, 81.1 and 14.7, respectively. The ANN models' MAPE, MSSE and RMSSE were 2.94%, 0.0015, and 0.039, respectively. The DNN's results revealed that the three types of contracts with an OEC of 100,000,000 SAR need more accurate. However, they are close to the optimum for the OEC of 10,000 SAR. This study provides the body of knowledge by developing an ANN and DNN model that enhances the accuracy of the OEC and narrows the discrepancy between the OEC and the lowest submitted offer. The owner and consultant should be able to use the study's findings to create more precise cost estimates and budget plans for better decision-making.

#### References

- Almonacid, F., Pérez-Higueras, P., Rodrigo, P. and Hontoria, L. (2013), "Generation of ambient temperature hourly time series for some Spanish locations by artificial neural networks", *Renewable Energy*, Vol. 51, pp. 285-291, doi: 10.1016/j.renene.2012.09.022.
- Alsugair, A.M. (2022), "Cost deviation model of construction projects in Saudi Arabia using PLS-SEM", Sustainability, Vol. 14 No. 24, 16391, doi: 10.3390/su142416391.
- Amiri, M. and Jensen, R. (2016), "Missing data imputation using fuzzy-rough methods", *Neurocomputing*, Vol. 205, pp. 152-164, doi: 10.1016/j.neucom.2016.04.015.
- Andridge, R.R. and Little, R.J. (2010), "A review of hot deck imputation for survey nonresponse", International Statistical Review, Vol. 78 No. 1, pp. 40-64, doi: 10.1111/j.1751-5823.2010.00103.x.
- Aydilek, I.B. and Arslan, A. (2013), "A hybrid method for imputation of missing values using optimized fuzzy c-means with support vector regression and a genetic algorithm", *Information Sciences*, Vol. 233, pp. 25-35, doi: 10.1016/j.ins.2013.01.021.
- Azadeh, A., Asadzadeh, S.M., Jafari-Marandi, R., Nazari-Shirkouhi, S., Khoshkhou, G.B., Talebi, S. and Naghavi, A. (2013), "Optimum estimation of missing values in randomized complete block design by genetic algorithm", *Knowledge-Based Systems*, Vol. 37, pp. 37-47, doi: 10.1016/j. knosys.2012.06.014.
- Badawy, M. (2020), "A hybrid approach for a cost estimate of residential buildings in Egypt at the early stage", Asian Journal of Civil Engineering, Vol. 21 No. 5, pp. 763-774, doi: 10.1007/s42107-020-00237-z.

Artificial and deep neural networks

ECAM 31,13	Badawy, M., Alqahtani, F. and Hafez, H. (2022), "Identifying the risk factors affecting the overall cost risk in residential projects at the early stage", <i>Ain Shams Engineering Journal</i> , Vol. 13 No. 2, 101586, doi: 10.1016/j.asej.2021.09.013.
	Baek, M., Ph, D., Ashuri, B. and Ph, D. (2019), Assessing Low Bid Deviation from Engineer's Estimate in Highway Construction Projects, Associated Schools of Construction, 2013, Denver, pp. 371-377.
98	Bala, R. and Singh, R.P. (2019), "Financial and non-stationary time series forecasting using LSTM recurrent neural network for short and long horizon", 10th international conference on computing, communication and networking technologies (ICCCNT), IEEE, pp. 1-7.

