Comparison of Four Time Series Forecasting Methods for Coal Material Supplies: Case Study of a Power Plant in Indonesia

Muhammad Rizki
Industrial Engineering Department
Faculty of Science and Technology,
UIN Sultan Syarif Kasim Riau
Pekanbaru, Indonesia
Muhammad.rizki@uin-suska.ac.id

Alex Wenda
Electrical Engineering Department
Faculty of Science and Technology,
UIN Sultan Syarif Kasim Riau
Pekanbaru, Indonesia
Alexwenda@uin-suska.ac.id

Farhan Dio Pahlevi
Industrial Engineering Department
Faculty of Science and Technology,
UIN Sultan Syarif Kasim Riau
Pekanbaru, Indonesia
11850210432@students.uin-suska.ac.id

Muhammad Isnaini Hadiyul Umam
Industrial Engineering Department
Faculty of Science and Technology,
UIN Sultan Syarif Kasim Riau
Pekanbaru, Indonesia
muhammad.isnaini@uin-suska.ac.id

Muhammad Luthfi Hamzah
Information System Department
Faculty of Science and Technology,
UIN Sultan Syarif Kasim Riau
Pekanbaru, Indonesia
muhammad.luthfi@uin-suska.ac.id

Sutoyo
Electrical Engineering Department
Faculty of Science and Technology,
UIN Sultan Syarif Kasim Riau
Pekanbaru, Indonesia
sutoyo@uin-suska.ac.id

***Abstract*— Coal is the main fuel in the production process at PT PJB UBJ O&M Tenayan. As a raw material, coal needs to be considered in terms of supply to prevent losses (depreciation in caloric content) in case of oversupply. This study aimed to compare four forecasting methods for coal material supply. The four methods of time series forecasting are the moving average method, the weighted moving average, the single exponential smoothing, and the linear regression. Forecasting error calculations used the smallest MAD, MSE, and MAPE error parameters, whereas the tracking signal was used to monitor the forecasting results. The data required were coal supply and demand. Based on the data processing obtained, results of this study show that the best method is linear regression with the results of the MAD value of 13,285.63, MSE of 228,778,800, and MAPE of 15.04%. Based on the results of the tracking signal, the forecasting results were within the control limits, which shows that the linear regression method is the best forecasting method that can be applied to control coal supply in the next period.**

Keywords—forecasting, linear regression, time series,

I. INTRODUCTION

Currently, coal remains the primary energy source for electric power in Indonesia as it enables the production of thermal electricity at the lowest-cost. Besides, Indonesia has coal resources of around 120.5 billion tons with 31.4 billion tons of coal reserves in 2015 [1]. PT PJB UBJ O&M Tenayan is one of the power plants in Indonesia located on 70 Ring Road, Tenayan Industrial Park, Tenayan Raya District, Pekanbaru City, Riau Province. PLTU Tenayan has a capacity of 2 x 110 MW using coal as the main fuel. It is the largest power plant in Riau providing 31% of energy generation and delivering 4% of power to Sumatra Island.

The most critical aspect in power plant operation is coal inventory. Tenayan's coal mostly comes from Palembang and Jambi. PJB Tenayan does not yet have a coal inventory policy so that inventory is often stored in excessive quantity. Inventory management is needed to prevent excess or shortage of coal demand and increased costs due to losses (coal calorie reduction due to being left too long in the field). Here is the data reception and consumption of coal at PT PJB UBJ O&M Tenayan in 2020.

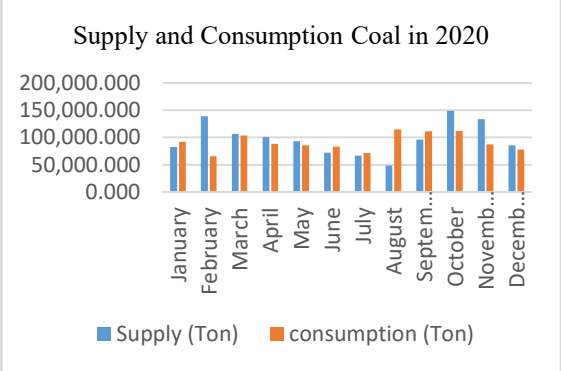

Fig. 1. Coal Supply and Consumtion in 2020

Based on the data presented above, it is clear that there is still an excess or shortage of coal supply. It is caused by several factors such as subsidence of the soil in the coal yard before the coal is combined with soil or the presence of rocks.

