

Research Article

# Predicting the Number of Train Passengers in Java Island using SARIMA Model

Dimas Wahyu Saputro<sup>1\*</sup>, Dhea Sukma Agustiana<sup>1</sup>, Fadia Dilla Sabine<sup>2</sup>, Lis Nurani<sup>3</sup>, Rezki P. Manullang<sup>4</sup>, Taj Shavira<sup>5</sup>, Mika Alvionita<sup>6</sup>, Luluk Muthoharoh<sup>7</sup>

<sup>1,2,3,4,5,6,7</sup> Data Science Study Program, Science Department, Sumatra Institute of Technology Jl. Terusan Ryacudu, Way Huwi, Kec. Jati Agung, Kabupaten Lampung Selatan, Lampung 35365

\* Corresponding author: [luluk.muthoharoh@sd.itera.ac.id](mailto:luluk.muthoharoh@sd.itera.ac.id)

Received: 5 August 2022; Accepted: 2 September 2022; Published: 23 August 2023 (40-48)

**Abstract:** The train at this time has become one of the most popular public transportation for medium and long-distance travel. The number of train passengers is difficult to predict during the holiday season. This study aimed to predict the number of train passengers using the Seasonal Autoregressive Integrated Moving Average (SARIMA) method. The stages used in this study include (1) dataset preparation, (2) preprocessing data, and (3) experimental testing and methods. The SARIMA model obtained is ARIMA(2,1,0)(0,1,2)[12] with an AIC value of 2379,265. A diagnostic model was carried out, and it was found that the model is quite good. So, the SARIMA method used in predicting passengers is accurate.

**Keywords:** Density-Based Spatial Clustering of Application Noise (DBSCAN), Quadrant Method, Ripley's K-Function, Silhouette Average Width

## Introduction

One of the vehicles of transportation used by the community to travel across cities or provinces is the train. Many indicators are used to describe Indonesia's economic conditions, but most of these indicators are released by the Central Statistics Agency (BPS) with a time lag. One of the strategic indicators that have a domino effect on other economic sectors is the number of train passengers [1]. The train has existed across the Java and Sumatra islands and has even become one of the most popular forms of transportation. According to a report from the Ministry of Transportation on the Java island in 2020, there are around 463 stations that serve the boarding/descending of train passengers [2].

One of the vehicles of transportation used by the community to travel across cities or provinces is the train. Many indicators are used to describe Indonesia's economic conditions, but most of these indicators are released by Central Statis Agency with a time lag. On eof strategic indicators that have adoino effect on other economic sectors is the number of train passenger. The train has existed across the Java and Smatra Islands and has even become one of the most popular forms of transportation. According to a report from the Ministry of Transportation on the Java Islan in 2020, there are around 463 stations that serve the boarding/descending of train passenger.

Along with the many enthusiasts of rail transportation, especially in Java, the prediction of the number of train passengers has become very important. Considering the aspects of price, comfort, security and so on, trains are still the choice for most Indonesian people. The train is one of the modes of transportation-oriented to the public interest and its customers. With its advantages, rail transportation is one of the products chosen by most Indonesian people [3].

Many indicators are used to describe Indonesia's economic conditions, but most of these indicators are released by the Central Statistics Agency (BPS) with a time lag. One of the strategic indicators that have a domino effect on other economic sectors is the number of train passengers. We are modelling to see the number of Train Passengers in Java using the past data to create/determine historical patterns of data to extrapolate these patterns in the future. While research related to forecasting the number of train passengers in Indonesia, Katabba (2021)

uses the SARIMA model (0,1,1) (2,1,1)12 and Arianto (2017) uses the ARIMA Box-Jenkins model (1,1, 12) in forecasting the actual value of the number of train passengers [1], [4], [5] .

The Seasonal Autoregressive Integrated Moving Average (SARIMA) method is a method that ignores independent variables in forecasting [6]. The SARIMA method has a dependent variable consisting of past and present data, which results in short-term forecasting[7]. The SARIMA method is a technique that observes changes in the dependent data and in a short period[8]. The plot will be flat (flat/constant) if the period is long enough. The SARIMA method uses past and present data from the dependent variable to produce accurate short-term forecasts [7]. The SARIMA method is under observations from time series (time series), which are statistically related (dependent)[9].