- Bańbura, M. and Modugno, M. (2014), "Maximum likelihood estimation of factor models on datasets with arbitrary pattern of missing data", *Journal of Applied Econometrics*, Vol. 29 No. 1, pp. 133-160, doi: 10.1002/jae.2306.
- Caltrans, Planning Cost Estimate (2006), available at: https://dot.ca.gov/SearchResults? q=Planning+Cost+Estimate.Pdf (accessed 11 April 2023).
- Camelo, H.D.N., Lucio, P.S., Leal Junior, J.B.V., de Carvalho, P.C.M. and Santosvondos, D.G. (2018), "Innovative hybrid models for forecasting time series applied in wind generation based on the combination of time series models with artificial neural networks", *Energy*, Vol. 151, pp. 347-357, doi: 10.1016/j.energy.2018.03.077.
- Carr, P.G. (2005), "Investigation of bid price competition measured through prebid project estimates, actual bid prices, and number of bidders", *Journal of Construction Engineering and Management*, Vol. 131 No. 11, pp. 1165-1172, doi: 10.1061/(asce)0733-9364(2005)131:11(1165).
- Chapman, R.J. (2001), "The controlling influences on effective risk identification and assessment for construction design management", *International Journal of Project Management*, Vol. 19 No. 3, pp. 147-160, doi: 10.1016/s0263-7863(99)00070-8.
- Chen, K.Y. and Wang, C.H. (2007), "A hybrid SARIMA and support vector machines in forecasting the production values of the machinery industry in Taiwan", *Expert Systems with Applications*, Vol. 32 No. 1, pp. 254-264, doi: 10.1016/j.eswa.2005.11.027.
- Corrêa, J.M., Neto, A.C., Teixeira Júnior, L.A., Franco, E.M.C. and Faria, A.E. (2016), "Time series forecasting with the WARIMAX-GARCH method", *Neurocomputing*, Vol. 216, pp. 805-815, doi: 10.1016/j.neucom.2016.08.046.
- Dhevi, A.S. (2014), "Imputing missing values using Inverse Distance Weighted Interpolation for time series data", 2014 Sixth International Conference on Advanced Computing (ICoAC), IEEE, pp. 255-259.
- Flyvbjerg, B., Bruzelius, N. and Rothengatter, W. (2003), *Megaprojects and Risk: an Anatomy of Ambition*, Cambridge university press.
- Gransberg, N.J. and Gransberg, D.D. (2020), "Public project construction manager-at-risk contracts: lessons learned from a comparison of commercial and infrastructure projects", *Journal of Legal Affairs and Dispute Resolution in Engineering and Construction*, Vol. 12 No. 1, doi: 10.1061/ (asce)la.1943-4170.0000339.
- Hochreiter, S. and Schmidhuber, J. (1997), "Long short term memory", *Neural Computation*, Vol. 9 No. 8, pp. 1735-1780, doi: 10.1162/neco.1997.9.8.1735.
- Hwang, S. (2011), "Time series models for forecasting construction costs using time series indexes", *Journal of Construction Engineering and Management*, Vol. 137 No. 9, pp. 656-662, doi: 10.1061/ (asce)co.1943-7862.0000350.
- Iwana, B.K. and Uchida, S. (2021), "An empirical survey of data augmentation for time series classification with neural networks", *PLoS ONE*, Vol. 16 No. Issue 7 July, e0254841, doi: 10. 1371/journal.pone.0254841.
- Jahren, C.T. and Ashe, A.M. (1990), "Predictors of cost-overrun rates", Journal of Construction Engineering and Management, Vol. 116 No. 3, pp. 548-552, doi: 10.1061/(asce)0733-9364(1990) 116:3(548).