We need a method of controlling coal inventory to avoid those problems. Coal supply control can be accomplished by utilizing forecasting tool, which can be used to determine the quantity of coal supply for next period [2].

Forecasting is an art and science that predicts future events based on historical data, which is put in the future using a mathematical model [3]. Forecasting time series data is an important subject in economics, business, and finance. Traditionally, there are several techniques to effectively forecast the next lag of time series data such as univariate Autoregressive (AR), univariate Moving Average (MA), Simple Exponential Smoothing (SES), and, more notably, Autoregressive Integrated Moving Average (ARIMA) with its many variations. [4]. Forecasting is a process that is not difficult, but the mistake of deciding the goal or what is desired will result in different results so that the accuracy of forecasting will be doubted [5]

978-1-6654-1224-7/21/$31.00 ©2021 IEEE

II. LITERATURE REVIEW

A. Forecasting

Forecasting is the art and science of determining future events based on historical evidence to be put in the future using a statistical model. One of the goals of forecasting in industrial management is to pin-point the short-term variations in demand so that manufacturing, labour, supply, and other requirements can be planned.

Forecasting is divided into three categories based on the fundamental characteristics, which are as follows [6].

1. Forecasting is classified based on the characteristics of its constituents.
    a. Subjective forecasting is forecasting based on the views or opinion of the person compiling it. In this case, the opinions or "judgments" of those who assembled them will decide whether the forecast findings are accurate or not.
    b. Objective forecasting is forecasting based on historical data that is real or actual, using techniques and methods for data analysis.
2. Forecasting is divided into two segments based on the forecast period.
    a. Long-term forecasting is forecasting conducted for a period of more than one and a half years or three semesters.
    b. Short-term forecasting is undertaken to predict demand in less than a year and a half.
3. Forecasting based on the type of data pattern is classified into two.
    a. Qualitative method is an approach used without a mathematical model, usually because the existing data is not representative enough to predict the future (long term forecasting).
    b. Quantitative method is an approach based on the availability of raw data along with a series of mathematical rules whose results are utilized to predict future.

Forecasting has several approaches, including the following.

1. Moving average

    Moving average is a method for combining a set of real values and updating them as new values become available [7]. The following is the equation that was used:

    $$Ft = MA_n = \frac{\sum_{i=1}^{n} At\text{-}1}{n} \quad (1)$$

    where Ft is the forecast for the time period t; MA_n is moving average of n periods; At-1 is the actual value in period $t - 1$; and n is the number of periods (data points) in the moving average.

2. Weighted moving average

    Weighted moving average is a forecasting method that is more responsive to changes. This is because the data from the new period are given greater weight. The weighted moving average can be expressed by the following formula [8].

    $$WMA = \frac{\sum (weight\ in\ period\ n)(demand\ in\ period\ n)}{\sum weight} \quad (2)$$

3. Single exponential smoothing

    Single exponential smoothing is a weighted moving average forecasting method where the data have weighted based on the exponential function. Single exponential smoothing can be expressed in a mathematical form as follows [9]:

    $$F_t = F_{t\text{-}1} + \alpha\,(A_{t\text{-}1} - F_{t\text{-}1}) \quad (3)$$

    where Ft is the new forecast; $Ft - 1$ is the previous forecast; and α is smoothing constant $(0 \leq \alpha < 1)$.

4. Linear regression

    Linear regression is a forecasting method with the technique of adjusting trend lines on a series of past data, then projecting them onto a line in the future. The line equation that approaches the form of linear data is as follows.

    $$y = a + bx \quad (4)$$

    To determine the values of a and b, it can be done with the following equation:

    $$b = \frac{\sum xy - N\,(\bar{x})\,(\bar{y})}{\sum x^2 - N\,(\bar{x})^2} \quad (5)$$

    where y is variable prediction (dependent variable); a is the slope of the line; b is y value if $x = 0$; x is independent variable with respect to time; $\bar{x}$ is average of the x values; $\bar{y}$ is average y-value; and n is the number of data or observations.

