EPiC Series in Built Environment
Volume 2, 2021, Pages 117–126
ASC 2021. 57th Annual Associated Schools of Construction International Conference

Predictive Capabilities of Supervised Learning Models Compare with Time Series Models in Forecasting Construction Hiring

Boong Yeol Ryoo and Milad Ashtab
Texas A&M University
College Station, TX

Abstract
The construction market is playing a massive role in the United States Gross Domestic Product (GDP). Among various industries, construction is a significant sector responsible for 4-8 percent of GDP. Like other sectors, construction markets are susceptible to demand fluctuations, which the economic recession can cause, political decisions, natural disasters, or outbursts of pandemics. The ability to predict the demand rate in the construction market could give the contractors and owners a better understanding of what they need in their short-term and long-term programs and make them more competitive by predicting the needs in workforce demand. The research selected Texas employment data as the focal point due to the size of the construction market and its workforce diversity.
Furthermore, Texas has been a hotspot for dozens of hurricanes, also affected by many political bills and economic Turmoil, making results more capable of further generalization. This research used three different methods to predict the total construction employment. Univariate models are applied to the datasets to forecast them based on their previous quantities. Three methods such as autoregressive integrated moving average time series models (ARIMA), supervised learning regressors, and the long-term-short-memories (LSTM), were applied to the construction hiring data extracted from the U.S. Bureau of Labor Statistics website. Generally, LSTM models had the most accurate predictions in most cases, except for Austin, where ARIMA models predicted the dataset accurately.

Key Words: Construction Hiring, Deep Learning, Workforce Migration, LSTM, Time-Series Models

Introduction

Analyzing and Predicting the construction market is a crucial competitive asset for contractors to decide on their future policies (Fan et al. 2010, Bureau 2020). The more predictable market helps the public and private sectors to have more realistic short-term and medium-term strategies and market plans (Egan 1998). Many research pieces indicate that economic growth is highly dependent on the construction industry (Giang and Pheng 2011, Chiang et al. 2015). Studies also suggest that the change in construction market output can significantly affect the economy. The volatility of the construction outputs eventuates in varying production levels for contractors. This can result in the contractors' bankruptcy who might struggle to maintain their cash flow (Jiang 2013). In terms of the construction outputs, Turmoil was a recurring event in many East Asian, European, and U.S. markets (Hua 1996). The Texas construction market is no exception in experiencing fluctuations in recent years. According to official reports, the Texas construction volume was nearly 1 billion dollars each month, and the unemployment rate varied between 3.5 to 6.5 in the last five years. Total Hiring in the industry was closing to a record-breaking 800,000 (nearly 10% of the national total 7,639,000) people during economic growth, and right before the COVID-19 pandemic in February 2020 was as low as 570,000 (Bureau, 2020). Moreover, the number of hired people in construction undergone plenty of sudden decreases and increases between 2000 and 2020. It was mostly concurrent with seasonal hurricanes that are used to hit the state during summer. Hurricanes also change general hiring patterns in Texas through labor migration. A combination of skilled and unskilled laborers travel from neighboring states and cities to cities hit by hurricanes to fill the temporary job opportunities in the area, which makes additional uncertainty for predicting employment level.

Hence, the Texas construction market has all the elements to become the pilot study for the relationship between the construction market and the economy. The size of the Texas construction market and fluctuation of the market is essential for a contractor to predict future Hiring and other economic variables to set their short, mid, and long-term plans effectively. Finally, the Texas oil and gas industry and its economic independence make it a completely different state compared to other states (Florida, Alabama, and Louisiana) with disaster-related inconsistencies in their construction market. So, forecasting the future in construction market outputs like construction put in place, Construction Cost Index (CCI), and employment accurately can help contractors put behind fluctuating situations. Thus, econometric models can forecast construction variables. Geographically research focuses on the state of Texas construction market. It will be a pilot for analyzing the whole U.S construction industry. Four significant metropolitan areas (Houston, Dallas, Austin, and San Antonio) were selected as the construction industry's main blocks in Texas. They carry on nearly 90 percent of the hiring load in Texas. They also represent varying geographical areas (economic, geographical, and occurrence of hurricanes).