Various research topics have been carried out to find how much the train passengers have increased, including using the Artificial Neural Network Method and SARIMA to estimate the number of train passengers in Indonesia. This research will help the government make appropriate improvements in the railway system for the future and evaluate potential losses due to COVID -19 [10]. By Using this algorithm. It is found that the root means squared error (RMSE) using Backpropagation Artificial Neural Network is smaller than SARIMA (2,1,0)(0,1,2)12. Evelina et al.'s research in 2013 was regarding the "Exponential Smoothing Holt-Winters Algorithm for predicting the number of passengers for the Medan-Rantau Prapat train" [9]. This algorithm can only produce up to 40% accuracy using the Seasonal Addition Method. Irma Eliza has developed "Predictions of the Number of Passengers Airplane PT. Angkasa Pura II Sultan Syarif Kasyim II Airport Pekanbaru Using the ARIMA Algorithm (0,1,1)" in 2011[10]. This study has the same pattern of forecasting results as the actual pattern. It produces an accuracy of 76.26% with only 72 data, so the level of accuracy will increase. Sinyo et al. have developed "Passenger Prediction at Manado's Sam Ratulangi International Airport Using the ARIMA Algorithm (1,1,1)" in 2015[11]. This research can produce a reasonable prediction accuracy rate of 80.12%.

Then the research of Jie et al. in 2017 has developed the Wavelet-ARIMA Algorithm to predict Beijing City Railway traffic [12]. The wavelet-ARIMA model can achieve an accuracy of 96%. Furthermore, the ARIMA-RBF Algorithm predicted the amount of rail traffic in the city of Jiuran-He and Si in 2013 [13]. This study uses the MAPE accuracy-test evaluation model, which can achieve a high level of accuracy, namely 98%. Then research was conducted using the SARIMA algorithm to predict vehicle traffic flow in the city of Bandung Billy and Lester in 2003[14]. The SARIMA algorithm can produce a higher level of accuracy, which is 91.4%, compared to the Random walk model, which is 89.6%. Regarding the above problems, this study this study aims to predict the number of train passengers on the island of Java using the SARIMA method. This method can indicate extensive data, produce accurate short-term predictions, be efficient in predicting time series, and deal with seasonal data trends where the number of passengers changes every day.

## Materials and Methods

### Participant

The data used is data on the number of train passengers on the island of Java from the Indonesian Central Statistics Agency [15]. The amount of data used is 168 data from January 2006 to December 2019. The dataset consists of one column, namely the column for the number of passengers in thousand units.

### Instrument

The method used is the SARIMA method (Seasonal Autoregressive Integrated Moving Average) which is an extension of the ARIMA method[15]. SARIMA is a method that ignores independent variables in forecasting [8]. Forecasting used or used is a parameter for forecasting a particular case that occurs. The SARIMA method is based on a variable's values that have occurred in the past [16]. SARIMA, which produces a prediction or prediction, is synthetic modelling based on data patterns that have been obtained historically. Then, the data in the past is used to create/determine historical patterns of data, which will then be used to extrapolate these patterns in the future [17].

ARIMA is also known as the Box-Jenkins time series method. ARIMA models are divided into three groups, including:

Autoregressive Model (AR), expressed as follows:

$$X_t = \mu' + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + e_t \quad \text{eq. 1}$$

Moving Average (MA), expressed as follows:

$$X_t = \mu' + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \dots - \theta_q e_{t-q} \quad \text{eq. 2}$$

The mixed model, the general model for a pure AR(1) and pure MA(1) process mixture, is expressed as follows:

$$\begin{aligned} X_t &= \mu' + \phi_1 X_{t-1} + e_t - \theta_1 e_{t-1} \\ \text{Or} \\ (1 - \phi_1 B) X_t &= \mu' + (1 - \theta_1 B) e_t \end{aligned} \quad \text{eq. 3}$$

In SARIMA modelling, the method we used was SARIMA with degrees AR (p), degrees of difference (d), and degrees MA (q). The model written in SARIMA can be written as (p, d, q), which has the general form as follows [14].

$$Y_t = \alpha + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \dots + \beta_p Y_{t-p} + \epsilon_t + \phi_1 \epsilon_{t-1} + \phi_2 \epsilon_{t-2} + \dots + \phi_q \epsilon_{t-q} \quad \text{eq. 4}$$

The modelling of SARIMA is basically divided into several stages, including Identification of the model, estimation of its parameters, testing for the model and diagnostics of the ARIMA [18].