B. Forecasting Error Measurement

Measurement of forecasting error is a measurement of forecasting accuracy based on the level of difference between forecasting results and actual demand. There are several ways to measure forecasting errors as follows [10].

1. MAD (Mean Absolute Deviation)

    MAD is the mean of the absolute values of deviation. Systematically, MAD is formulated as follows [7].

    $$MAD = \sum \left| \frac{At - Ft}{n} \right| \quad (6)$$

2. MSE (Mean Square Error)

    MSE is determined by adding the squares of all forecast error values for each period and dividing the cumulative number of forecast periods by the number of forecast periods. MSE is formulated as follows.

    $$MSE = \sum \frac{(At - Ft)^2}{n} \quad (7)$$

3. MAPE (*Mean Absolute Percentage Error*)

    MAPE is the average absolute error value over a certain period then multiplied by 100% to get the result in percentage. MAPE is formulated as follows.

    $$MAPE = 100\%\ x\ \sum \left| At - \frac{Ft}{At} \right| \quad (8)$$

C. Forecasting Result Verification

Forecasting results need to be verified to see whether the forecasting results using the selected method will represent the data. Verification of forecasting results can be done with a tracking signal.

The tracking signal is the result of the running sum of the forecast error (RSFE) divided by the Mean Absolute Deviation (MAD), which is used to determine the comparison of the actual value with the forecast value. The value of the tracking signal can be calculated using the following formula.

$$TS = \frac{RSFE}{MAD} \quad (9)$$

According to [7], if the results of the tracking signal value are within the range of ± 6, then the forecast is good. Meanwhile, if the results of the tracking signal value are outside the range of ± 6, the forecasting results are called underforecasting (TS < -6) or overforecasting (TS > 6), so other best forecasting methods are needed.

III. Methodology

This study used time series forecasting methods based on coal usage data at PT PJB UBJ O&M Tenayan in 2020 which was taken directly from the company as follows.

TABLE I. COAL CONSUMPTION IN 2020

No.	Month	Consumption (Ton)
1	January	91,825.00
2	February	65,844.00
3	March	103,956.00
4	April	88,372.00
5	May	85,767.00
6	June	83,282.00
7	July	71,319.00
8	August	114,811.00
9	September	111,375.00
10	October	111,904.00
11	November	87,202.00
12	December	77,907.00

The following is coal reception data at PT PJB UBJ O&M Tenayan in 2020.

TABLE II. COAL SUPPLY IN 2020

No.	Month	Reception (Ton)
1	January	82,854.359
2	February	138,694.045
3	March	106,373.564
4	April	100,815.705
5	May	93,272.202
6	June	72,195.952
7	July	67,000.795
8	August	48,581.560
9	September	96,182.832
10	October	149,021.180
11	November	133,649.562
12	December	85,238.064

The forecasting methods used in this study are as follows.

1. Moving average
   This method was carried out based on a period starting from the 1st period to the 9th period.
2. Weighted moving average
   This method was carried out based on a period starting from the 1st period to the 9th period with the greatest weighting in the latest period for each calculation.
3. Single exponential smoothing
   This method was performed using a smoothing constant starting from $\alpha = 0.1$ to $\alpha = 0.9$.
4. Linear regression
   This method was done by calculating according to the existing formula with variable X showing the time sequence from January to December, while variable Y shows the use of coal.

Data processing was done by manual calculations and using QM software. Determination of the chosen method was taken based on the smallest error value and verification of the forecasting results with the tracking signal.

IV. Results and Discussion

Based on the data processing that has been done, the following results are obtained.

A. Results of data processing using the moving average method

TABLE III. RECAPITULATION OF THE MAD, MSE, AND MAPE VALUES USING THE MOVING AVERAGE METHOD

Period	MAD	MSE	MAPE
Period 2	15,907.65	374,942,000	16.926%
Period 3	16,159.41	381,716,500	17.369%
Period 4	17,367.5	407,840,100	18.676%
Period 5	17,182.46	351,474,200	18.076%
Period 6	19,066.56	409,152,900	19.827%
Period 7	19,450.26	433,850,400	19.012%
Period 8	17,508.63	336,168,200	17.843%
Period 9	14,786.0	261,733,200	15.964%

Based on the results of forecasting data processing using the moving average method, the lowest MAPE value is in period 9, namely 15.96%. This shows that if this method is used, forecasting can be done with a 9-month moving average.

