Modelling the Price Forecast for Construction Steel: A Case Study in EPC Company

[1]Alfa Firdaus*, [1]Uly Amrina
[1]Universitas Mercu Buana

Abstract— The EPC (Engineering Procurement Construction) industry is one of the most dynamic industries. The problems faced are related to market conditions that often change, short construction periods, and fluctuations in material prices that are difficult to predict. This dynamic requires an appropriate forecasting model, which can predict the pattern of material price movements and anticipate the occurrence of fluctuations in the future. This research aims to get the best price during the project tender process. This study model the forecasting of construction iron prices in the future by considering the historical pattern of construction iron price data, the value of foreign exchange rates, and the price of billets as raw materials for construction iron. The forecasting procedure used is nonparametric, which involves several statistical tests such as cross-correlation, linearity, and error assessment. The results of this study can be a firm reference for the price value of construction iron, which makes it easier for management to determine an accurate and competitive project value.

Index Terms—construction steep, EPC company, forecasting, price.

I. INTRODUCTION

COVID-19 has had a significant impact on the world economy, especially in the first six months of 2020. This impact was also felt in developing the EPC (Engineering Procurement Construction) industry. Changes in economic conditions affect the management of EPC projects, especially in terms of the dynamics of market demand and fluctuations in material prices. Several property developers have reviewed price adjustments in line with rising iron and steel prices at retail. Previously, local iron and steel producers admitted an increase in selling prices at the consumer level from 20% to 50% due to rising raw material costs in the world. In addition, demand for iron and steel also increased amid limited supply. The local iron and steel industry projects that the selling price will continue to climb following the price of raw materials.

The increase in iron and steel prices will affect property products' selling prices and affect many other iron and steel elements [1]. Property projects are affected, as also projects in the construction sector. Therefore, some EPC companies are reviewing accurate price predictions for their products. One of the strategies is to choose the proper forecasting method to predict

* Corresponding Authour:alfafirdaul@gmail.com

the pattern of material price movements and anticipate the occurrence of fluctuations in the future.

This research was developed to assist EPC companies in modelling construction iron price forecasts in the future by considering the historical patterns of construction iron price data, foreign exchange rates, and billet prices as construction iron raw materials. This forecasting model will be helpful to get the best competitive price during the project tender process [2]. The forecasting procedure used is nonparametric, which involves several statistical tests such as cross-correlation, linearity, and error assessment [3]. Further discussion will describe the research methods and results.

II. RESEARCH METHOD

This study uses a nonparametric approach with the procedure as visualized in Figure 1.

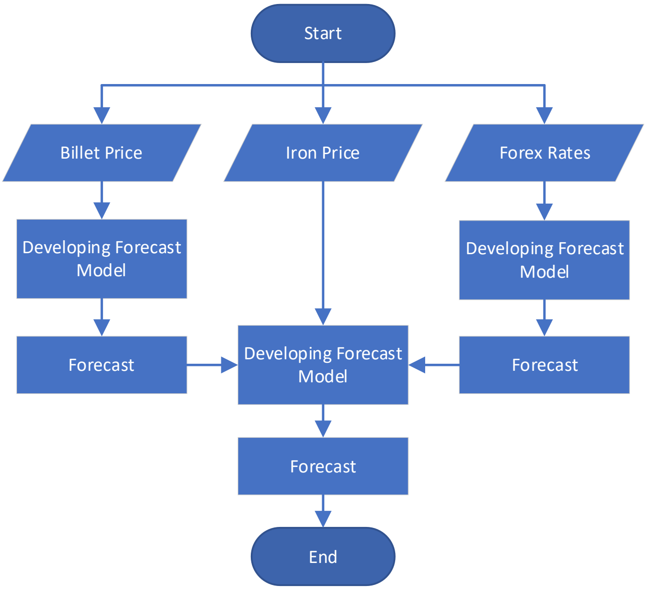


Figure 1. Research steps of nonparametric forecasting

Based on the flow chart in Figure 1 above, the forecasting procedure used goes through several stages as follows:

1. Build a Billet Price forecasting model and make a forecast for the next month.
2. Build a Foreign Exchange Rate or Central Bank (Bank Indonesia / BI) Mid Rate forecasting model and make a forecast for the next month.
3. Build a forecasting model for Iron Price by involving the influence of Billet Price and BI Mid Rate. Then, forecast Iron Prices for the next month based on the results in step 3 using the input data for Billet Prices and BI Mid Rates obtained in steps 1 and 2.