- Junger, W.L. and De Leon, A.P. (2015), "Imputation of missing data in time series for air pollutants", *Atmospheric Environment*, Vol. 102, pp. 96-104, doi: 10.1016/j.atmosenv.2014.11.049.
- Kardakos, E.G., Alexiadis, M.C., Vagropoulos, S.I., Simoglou, C.K., Biskas, P.N. and Bakirtzis, A.G. (2013), "Application of time series and artificial neural network models in short-term forecasting of PV power generation", 48th International Universities' Power Engineering Conference. doi: 10.17632/t2zk3xnt8y.5.
- Khashei, M. and Bijari, M. (2011), "A novel hybridization of artificial neural networks and ARIMA models for time series forecasting", *Applied Soft Computing Journal*, Vol. 11 No. 2, pp. 2664-2675, doi: 10.1016/j.asoc.2010.10.015.
- Khashei, M., Bijari, M. and Raissi Ardali, G.A. (2009), "Improvement of auto-regressive integrated moving average models using fuzzy logic and artificial neural networks (ANNs)", *Neurocomputing*, Vol. 72 Nos 4-6, pp. 956-967, doi: 10.1016/j.neucom.2008.04.017.
- Li, M., Baek, M. and Ashuri, B. (2021), "Forecasting ratio of low bid to owner's estimate for highway construction", *Journal of Construction Engineering and Management*, Vol. 147 No. 1, pp. 1-12, doi: 10.1061/(asce)co.1943-7862.0001970.
- Li, M., Zheng, Q. and Ashuri, B. (2022), "Predicting ratio of low bid to owner's estimate using feedforward neural networks for highway construction", *Construction Research Congress*, Vol. 2022, pp. 340-350.
- Little, R.J. and Rubin, D.B. (2019), Statistical Analysis with Missing Data, Vol. 793, John Wiley & Sons.
- Lobato, F., Sales, C., Araujo, I., Tadaiesky, V., Dias, L., Ramos, L. and Santana, A. (2015), "Multiobjective genetic algorithm for missing data imputation", *Pattern Recognition Letters*, Vol. 68, pp. 126-131, doi: 10.1016/j.patrec.2015.08.023.
- Loy, J. (2019), Neural Network Projects with Python: The Ultimate Guide to Using Python to Explore the True Power of Neural Networks through Six Projects, Packt Publishing.
- Mahamid, I. (2018), "Critical determinants of public construction tendering costs", International Journal of Architecture, Engineering and Construction, Vol. 7 No. 1, doi: 10.7492/ijaec.2018.005.
- MATLAB (2021), "Train Network for Time Series Forecasting Using Deep Network Designer", MATWORKS.
- Maweu, B.M., Shamsuddin, R., Dakshit, S. and Prabhakaran, B. (2021), "Generating healthcare time series data for improving diagnostic accuracy of deep neural networks", *IEEE Transactions on Instrumentation and Measurement*, Vol. 70, pp. 1-15, doi: 10.1109/TIM. 2021.3077049.
- Naim, I., Mahara, T. and Idrisi, A.R. (2018), "Effective short-term forecasting for daily time series with complex seasonal patterns", *Procedia Computer Science*, Vol. 132, pp. 1832-1841, doi: 10.1016/j. procs.2018.05.136.
- Pai, P.F. and Lin, C.S. (2005), "A hybrid ARIMA and support vector machines model in stock price forecasting", Omega, Vol. 33 No. 6, pp. 497-505, doi: 10.1016/j.omega.2004.07.024.
- Pasini, A. (2015), "Artificial neural networks for small dataset analysis", *Journal of Thoracic Deisease*, Vol. 7 No. 5, pp. 953-960, doi: 10.3978/j.issn.2072-1439.2015.04.61.
- Rashid, K.M. and Louis, J. (2019), "Times-series data augmentation and deep learning for construction equipment activity recognition", *Advanced Engineering Informatics*, Vol. 42, 100944, doi: 10. 1016/j.aei.2019.100944.
- Rubio, A., Bermúdez, J.D. and Vercher, E. (2016), "Forecasting portfolio returns using weighted fuzzy time series methods", *International Journal of Approximate Reasoning*, Vol. 75, pp. 1-12, doi: 10. 1016/j.ijar.2016.03.007.
- Saqer, F.(2020), "Development of cost estimation model for ministry of youth and sports affairs construction projects A case study from Kingdom of Bahrain", Submitted by: Master thesis, University of Bahrain.

Artificial and deep neural networks

ECAM 31,13	Shao, C., Fang, F., Bai, F. and Wang, B. (2014), "An interpolation method combining Snurbs with window interpolation adjustment", 2014 4th IEEE International Conference on Information Science and Technology, IEEE, pp. 176-179.
	Citaren D. Deleveri A. Never A. Dec. M. and Association D. (2015). "A superson of similarity of

Sitaram, D., Dalwani, A., Narang, A., Das, M. and Auradkar, P. (2015), "A measure of similarity of time series containing missing data using the mahalanobis distance", 2015 Second International Conference on Advances in Computing and Communication Engineering, IEEE, pp. 622-627.