B. Results of data processing using the weighted moving average method

TABLE IV. RECAPITULATION OF THE MAD, MSE, AND MAPE VALUES USING THE WEIGHTED MOVING AVERAGE METHOD

Period	MAD	MSE	MAPE
Period 2	14,951.27	370,349,200	15.82%
Period 3	14,551.98	334,803,100	15.69%
Period 4	16,705.23	378,130,200	17.99%
Period 5	18,158.14	396,723,400	19.38%
Period 6	20,035.93	440,359,900	21.17%
Period 7	20,611.35	466,163,000	20.6%
Period 8	17,692.96	324,570,500	18.56%
Period 9	16,093.52	271,405,800	17.84%

Based on the results of forecasting data processing using the weighted moving average method, the lowest MAPE value is in period 3, which is 15,695. This shows that if this method is used, forecasting can be done with a 3-month moving average.

C. *Results of data processing using the single exponential smoothing method*

TABLE V. RECAPITULATION OF MAD, MSE, AND MAPE VALUES USING THE SINGLE EXPONENTIAL SMOOTHING METHOD

Alpha	MAD	MSE	MAPE
α = 0.1	14,838.52	287,124,000	16.69%
α = 0.2	15,316.29	308,196,100	17.17%
α = 0.3	15,687.23	327,464,400	17.59%
α = 0.4	15,928.57	345,176,400	17.88%
α = 0.5	16,042.77	361,464,400	18.03%
α = 0.6	16,044.86	377,014,400	18.04%
α = 0.7	15,960.76	393,156,600	17.95%
α = 0.8	15,826.3	411,766,500	17.79%

Based on the results of data smoothing using the single exponential smoothing method, the MAPE in the field is 16.69% when the constant of smoothing of 0.1 is used.

D. *Results of data processing using the linear regression method*

TABLE VI. RESULTS OF DATA PROCESSING USING THE LINEAR REGRESSION METHOD

MAD	MSE	MAPE
13,285.63	228,778,800	15.04%

Based on the results of forecasting data processing using the linear regression method, the MAPE value is 15.04%. Therefore, the chosen forecasting method is the linear regression method because it has the lowest MAPE value among other methods, namely 15.04%.

E. *Verifying the results of the forecast value*

Verification of the results of the forecasting value was carried out by tracking the signal against the selected method, namely the linear regression method. Results of the tracking signal are as follows.

TABLE VII. RESULTS OF THE TRACKING SIGNAL

No.	Month	Consumption (Ton)	Forecast (Ton)	Tracking Signal
1	January	91,825.00	85,650.65	1
2	February	65,844.00	86,646.95	-1.08
3	March	103,956.00	87,643.27	0.12
4	April	88,372.00	88,639.57	0.13
5	Mei	85,767.00	89,635.88	-0.26
6	June	83,282.00	90,632.18	-1.07
7	July	71,319.00	91,628.48	-2.81
8	August	114,811.00	92,624.8	-0.65
9	September	111,375.00	93,621.1	0.77
10	October	111,904.00	94,617.41	2.05
11	November	87,202.00	95,613.71	1.46
12	December	77,907.00	96,610.02	0

For more details, the results of the tracking signal can be seen in the following graph.

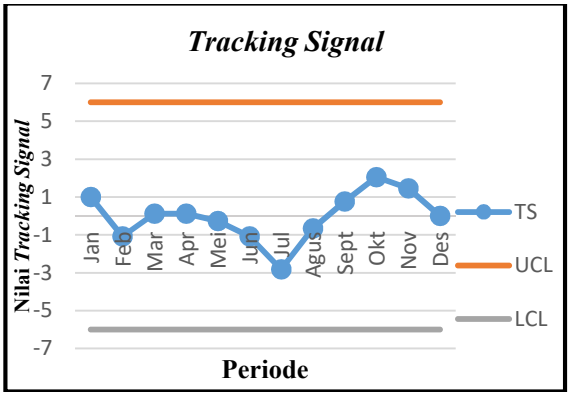

Fig. 2. Results of the tracking signal calculation

Based on the tracking signal chart above, it can be concluded that the chosen method is the best method because the value of the tracking signal is still within the control limit range of ± 6 so that the accuracy of forecasting is acceptable and no other method is needed.