The research aims to estimate uncertainty for contractors in their short-term hiring plans by developing both statistical and artificial intelligence models. The models' objective is to predict employment in the coming 12 months based on monthly historical data and compare the developed models based on their accuracy. While the models give us a better understanding of construction output quantity in the future, it is necessary to mention that the results are subject to change due to natural disasters or a widespread crisis (Makridakis et al., 2009). Although some models could capture the patterns related to natural disasters like hurricanes, there are always uncertainties like pandemics that would make the models less accurate in real-time forecasting.

Related Researches

Previous related works reviewed in this research engaged with one or both of the following: The first one is the works that primarily focused on drawing a connection between construction output with the economy on a micro and micro scale. The second essential part of the literature focused on methods, models, and algorithms that can forecast time series and apply them in the construction industry. Predicting economic variables in the construction market had been an area of interest among a variety of researchers. According to Oshodi et al. (2020), who performed a systematic Literature review on the subject, the unemployment rate, Construction Price Index (CPI), interest rate, and GDP are indicators of construction output in previous studies. Generating predictive statistical models for construction output on a global scale dates back to the last years of the 1970s. Statistical regressive, autoregressive, and integrated models were among the tools that have been used to forecast construction output. Killingsworth Jr. (1990) implemented a regression model to evaluate industrial construction demand. Hua and Pin (2000) also applied time series models to figure out the future of price and productivity in the Singapore construction market. Shahandashti and Ashuri (2013) used vector error correction (VER) to predict the monthly amount of the CCI published by Engineering News-Record (ENR). Shahandashti and Ashuri (2016) also applied the Granger causality test and VER to predict the monthly National Highway Construction Cost Index (NHCCI). Machine learning, deep learning, and genetic algorithm are also used to predict construction output. A model that combined Neural Networks (N.N.) and Genetic Algorithm (G.A.) was used by Goh (2000) in predictingSingapore's residential market demand level. Lam and Oshodi (2016) also used neural networks to predict the volume of construction work. They compared it to the Auto-Regressive time series model and Support Vector regressors prevalent in the econometric realm. Cao and Ashuri (2020) used the univariate sequence to sequence Long Short-Term Memory Models (LSTM) to predict the Highway Construction Cost Index (HCCI) in Texas. They compared it with the ARIMA model prediction on HCCI. The mentioned studies' scope was limited to the stable economic situation in their area, and the models were not predictive in a more volatile situation.

Hence, some studies focused on measuring the effect of extreme events on construction output. Goh (2005) was among the pioneers who used more dynamic models by intervening datasets through possible economic and disaster hinders that can happen in a period. Jiang and Liu (2011) used vector error correction to form a multivariate model capable of predicting construction demand and draw a positive correlation between construction demand and economic growth. Jiang (2013) also measures the significance of the relationship between construction demand with economic growth, employment growth, and demographic change. Khodahemmati and Shahandashti (2020) also used ARIMA models to predict material costs in post-disaster conditions. Ahmadi and Shahandashti (2020) also used spatial autoregressive models and spatial vector error models for panel data to predict demand surge in post-disaster situations in Texas. The research uses three different sets of tools (Time Series Models, Supervised Learning Regressors, and LSTM to find the best predictive, which can be unique in each of four metropolitan areas or even from time to time. This research also contributes to the current literature by applying statistical and Artificial Intelligence-Based predictive models on a short-term scale rather than trend prediction or long-term forecasts. Accurate short-term predictions would be essential for mitigating early stages impacts of extreme events and, most significantly, hurricanes. The LSTM networks and Supervised learning regressors are expected to capture post-disaster patterns and outperform statistical models in more volatile datasets.

Research methodology

The research has four main parts: Data gathering, developing models, and evaluating their performance. Then, select the best-performing model for each city. Root Mean Square Error (RMSE) and Mean Average Percent Error (MAPE) were selected as the models' performance evaluation criteria. Figure 1 summarizes the overall process of the research.

Data gathering and preprocessing

Available data for construction hiring between 1990 and 2020 is available on the Census Bureau website and provided monthly for metropolitan areas in each state, including Texas (Bureau, 2020). As we discussed in this research, Austin, Dallas, San Antonio, and Houston data were subjected to investigation and analysis. The rationale for focusing on these cities was their share of Texas construction hiring (90%). Each of the cities functions as the economic hub with satellite cities around it. Furthermore, cities represent geographical areas inside Texas with varying neighboring and different exposure to the hurricane. The first step in analyzing the construction hiring time series was to decompose them—trend, seasonality, and residuals of the datasets.