Time series models create patterns in the past movements of variables and use that information to predict future movements [4], [5]. Often we are faced with time-series data patterns that are difficult to specify parametrically. Therefore, we need a flexible technique (flexible modelling) to have a forecasting model that follows the existing data pattern [6]. Several techniques are flexible in application. The nonparametric Fourier is the most practical technique for time series data [7], [8].

The Fourier series is a series of simple periodic functions built by the trigonometric sine-cosine function or the exponential function x [9]. The level of flexibility of this Fourier approach makes modelling or forecasting using this technique very suitable when dealing with data patterns that are difficult to specify parametrically [10], [11].
Suppose there is a flexible time series data model as follows:

$$y(t) = f(t) + \boldsymbol{X\beta} + \varepsilon \quad \text{.................................} \quad (1)$$

with:

t	:	Time variable
y(t)	:	Response variable at time t
X	:	The matrix containing input variable information (covariate at time t)
$\boldsymbol{\beta}$	:	Coefficient of variable *X*
f(t)	:	Smooth function of time variable
$\boldsymbol{\varepsilon}$	:	The error of the model

In general, the time series data modeling approach using the Fourier transform based on the above model has the following form:

$$y(z) \approx \alpha_0 + \alpha_1 z + \alpha_1 z^2 + \sum_{q=1}^{Q} \left(\lambda_q \sin(qz) + \delta_q \cos(qz)\right) + \boldsymbol{X\beta} + \varepsilon \quad \text{...............} \quad (2)$$

It should be noted that in the above model, the notation z is used, so before the formation of the Fourier Series model, the t variable is first converted to the z variable as follows:

$$z = \frac{2\pi(t - \min(t))}{(\max(t) - \min(t))} \quad \text{...............} \quad (3)$$

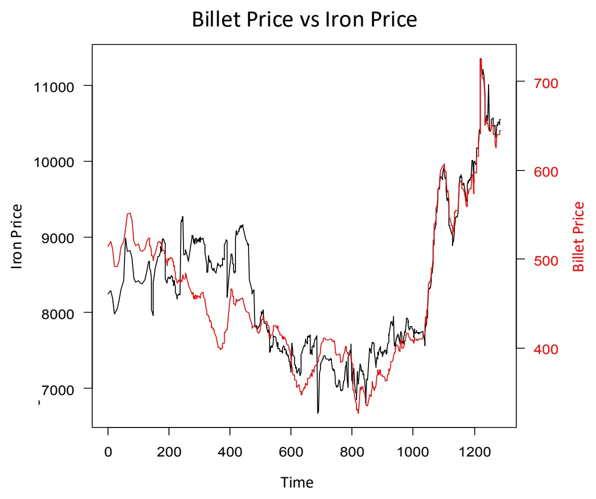
If we are dealing with a single time series modelling, not involving other variables as input variables (not involving covariates), then the above model becomes simpler as follows:

$$y(z) \approx \alpha_0 + \alpha_1 z + \alpha_1 z^2 + \sum_{q=1}^{Q} \left(\lambda_q \sin(qz) + \delta_q \cos(qz)\right) + \varepsilon$$

III. RESULT AND DISCUSSION

A. Identification the Model

Figure 2 shows a similar pattern for Billet Prices with Prices for Iron, which shows a reasonably high correlation between the two variables. A slightly different pattern is found in the BI Mid Rate with Iron Prices. There is also an interesting phenomenon. There is a high spike in the BI Mid Rate, which shows that there are interventions or outliers from the data.