A way to identify seasonal models is by checking the autocorrelation and partial autocorrelation of the system. The steps for identifying modelling include the following:

1. Making a plot to check whether the seasonal pattern is in the seasonal pattern or not
2. Perform stationarity on the data. The function of the existence of stationarity in the data is to see whether the data is stationary in its variation or stationary in its mean. In this case, if it is standing in its normal, it must be different and vice versa. A transformation must be carried out if it is not static in its variation.
3. Plot the ACF and PACF. This is done to see if the data created is stationary after the differencing process is carried out or after the transformation process is carried out. And another use of plotting on ACF and PACF is to predict a possible model following the data used.

In time-series data, there are several patterns, namely trend patterns, horizontal patterns, cyclical patterns and seasonal patterns. The SARIMA method uses seasonal data patterns, which are repeated in a fixed time interval. Identify seasonal patterns using autocorrelation on two or three time-lags significantly different from zero for static data [19]. Autocorrelation, which is substantially different from zero, indicates the presence of a pattern in the data. One must look at the high autocorrelation to recognize the existence of seasonal factors. To deal with seasonality, the short standard notation is:

$$ARIMA(p, d, q) (P, D, Q)^S \quad \text{eq. 5}$$

where:

(p,d,q) : the non-seasonal part of the model

(P,D,Q) : seasonal part of the model

S : number of periods per season

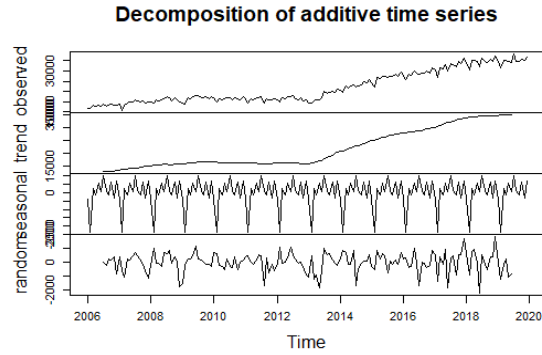
There are two basic ways to get these parameters:

- a. Trial and error by testing several different values and selecting the one value that minimizes the sum of squared residuals.
- b. Iterative refinement, selecting an initial estimate and then letting the computer program refine the estimate iteratively.

## Result and Discussion

### Decomposition of Time Series

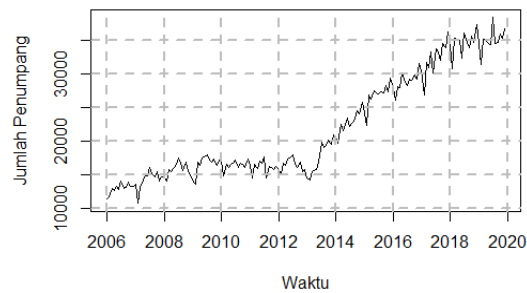
Decomposition Time Series is to parse a time series and separate it into its constituent components, which are often trend, random, and seasonal components. In R programming, the plot(decompose(data\_ts)) command is used. After executing this line of code, it will produce a plot, as shown in Figure 1. The time series is not stationary due to seasonal and trend effects.



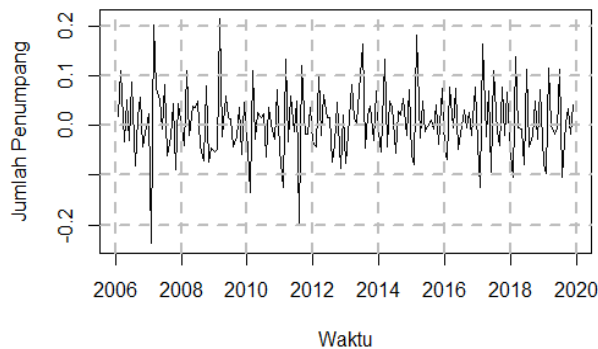
**Figure 1.** Decomposition of Time Series

### Stationary Test

Stationary is a time-series data that tends to move toward the mean and variance. The number of train passengers dataset on Java island was initially not stationary. This is reflected in the graph in Figure 3, so it is necessary to do differencing logs. The Augmented Dickey-Fuller (ADF) test is carried out for the differencing record to be carried out correctly. The ADF test is carried out on time-series data to determine whether the time series data is stationary. The ADF test was carried out on the dataset so that the p-value of the ADF test was 0.745, where the p-value was more significant than 0.05, meaning that it failed to reject the null hypothesis or the dataset was not stationary. In other words, time-series data is time-dependent and does not have constant variance over time. This is following Figure 2. Then do differencing logs, and check the dataset with the ADF test. The p-value is 0.01, where the p-value is less than 0.05, meaning that it rejects  $H_0$  or the dataset is stationary. This is following Figure 3. Stationary is indicated by the average and variance of the data on the number of train passengers on the island of Java which is constant from time to time.