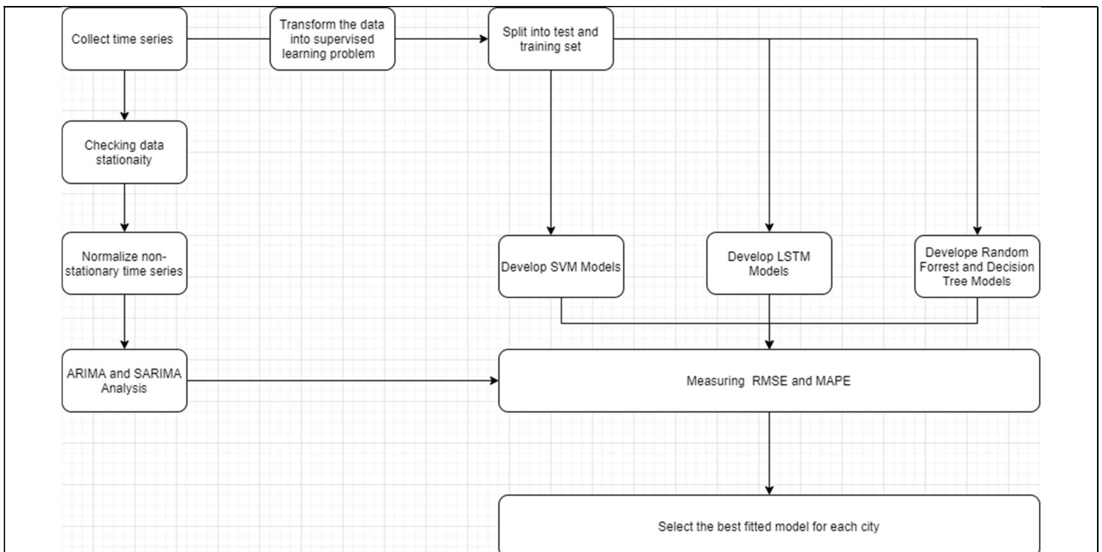

Figure 1 Overall process of the research

To perform further predictive steps, the data needed to be tested from a stationarity point of view. Augmented Dickey-Fuller (ADF) test perform on cost hiring data to determine whether or not the dataset is ready for different time series. For instance, in Houston, the significant level for the test is considered to be 0.01. As a result, the P-value for the decomposed dataset's residuals in the ADF test is 0.693611, which is more than a significant level. So, the null hypothesis, which is the non-stationarity of the raw dataset, cannot be rejected. As the construction hiring data was not stationary logarithm of the dataset was used as the input of time series analysis, the results were extracted independently. P-value is less than 0.01, which rejects the null hypothesis and shows that the changed datasets are stationary. The same data preprocessing method was performed on Dallas, Austin, and San Antonio Data as well. Time series can be decomposed as follows:

$$y_t = S_t + T_t + R_t \qquad (1)$$

Where y_t is the actual value of time-series in time t. S_t is the seasonal component of data and T_t represents the trend for the given timespan and R_t represents the residual of the data. Predicting the hiring data needs models that can capture trends and seasonality, and internal patterns of change in the dataset. Time series predictive models like ARIMA, SARIMA are also applied to predict the hiring data. A grid search approach was selected to select hyperparameters of all the mentioned models. The models with the least Akaike Information Criteria (AIC) were selected as the primary indicator of the hyperparameters' selected set. Supervised learning methods are also applied to predict datasets. Decision Tree, Random Forrest, Support Vector Machines, and Long Short-Term Memory based regressive models fitted by train dataset for Dallas, Houston, Austin, and SanAntonio. Base on the RMSE of test data comparing to predictions, the most effective models selected and applied to predict five years look ahead. Results of look-ahead forecast compared with the current admission level for construction-related programs in Texas.