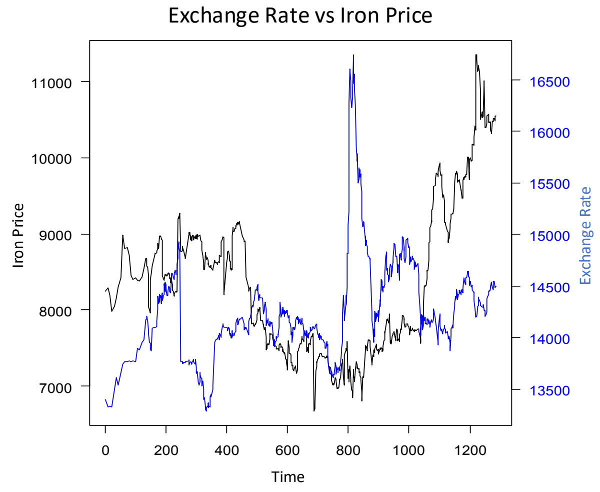


Figure 2. Data pattern - (Above) Billet Price (red) with Iron Price (black) and (Bottom) BI Mid Rate (blue) with Iron Price (black)

The next stage is testing whether there is a correlation between predictors. This test aims to show whether there is a potential for multicollinearity. The test results show that the general pattern is different, so the correlation between the variables above is thought to be not too large. This result makes it possible to use the two variables above as predictors in the forecasting model for Iron Prices.
Meanwhile, the results of the cross-correlation test have a significant value (above 0.8) and are relatively the same (the largest is at lag 0). These results indicate no time difference (lag) to influence the model of the Billet Price on the Price of Iron.
The last test is the linearity test. Teraesvirta test results showed a p-value <5%. It means that with an error risk of 5% (even for 1%), it appears that H0 is rejected, which means that all patterns are nonlinear.

B. Develop A Forecasting Model

B.1. Forecasting of Billet Price

The first step to making Fourier modelling from Billet Price is determining the optimal Q value using Generalized Cross-Validation (GCV) criteria. The calculation results show a minimum GCV value of 159.8656 with a Q of 12. Thus, the Fourier Series Nonparametric

Regression model that is formed to predict Billet Prices based on the period with a bandwidth value of Q=12 is as follows:

$$y(z) \approx 635,5991 - 205,7458\, z + 35,7736 z^2 + \sum_{q=1}^{12} \left(\lambda_q \sin(qz) + \delta_q \cos(qz) \right) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (4)$$

with the respective values of q and q for q = 1,2,…, 12 are as shown in Table 1.

Table 1. The coefficient value of $\boldsymbol{\lambda_q}$ and $\boldsymbol{\delta_q}$ for *q=1,2, …,12* from the Fourier Model of Billet Price

λ			δ	
λ_1	49,31687		δ_1	-45,8951
λ_2	-12,8417		δ_2	-2,76652
λ_3	-7,58535		δ_3	-15,6492
λ_4	2,768302		δ_4	-22,1209
λ_5	-5,72452		δ_5	0,201446
λ_6	5,902521		δ_6	-5,54705
λ_7	-1,65807		δ_7	6,902292
λ_8	-14,2557		δ_8	-6,24427
λ_9	-7,55838		δ_9	-2,8493
λ_{10}	1,368928		δ_{10}	-14,1607
λ_{11}	-0,78645		δ_{11}	-8,92979
λ_{12}	2,426225		δ_{12}	-0,66811

The model in equation (4) has an Adjusted R-squared value of 0.9792. It indicates that 97.92% of the total variance of the Billet Price variable can be explained by the model formed. The AIC obtained is 5.07 with mean squared error, MSE = 148.6651 and Root Mean Squared Error = 12.19283.