**Figure 2.** Number of Train Passengers



**Figure 3.** Number of Train Passengers After Differencing

### Determine of Order

After checking the stationary data set, the ARIMA method should estimate the dataset. We observed the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots. ACF checks the autoregressive value while PACF checks the Moving Average value. Data that has Autoregressive (AR) or Moving Average (MA) is to see the pattern of the ACF and PACF. The patterns formed in ACF and PACF are the Cut Off and dying down patterns. From these patterns, it can then be determined the best ARIMA model, according to Figure 4.

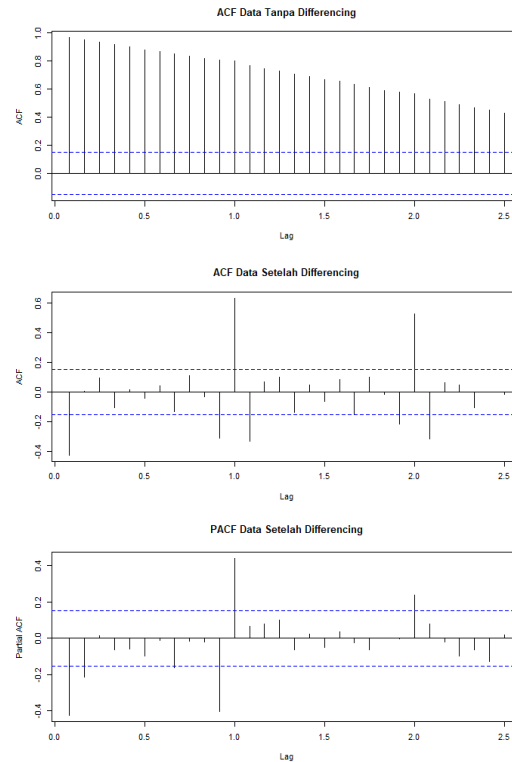


Figure 4. ACF and PACF Plot

The values of  $p$ ,  $d$ , and  $q$  can be seen from the ACF and PACF plots. However, in the R language, there is a library to get these values that will be used for ARIMA. Auto Arima is one of the packages or libraries in R that makes it easy for users, so there is no need to search for the best ARIMA model manually. By using Auto Arima, users are required to import the Forecast library.

```
> arimaAP <- auto.arima(data_ts, trace=TRUE, ic="aic")
Fitting models using approximations to speed things up...
ARIMA(2,1,2)(1,1,1)[12] : 2386.179
ARIMA(0,1,0)(0,1,0)[12] : 2424.681
ARIMA(1,1,0)(1,1,0)[12] : 2392.883
ARIMA(0,1,1)(0,1,1)[12] : 2385.297
ARIMA(0,1,1)(0,1,0)[12] : 2407.306
ARIMA(0,1,1)(1,1,1)[12] : 2390.862
ARIMA(0,1,1)(0,1,2)[12] : 2384.577
ARIMA(0,1,1)(1,1,2)[12] : 2385.317
ARIMA(0,1,0)(0,1,2)[12] : 2408.152
ARIMA(1,1,1)(0,1,2)[12] : 2382.915
ARIMA(1,1,1)(0,1,1)[12] : 2384.756
ARIMA(1,1,1)(1,1,2)[12] : 2388.137
ARIMA(1,1,1)(1,1,1)[12] : 2390.359
ARIMA(1,1,0)(0,1,2)[12] : 2384.636
ARIMA(2,1,1)(0,1,2)[12] : 2380.63
ARIMA(2,1,1)(0,1,1)[12] : 2383.463
ARIMA(2,1,1)(1,1,2)[12] : 2387.489
ARIMA(2,1,1)(1,1,1)[12] : 2389.674
ARIMA(2,1,0)(0,1,2)[12] : 2379.265
ARIMA(2,1,0)(0,1,1)[12] : 2381.829
ARIMA(2,1,0)(1,1,2)[12] : 2385.843
ARIMA(2,1,0)(1,1,1)[12] : 2387.976
ARIMA(3,1,0)(0,1,2)[12] : 2381.437
ARIMA(3,1,1)(0,1,2)[12] : 2383.306

Now re-fitting the best model(s) without approximations...
ARIMA(2,1,0)(0,1,2)[12] : 2567.015
Best model: ARIMA(2,1,0)(0,1,2)[12]
```

Figure 5. AIC Value of Each Model

When using Auto Arima, it will display the AIC value. Each model used can be measured or compared by looking at the Akaike Information Criterion (AIC). \*AIC is a general method for determining how well a model fits the data, and the smallest AIC value is the best fitting model. \*When viewed from the data above, the lowest AIC is when ARIMA(2,1,0)(0,1,2)[12] with an AIC value of 2379,265. Therefore, the ARIMA model will be used for forecasting according to Figure 5.