After obtaining the selected method, namely the linear regression method, forecasting was carried out for the next period. Table VIII displays the forecast results for the next period.

TABLE VIII. 2021 FORECASTING RESULTS USING THE LINEAR REGRESSION METHOD

No.	Month	Forecast (ton)
1	January	97,606.33
2	February	98,602.63
3	March	99,598.94
4	April	100,595.2
5	May	101,591.5
6	June	102,587.9
7	July	103,584.2
8	August	104,580.5
9	September	105,576.8
10	October	106,573.1
11	November	107,569.4
12	December	108,565.7

V. CONCLUSION

In this study, results of measurement using moving average, weighted moving average, single exponential smoothing, and linear regression methods were compared to determine which one has smallest errors. The forecasting results show that the lowest MAPE in this case is linear regression method which has 14,04%. Based on the results of the tracking signal, the forecasting results were verified to be within the control limits, which is in the range of ± 6.

The contribution of this research is comparing four forecasting methods and its direct application in forecasting field that has not yet been applied by the company in determining the amount of coal inventory. This company can choose linier regression method as the method to forecast how to control inventory of coal supply.

REFERENCES

[1] Aditya, I., Simaremare, A. A., & Hudaya, C. (2019, April). Study of Coal Inventory Planning Analysis in a Coal-Fired Power Plant Using Continuous and Periodic Review. In 2019 IEEE 2nd International Conference on Power and Energy Applications (ICPEA) (pp. 33-36). IEEE.

[2] Rahmi, Amalliah, and Taufiq Taufiq. "Peramalan Persediaan Material Batubara Dengan Metode Single Exponential Smoothing." Progresif: Jurnal Ilmiah Komputer 13.1 (2017).

[3] Khiram, Hibatul, and Wirdah Irawati. "Analisis Perencanaan Sistem Distribusi pada PT. Lafarge Cement Indonesia Aceh Besar." Jurnal Ilmiah Mahasiswa Ekonomi Manajemen 2.1 (2017): 118-134.

[4] Siami-Namini, S., Tavakoli, N., & Namin, A. S. (2018, December). A comparison of ARIMA and LSTM in forecasting time series. In 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA) (pp. 1394-1401). IEEE.

[5] Syah, R., Nasution, M. K., & Elveny, M. (2021). Sensitivity of MAPE using detection rate for big data forecasting crude palm oil on k-nearest neighbor. International Journal of Electrical & Computer Engineering (2088-8708), 11(3).

[6] Robial, Siti Muawanah. "Perbandingan Model Statistik Pada Analisis Metode Peramalan Time Series:(Studi Kasus: Pt. Telekomunikasi Indonesia, Tbk Kandatel Sukabumi)." Santika: Jurnal Ilmiah Sains Dan Teknologi 8.2 (2018): 823-838.

[7] Saptaria, Lina. "Analisis Peramalan Permintaan Produk Nata De Coco Untuk Mendukung Perencanaan Dan Pengendalian Produksi Dalam Supply Chain Dengan Model Cpfr (Collaborative Planning, Forecasting, And Replenishment)." Jurnal Nusantara Aplikasi Manajemen Bisnis 2.2 (2017): 130-141.

[8] Mardhiyah, Zulfa, Rio Aurachman, and Putu Giri Artha Kusuma. "Penentuan Jumlah Perencanaan Permintaan Terhadap Produk Aqua Dengan Metode Peramalan Time Series (studi Kasus Pada Pt Tirta Investama Bandung)." eProceedings of Engineering 7.2 (2020).

[9] Nugraha, Eucharistia Yacoba, and I. Wayan Suletra. "Analisis Metode Peramalan Permintaan Terbaik Produk Oxycan pada PT. Samator Gresik." Jurnal Seminar dan Konferensi Nasional IDEC. 2017

[10] Mahardhika, Adri Dharmesta, and Novie Susanto. "Peramalan Perencanaan Produksi Terak Dengan Metode Exponential Smoothing with Trend Pada PT. Semen Indonesia (Persero) Tbk." Industrial Engineering Online Journal 6.1 (2017).