Based on the estimation of the model above, the next step is to perform the process of forecasting the value of the Billet Price. The time information must first be transformed according to the z equation above by using the minimum (t =1 and maximum (t)=1285 so that the following transformation equation is obtained:

$$t^* = \frac{2\pi(t-2)}{(1285-1)} = \frac{2\pi(t-2)}{1284} \quad \ldots\ldots\ldots\ldots\ldots\ldots.. \quad (5)$$

The next stage is forecasting. The results of the Billet Price forecast for the next 24 days are shown in Figure 3.

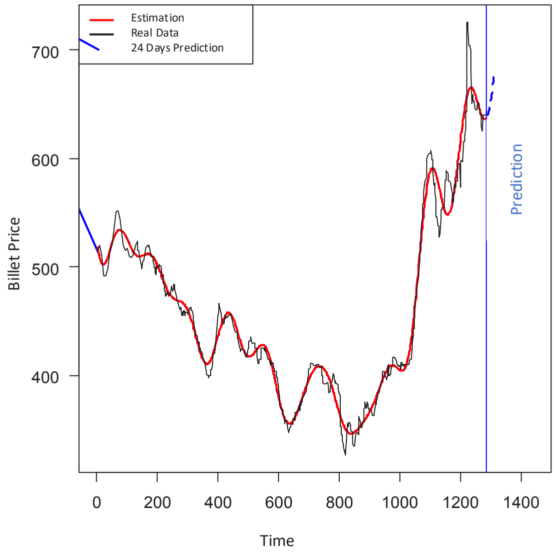


Figure 3. Estimated plot along with actual data of Billet Price for the next 24 days

B.2. Mid BI Rate Forecast

The calculation results of the minimum GCV value of 70898.21 with a Q of 8. Thus, the Fourier Series Nonparametric Regression model that is formed to predict the Mid BI Rate based on the period with a bandwidth value of Q = 12 is as follows:

$$y(z) \approx 11221{,}866 + 2357{,}918\, z - 334{,}994\, z^2 + \sum_{q=1}^{8}\left(\lambda_q \sin(qz) + \delta_q \cos(qz)\right) \quad \text{.....................} (6)$$

with the respective values of q and q for q = 1,2,…, 12 are as shown in Table 2.

Table 2. The coefficient value of $\boldsymbol{\lambda_q}$ and $\boldsymbol{\delta_q}$ for *q=1, .. ,8* from the Fourier model of Mid BI Rate

$\boldsymbol{\lambda}$	
$\boldsymbol{\lambda_1}$	188,6633
$\boldsymbol{\lambda_2}$	315,4132
$\boldsymbol{\lambda_3}$	252,0437
$\boldsymbol{\lambda_4}$	-113,086
$\boldsymbol{\lambda_5}$	6,672905
$\boldsymbol{\lambda_6}$	-96,6108
$\boldsymbol{\lambda_7}$	130,6118
$\boldsymbol{\lambda_8}$	224,6594

$\boldsymbol{\delta}$	
$\boldsymbol{\delta_1}$	1250,525
$\boldsymbol{\delta_2}$	198,3368
$\boldsymbol{\delta_3}$	240,0651
$\boldsymbol{\delta_4}$	-81,7077
$\boldsymbol{\delta_5}$	82,04163
$\boldsymbol{\delta_6}$	160,9205
$\boldsymbol{\delta_7}$	-165,893
$\boldsymbol{\delta_8}$	82,08682

Based on this model, it has an Adjusted R-squared of 0.7253, which indicates that 72.53% of the total variance of the Mid BI rate variable can be explained by the model formed. The AIC obtained is 11.16836 with mean squared error, MSE = 67383.91 and Root Mean Squared Error = 259,5841. The next stage is forecasting the rate. The Mid BI Rate forecast results for the next 24 days are shown in Figure 4.