Transform time-series into a supervised learning problem

The first step to apply supervised learning regression to forecast the time series is to transform it into sets of independent variables defining a dependent variable. In this case, the problem is univariate, and the goal is to predict construction hiring data based on its previous performance. A function is defined to generate 12 timesteps back in time through data (t-1, t-2, …, t-12) as independent variables. The function also generated look ahead timesteps for a given number of months (t, t+1, …, t+11). Each of the look-ahead timesteps considered as the dependent variable in statistical and artificial intelligence models. The models use the last 12-month data to predict the next 12-month point by point. Having 363 data points (each month from January 1990 to March 2020) and putting aside the first 12 months as they did not have enough time lags, a 351 data point was divided into test and training datasets with a 1:2 proportion. (first 218 for test and 133 rest for the train). To make a time-series stationary MinMaxScaler function used to normalize them in range (0,1). In contrast to time series analysis, data stationarity is not essential in the supervised learning models. However, there was a need to make the searching space smaller to make it easier for supervised learning methods to transfer weights or select the best hyperparameter. The Hiring data simplified the numbers between (0,1) using the MinMaxScaler function from the Sklearn package. After applying models to the normalized sets of data predictions, which were between (0,1) transformed into meaning hiring forecasts using the inverted function of MinMaxScaler. Hence, the converted data points were more compatible with the test dataset and generating less and more interpretable RMSE and MAPE values. For each of the supervised learning regressors that have been selected for predicting Hiring data, the Grid Search approach was utilized to select the best hyperparameters for each model based on its performance (RSS, RMSE, and MAPE). In the time t of the time series, if the actual values in the test dataset are y_t and predicted values are y^_t. T is the total number of values in the test data set formula (2) and (3) shows RMSE and MAPE.

$$RMSE = \sqrt{\frac{\sum_{t=1}^{T}(y_t - \hat{y}_t)^2}{T}} \quad (2)$$

$$MAPE = \frac{\sum_{t=1}^{T}\frac{y_t - \hat{y}_t}{y_t}}{T} \quad (3)$$

Results

TimeSeries Models

While generating SARIMA and ARIMA models, it is crucial to check decomposed predicted values. ARIMA and SARIMA are supposed to capture seasonality, trend, and cycles. Ideally, if the models work correctly, the errors(residuals) should be uncorrelated and normally distributed. If our model residuals fail to satisfy normality, it shows the model can be tuned even more accurately.

Figure 4 shows the diagnostic models for the Tuned SARIMA prediction of residuals for Houston while the one with the least RMSE among other models. The figures for all the cities indicate the normality of predicted values residuals. The top right figures for all four cities are showing Kernel Distribution Estimation (KDE) line approximately follows a normal distribution with the mean of zero and standard deviation 1(N(0,1)). The normal Q.Q. Plots show that the residuals of predictions follow a linear pattern of a sample is taken from the N(0,1) distribution. Standardize residuals and correlogram charts also do not show any seasonality and correlation among the residuals. All this evidence indicates residuals' normality, which shows that the tuned SARIMA model cannot improve further.

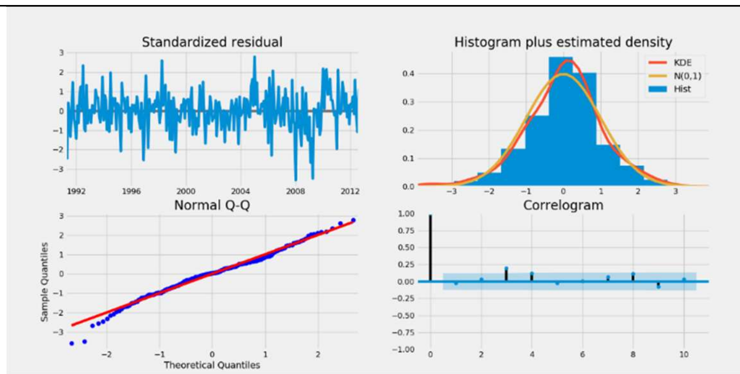

Figure 2 Diagnostic Charts for Sarimax Model

Among the four cities, Austin had the most accurate SARIMA model. The technical reason for that was a consistent ascending trend of the dataset with recognizable seasonalities, making the Austin dataset an ideal sample for time-series models to predict. In contrast, Dallas, Houston, and San Antonio datasets had volatilities beyond seasonality and trends. A significant example of that had shown in Figure 3. The SARIMA model lost track of a significant drop in the Houston dataset in 2017 for a couple of months.Supervised learning regressors applied to see if there is a potential to predict hiring data better.