- Srivastava, N., Hinton, G., Krizhevsky, A. and Salakhutdinov, R. (2014), "Dropout: a simple way to prevent neural networks from overfitting", Journal of Machine Learning Research, Vol. 15 No. 1, pp. 1929-1958.
- Strike, K., El Emam, K. and Madhavii, N. (2001), "Software cost estimation with incomplete data", IEEE Transactions on Software Engineering, Vol. 27 No. 10, pp. 890-908, doi: 10.1109/32. 962560
- Torres, J.F., Hadjout, D., Sebaa, A., Martínez-Álvarez, F. and Troncoso, A. (2021), "Deep learning for time series forecasting: a survey", Big Data, Vol. 9 No. 1, pp. 3-21, doi: 10.1089/big.2020.0159.
- Withington, L., Diaz Pardo de Vera, D., Guest, C., Mancini, C. and Piwek, P. (2021), "Artificial neural networks for classifying the time series sensor data generated by medical detection dogs", Expert Systems with Applications, Vol. 184, 115564, doi: 10.1016/j.eswa.2021.115564.
- WSDOT (Washington State DOT) (2011), "The gray notebook (GNB)", available at: https://wsdot.wa. gov/search/#g=The%20gray%20notebook%20(GNB (accessed 11 April 2023).
- Wu, S.F., Chang, C.Y. and Lee, S.J. (2015), "Time series forecasting with missing values", 2015 1st International Conference on Industrial Networks and Intelligent Systems (INISCom), IEEE, pp. 151-156.
- Zayed, T. (2001), "Assessment of productivity for concrete bored pile construction", PhD Diss, Purdue University.
- Zhang, Y.M., Wang, H., Bai, Y., Mao, J.X. and Xu, Y.C. (2022), "Bayesian dynamic regression for reconstructing missing data in structural health monitoring", Structural Health Monitoring, Vol. 21 No. 5, pp. 2097-2115, doi: 10.1177/14759217211053779.
- Zhao, L., Mbachu, J. and Liu, Z. (2020), "Modelling residential building costs in New Zealand: a timeseries transfer function approach", Mathematical Problems in Engineering, 2020, Vol. 2020, pp. 1-18, doi: 10.1155/2020/7028049.
- Zhao, L., Mbachu, J. and Zhang, H. (2019), "Forecasting residential building costs in New Zealand using a univariate approach", International Journal of Engineering Business Management, Vol. 11, 184797901988006, doi: 10.1177/1847979019880061.

#### About the authors

100

Abdulmohsen Salih Almohsen holds a bachelor's degree in civil engineering from King Saud University, Saudi Arabia. He holds a master's degree from the University of Calgary, Alberta, Canada, in construction logistics management. He obtained his PhD certificate from the University of Calgary, Alberta, Canada, in Strategic Risk Analysis in Joint Ventures. He is an assistant professor of civil engineering at King Saud University. He held several administrative and academic positions at King Saud University. His research interests focus on construction and project management, construction contracts, application of technology in construction, contract and cost management, logistics management, sustainable development, project control and risk management.

Naif M. Alsanabani got his BSc degree in civil engineering from Sana'a University, Republic of Yemen. He received his MSc degree and PhD degree from King Saud University, Riyadh, Saudi Arabia, in geotechnical engineering. His research interests are in soil dynamic, dynamic structure-soil interaction, deep Foundation and operational research.

Abdullah M. Alsugair got his BSc degree in civil engineering from King Saud University (KSU), Saudi Arabia. He received his MSc degree from University of Colorado at Boulder, CO, USA, in construction engineering and management. He received his PhD degree from Texas A&M University, College Station, Texas, USA, in construction engineering and management. He is an associate professor

in civil engineering at KSU. He served at several administrative and academic positions at KSU among	Artificial and
them vice president for projects. His research interests in construction contract administration,	deen neural
construction automation including BIM application, value engineering and sustainability.	uccp neural
Khalid S. Al-Gahtani got his BSc and master's degree in civil engineering from King Saud University	networks
(KSU). He received his PhD degree from the State University of New York at Buffalo in construction	
engineering and management. He is a professor for civil engineering at KSU. He served at several	
administrative and academic positions at KSU. His research interests in analyzing delay claims using	
the CPM schedule, construction automation including BIM application, value engineering,	101
sustainability and risk management. Khalid S. Al-Gahtani is the corresponding author and can be	
contacted at: kgahtani@ksu.edu.sa	

For instructions on how to order reprints of this article, please visit our website: www.emeraldgrouppublishing.com/licensing/reprints.htm Or contact us for further details: permissions@emeraldinsight.com