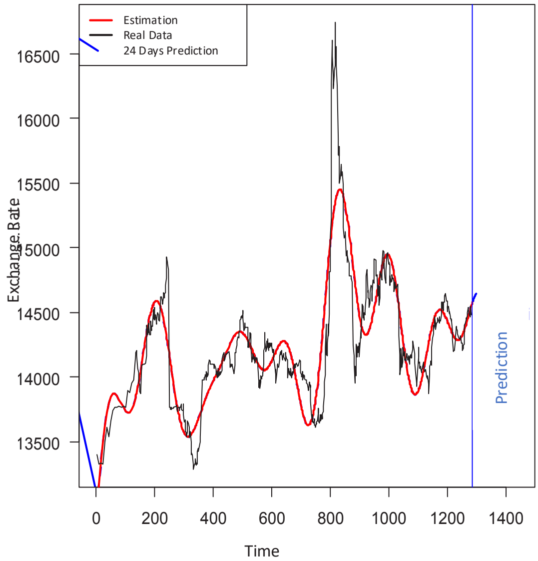


Figure 4. Estimated plot along with actual data of the Mid BI Rate for the next 24 days

B.3. Iron Price Forecast

If forecasting was carried out for each predictor variable (covariate) in the previous two stages, then in this section, forecasting will be carried out for the response variable using forecasting information on the two predictor variables. In this section, a modelling of the Iron Price will be made by including information on the Billet Price and BI Mid Exchange rate as input variables. As in the data exploration section, the presence of lag has little effect on changes in the magnitude of the cross-correlation; even for the Billet Price variable, the cross-correlation for zero lag has the most significant value. Therefore, the effect that in this modelling, the influence of the two variables will only be involved at zero lag. In other words, we will model the Iron Price using information from the Billet Price and BI Mid Rate at the same time, with the same model as follows:

$$y(z) \approx \alpha_0 + \alpha_1 z + \alpha_1 z^2 + \beta_1 X_1(z) + \beta_2 X_2(z) + \sum_{q=1}^{Q} \left(\lambda_q \sin(qz) + \delta_q \cos(qz)\right) + \varepsilon$$

………………………………………………………….. (7)

Where X1(z) is the input variable for the Billet Price at time t (transformed into z), and X2(z) is the input for the BI Mid Rate Variable at time t.

Based on the calculation, the minimum GCV value is 27148.68 with a Q of 8. Thus, the Fourier Series Nonparametric Regression model formed with a bandwidth value of Q=8 is as follows:

$$y(z) \approx 1332{,}88378 + 413{,}32942\, z - 47{,}41360\, z^2 + 10{,}40270\, X_1(z) + 0{,}10917\, X_2(z) + \sum_{q=1}^{8} \left(\lambda_q \sin(qz) + \delta_q \cos(qz)\right)$$

………………………………………………………….. (8)

With the value of λ_q and δ_q for q = 1,2,…, 12 are as shown in Table 3.

Table 3. The coefficient value of $\boldsymbol{\lambda_q}$ and $\boldsymbol{\delta_q}$ for *q=1, .. ,8* from the Fourier model of Iron Price

$\boldsymbol{\lambda}$			$\boldsymbol{\delta}$	
$\boldsymbol{\lambda_1}$	423,2411		$\boldsymbol{\delta_1}$	157,2683
$\boldsymbol{\lambda_2}$	-79,8651		$\boldsymbol{\delta_2}$	-211,502