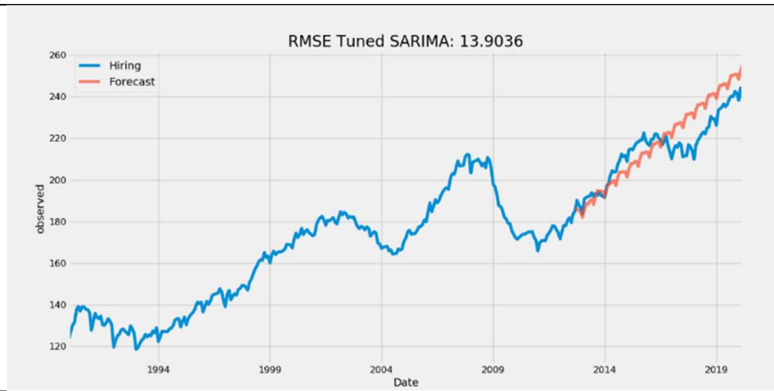

Figure 3 Sarimax Model Performance

Supervised Learning Models

Decision Tree and Random Forrest models predicted earlier parts of the test dataset with utmost accuracy and then ceased to predict the rest of the data. The only exception was San Antonio, where the predictions caught the trends and had the least MAPE and RMSE. The only difference between San Antonio Hiring data and other big cities in Texas is that none of the test dataset points are higher than the maximum or minimum of train data. So, the test data for San Antonio is less volatile, and as a result, the models can capture patterns and predict more effectively. SVM regressor had a better performance in finding the general direction of hiring and construction market demand. San Antonio and Houston had the least RMSE, and MAPE and the model were able the capture significant changes in the trend smoothly in both cities. For Dallas and Austin, as there is constant growth, the model captures the trend but cannot adjust to the market's rapid growth, and error level increases in more recent dates. LSTM had the best performance among all the selected methods. Different numbers of layers of LSTM were examined on the datasets alongside the grid search of hyperparameters to make the predictions more holistic. The table shows the best performance one, two, and three layers of LSTM through grid-search. The one-layer architecture proved to be the most accurate, and it had the least RMSE and MAPE. In the selected one-layer architecture, the Loss function for test and train data is similar It shows a minimal loss for both in all four cities, which indicates that the model fitted adequately. No overfitting or underfitting is associated with datasets Table 1, Three, Six, and twelve-month look represent ahead predictions based on the last 12-month hiring data for Random Forrest, SVM, and LSTM Models and tuned SARIMA.

Table 1 Accuracy of machine learning regressors in comparison to SARIMA performance

	CITY	Houston		Austin		Dallas		San Antonio	
Method	# month	RMSE	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE	MAPE
SARIMA	average	9.40	N/A	**0.99**	N/A	5.88	N/A	3.23	N/A
RF	1	12.848	4	10.325	4	9.519	10	**1.25**	1
RF	3	16.365	5	11.017	12	13.734	6	**2.296**	2
RF	6	21.893	7	13.059	15	16.878	8	**3.141**	4
RF	12	26.332	8	14.560	17	22.010	10	3.885	5
SVM	1	**6.92**	2	7.591	5	6.150	9	**1.938**	3
SVM	3	**7.348**	3	6.072	9	7.350	5	**2.745**	4
SVM	6	10.618	4	6.027	9	8.376	5	**3.014**	5
SVM	12	13.223	5	6.502	9	10.375	7	3.350	5
LSTM	1	**4.757**	1	2.803	4	**3.705**	2	**1.310**	2
LSTM	3	**6.828**	2	3.129	4	**5.261**	3	**1.783**	2
LSTM	6	11.772	4	4.719	7	7.710	4	**2.262**	3
LSTM	12	20.601	7	8.004	11	13.975	8	**3.313**	5

Conclusion

A group of statistical time series analysis methods and supervised learning algorithms were applied to investigate if the trend in Hiring's future market demand can be predicted. Based on each city's characteristics, different models proved to be sufficient in predicting upcoming values. The results for tuned SARIMA in Austin even outperformed LSTM models. The high value of seasonality in Austin data and steady growth in recent years helped tuned SARIMA to capture and predict future values correctly. Decision Tree and Random Forrest Regressors' prediction was inconsistent for predicting the hiring dataset. They ceased to predict from the beginning of 2017 when hiring data as construction hiring reach its all-time high suggests. The issue is inherent to these three algorithms' technical capabilities as they struggle to predict datasets that are reaching their extreme maximum or minimum. Hence, They are not appropriate models for forecasting the Texas construction market's future. Except for Austin, SVM and LSTM had the most accurate hiring values predictions in the coming month. The input data had 12 dimensions, and SVM's ability to work with high-dimensional data was the main reason for its decent performance. On the other hand, the large datasets are among the main weaknesses of SVM, which did not interrupt the algorithm's performance. LSTM's capability to capture inner states and data sequence patterns was the leading cause for its accurate performance in predicting the dataset. As mentioned in the methodology, LSTM has various possible input and output alternatives. However, as this research sought predictive ability in specific value in the short term, the LSTM architecture was designed to get a data sequence and predict a point. The analysis showed that the Texas hiring data is predictable based on its historical records, specifically with the LSTM that is a deep learning algorithm for short-term forecasting.

Future works to complete the current research could focus on other economic factors to predict construction hiring and then calibrate deep learning models with multivariate inputs. Also, the sequence to sequence models can be used to predict long-term uncertainties in the construction market. Finally, intervention time series models can be used to measure the effect of extreme events in construction hiring. Then, the results can be applied to calibrate current predictive models. A potential application of such models in the industry can the contractor's hiring plans more accurate.

References

AHMADI, N. & SHAHANDASHTI, M., 2020. Characterizing Construction Demand Surge Using Spatial Panel Data Models. Natural Hazards Review, 21, 04020008.

BUREAU, U. S. C., 2020. Construction Spending, U.S. Census Bureau.

CAO, Y. & ASHURI, B., 2020. Predicting the Volatility of Highway Construction Cost Index Using Long Short-Term Memory. Journal of Management in Engineering, 36, 04020020.

CHIANG, Y.-H., TAO, L. & WONG, F. K., 2015. Cau sal relationship between construction activities, employment and GDP: The case of Hong Kong. Habitat international, 46, 1-12.

EGAN, J., 1998. Rethinking construction, Department of Environment, Transport and the Region.

FAN, R. Y., N.G., S. T. & WONG, J. M. 2010. Reliability of the Box–Jenkins model for forecasting construction demand covering times of economic austerity. Construction Management and Economics, 28, 241-254.

GIANG, D. T. & PHENG, L. S., 2011. Role of construction in economic development: Review of key concepts in the past 40 years. Habitat international, 35, 118-125.

GOH, B.-H. 2000. Evaluating the performance of combining neural networks and genetic algorithms to forecast construction demand: the case of the Singapore residential sector. Construction Management & Economics, 18, 209-217.

GOH, B. H., 2005. The dynamic effects of the Asian financial crisis on construction demand and tender price levels in Singapore. Building and Environment, 40, 267-276.

HOCHREITER, S. & SCHMIDHUBER, J. LSTM can solve hard long time lag problems. Advances in neural information processing systems, 1997. 473-479.

HUA, G. B., 1996. Residential construction demand forecasting using economic indicators: a comparative study of artificial neural networks and multiple regression. Construction Management and Economics, 14, 25-34.

HUA, G. B. & PIN, T. H., 2000. Forecasting construction industry demand, price and productivity in Singapore: the BoxJenkins approach. Construction Management and Economics, 18, 607-618.

JIANG, H., 2013. Econometric techniques for estimating construction demand in Australia. Deakin University.

JIANG, H. & LIU, C., 2011. Forecasting construction demand: a vector error correction model with dummy variables. Construction Management and Economics, 29, 969-979.

KHODAHEMMATI, N. & SHAHANDASHTI, M., 2020. Diagnosis and Quantification of Postdisaster Construction Material Cost Fluctuations. Natural Hazards Review, 21, 04020019.

KILLINGSWORTH JR, R. A., 1990. A preliminary investigation into formulating a demand forecasting model for industrial construction. Cost Engineering, 32, 11.

LAM, K. C. & OSHODI, O. S., 2016. Using univariate models for construction output forecasting: Comparing artificial intelligence and econometric techniques. Journal of Management in Engineering, 32, 04016021.

MAKRIDAKIS, S., HOGARTH, R. M. & GABA, A., 2009. Forecasting and uncertainty in the economic and business world. International Journal of Forecasting, 25, 794-812.

OSHODI, O., EDWARDS, D. J., OLANIPEKUN, A. O. & AIGBAVBOA, C. O., 2020. Construction output modeling: a systematic review. Engineering, Construction and Architectural Management.

SHAHANDASHTI, S. & ASHURI, B., 2013. Forecasting engineering news-record construction cost index using multivariate time series models. Journal of Construction Engineering and Management, 139, 1237-1243.

SHAHANDASHTI, S. & ASHURI, B., 2016. Highway construction cost forecasting using vector error correction models. Journal of management in engineering, 32, 04015040.