λ_3	-124,826
λ_4	-1,71669
λ_5	12,29695
λ_6	27,98939
λ_7	-36,3833
λ_8	-40,9442

δ_3	132,2417
δ_4	154,9474
δ_5	-138,337
δ_6	89,92273
δ_7	-80,5952
δ_8	38,69738

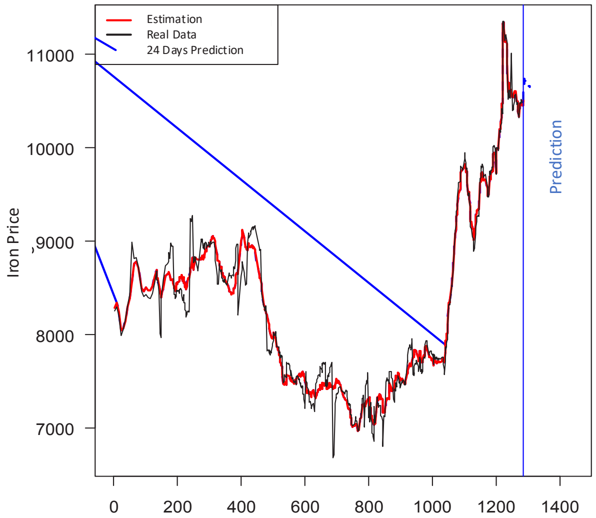


Figure 5. Estimated plot along with actual data of Iron Price for the next 24 days

This model has an Adjusted R-squared of 0.974, which indicates that 97.40% of the total variance of the Billet Price variable can be explained by the model formed. The AIC obtained is 10.2083 with a mean squared error, MSE = 25663.31 and a Root Mean Squared Error = 160.1977. The next stage is forecasting. In the last stage, the testing data is checked. In this case, it is compared with the actual data on the price of iron with the forecasted data. Based on the updated data, the MAPE value is 0.48%. Iron Price forecast results for the next 24 days are shown in Figure 5.

C. Discussion

The development of prices for iron ore, scrap, and international billets as raw materials for the manufacture of iron and steel materials in Indonesia is increasingly showing significant changes every month. These changes cause the prices for iron and steel materials in Indonesia to change, either increasing or decreasing the prices of iron and steel materials themselves. The iron and steel materials produced in Indonesia still require imported raw materials in iron ore, scrap and billets. This condition makes the prices for iron and steel materials in Indonesia strongly influenced by the prices of raw materials originating in the supplying country. In importing countries, prices of raw materials are influenced by factors such as natural factors, government policies, natural disasters, and economic influences from other countries.

Iron ore is one of the most volatile commodities as prices rise and fall. In just one month, iron ore, which fuels China's massive steel industry, surged to a record high. Structural problems affecting China's iron ore supply remain prominent. Historically low inventories, insufficient shipments in recent weeks from major suppliers in Australia and Brazil, and possible domestic production curbs following the recent mining accident. At least in the short term, China appears to have decided that falling raw material prices, including iron and steel, are its most pressing concern. To that end, Beijing has stepped up its efforts to curb commodity

inflation in recent weeks, focusing on preventing wild speculation, at least publicly, as pushing prices higher.
One of the ironies of the policy is the critical role that speculators play in reducing volatility by increasing liquidity in the futures market. Simply put, more buyers and sellers means price adjustments are less sudden. While the policy outlook for the iron and steel market remains bleak, summer was a slow season for construction and manufacturing in China, which should weigh on iron prices. In the longer term, China's efforts to suppress prices will find support from a cyclical turnaround in the economy as measures to support the economy is suspended. Steel demand may moderate as China exits large-scale stimulus and seeks growth driven by domestic consumption.
In Indonesia itself, to bring in raw materials to be processed into iron and steel materials, one must also look at the current exchange rate of the rupiah against the US dollar. Because this significantly affects the prices for iron and steel materials in Indonesia itself. If the condition of the rupiah exchange rate against the US dollar is experiencing a decline, the impact on the price of raw materials to be purchased will be even higher. In addition to purchasing raw materials for the manufacture of iron and steel materials, the rupiah's exchange rate against the US dollar also affects the purchase of natural gas as an energy source for the manufacture of iron and steel materials in Indonesia.
The company has to go through the production stage to get iron and steel material products that can be used in the market. Changes in the price of iron and steel materials in Indonesia are also driven by the production cost factor, which in the production process still has to look at the adjustment of the applicable basic electricity tariff (TDL). Electricity is needed to support the production process of making iron and steel materials in Indonesia as a support tool for forming iron and steel materials. The production costs for iron and steel materials in Indonesia cost much money. Therefore, with various factors influencing changes in iron and steel prices in Indonesia, the prices of iron and steel materials in Indonesia also experience development.
In this study, optimization is carried out based on the price formula designed for iron concrete. The formula is a regression model consisting of the dependent variable (Y), the price index of concrete iron, and the independent variable (X), the BI mid-exchange rate index, multiplied by the billet price index. The price formula calculation is done by first determining the baseline. The billet price forecast and the mid BI exchange rate are calculated using the Fourier Series Nonparametric Regression model as a reference for the model. The Mean Absolute Percentage Error (MAPE) is used to measure forecasting performance. The result of the MAPE is 0.48%.

IV. Conclusion

This study proves that statistically, the developed model is suitable for describing the pattern of fluctuations in the price of iron and concrete. The Fourier Series Nonparametric Regression Model, as one of the methods of procurement analytics, is a powerful cost estimation tool as a negotiating capital for the SCM division and helps achieve negotiation results that are close to the target price. The formula can be adjusted to increase the potential efficiency against the baseline price and can be designed so that the price is always below the market price. In a broader scope, procurement analytics impacts competitive advantage, tender winning strategies and economies of scale to increase the bargaining power of EPC companies because steel price parameter data can be reviewed in real-time and are continuously updated.

V. REFERENCES

[1] S. H. Sa'diyah, "Peramalan Harga Besi Beton Menggunakan Fungsi Transfer, Autoregressive Distributed Lag (Ardl), Dan Support Vector Regression (Svr)," Institut Teknologi Sepuluh Nopember, 2018.

[2] A. M. Husein, F. R. Lubis, and M. K. Harahap, "Analisis Prediktif untuk Keputusan Bisnis: Peramalan Penjualan," *Data Sciences Indonesia (DSI),* vol. 1, no. 1, pp. 32-40, 2021.

[3] J. Frois Caldeira, R. Gupta, M. T. Suleman, and H. S. Torrent, "Forecasting the term structure of interest rates of the BRICS: Evidence from a nonparametric functional data analysis," *Emerging Markets Finance and Trade,* pp. 1-18, 2020.

[4] S. Siami-Namini, N. Tavakoli, and A. S. Namin, "A comparison of ARIMA and LSTM in forecasting time series," in *2018 17th IEEE international conference on machine learning and applications (ICMLA)*, 2018: IEEE, pp. 1394-1401.

[5] R. Patria and S. Sudarto, "Integrasi forecasting pada rantai pasok manufaktur komponen otomotif Jepang di Indonesia dengan penerapan metode classic dan regresi," *Operations Excellence,* vol. 12, no. 3, pp. 386-397, 2020.

[6] P. Sun, I. Kim, and K. Lee, "Flexible weighted dirichlet process mixture modelling and evaluation to address the problem of forecasting return distribution," *Journal of Nonparametric Statistics,* vol. 32, no. 4, pp. 989-1014, 2020.

[7] A. Tompkins and F. Ramos, "Fourier feature approximations for periodic kernels in time-series modelling," in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2018, vol. 32, no. 1.

[8] P. J. Pritz, D. Perez, and K. K. Leung, "Fast-fourier-forecasting resource utilization in distributed systems," in *2020 29th International Conference on Computer Communications and Networks (ICCCN)*, 2020: IEEE, pp. 1-9.

[9] P. Sun, N. AlJeri, and A. Boukerche, "A fast vehicular traffic flow prediction scheme based on fourier and wavelet analysis," in *2018 IEEE Global Communications Conference (GLOBECOM)*, 2018: IEEE, pp. 1-6.

[10] E. Yukseltan, A. Yucekaya, and A. H. Bilge, "Hourly electricity demand forecasting using Fourier analysis with feedback," *Energy Strategy Reviews,* vol. 31, p. 100524, 2020.

[11] M. Mardianto, S. H. Kartiko, and H. Utami, "Forecasting Trend-Seasonal Data Using Nonparametric Regression with Kernel and Fourier Series Approach," in *Proceedings of the Third International Conference on Computing, Mathematics and Statistics (iCMS2017)*, 2019: Springer, pp. 343-349.