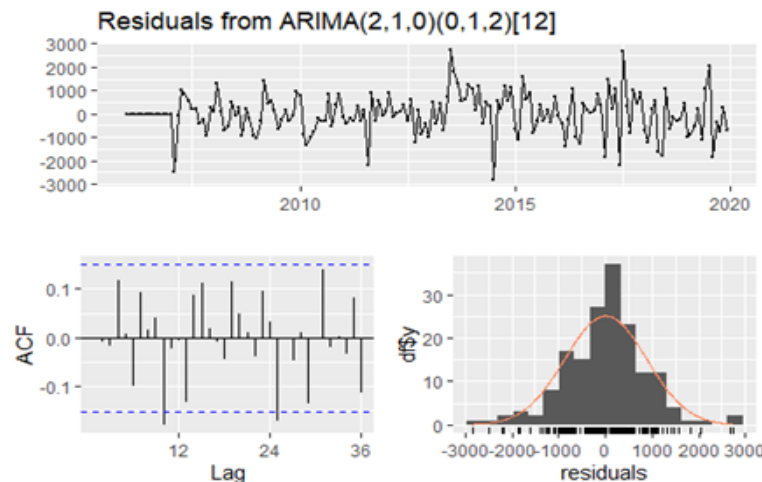
```
> arimaAP
Series: data_ts
ARIMA(2,1,0)(0,1,2)[12]

Coefficients:
      ar1      ar2      sma1      sma2
      -0.4406 -0.2333 -0.4988 -0.1203
s.e.    0.0800  0.0806  0.0865  0.0846

sigma^2 = 848530: log likelihood = -1278.51
AIC=2567.01  AICC=2567.42  BIC=2582.23
```

**Figure 6.** AIC, AICC and BIC Values of the Best Models

After getting the best model value, a diagnostic model is carried out. To perform diagnostics for residuals and acf will use the ggfortif library. It can be seen in Figure 7 that the ARIMA model above is quite good because the residual is centred around the zero value.



**Figure 7.** Diagnostic Model

Based on the plot above, it can be seen that the residuals do not follow the normal distribution. Furthermore, from the ACF and PACF plots, it can be seen that there is no significant lag. It shows that there is no autocorrelation symptom in the residuals. Furthermore, a formal assumption test will be carried out to make sure again.

Data Normality Formal Test

H0 : residual is normally distributed

H1: the remainder does not follow the normal distribution

```
> sisaan <- arimaAP$residuals
> # Uji formal normalitas data
> ks.test(sisaan,"pnorm")

One-sample Kolmogorov-Smirnov test

data: sisaan
D = 0.49911, p-value < 2.2e-16
alternative hypothesis: two-sided
```

**Figure 8.** Data Normality Formal Test

The test results, the p-value <0.05, means that H0 is rejected, which means the remainder is not normally distributed.

#### Residual Middle Value Test

$$H_0: \mu = 0$$

$$H_1: \mu \neq 0$$

```
> t.test(sisaan, mu = 0, alternative = "two.sided")

One Sample t-test

data: sisaan
t = 0.083295, df = 167, p-value = 0.9337
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
-127.7869 139.0445
sample estimates:
mean of x
5.628835
```

**Figure 9.** Residual Middle Value Test

The test results, the p-value > 0.05, means that H0 is not rejected, which means that the median value of the remainder is equal to zero.

#### Autocorrelation Test

$$H_0 = 0, \text{ there is no autocorrelation}$$

$$H_1 \neq 0 \text{ there is autocorrelation}$$

```
> # Uji autokorelasi
> Box.test(sisaan, lag = 23 ,type = "Ljung")

Box-Ljung test

data: sisaan
X-squared = 24.745, df = 23, p-value = 0.3636
```

**Figure 10.** Residual Middle Value Test

The test results, the p-value > 0.05, means that H0 is not rejected, which means that there are no symptoms of autocorrelation.

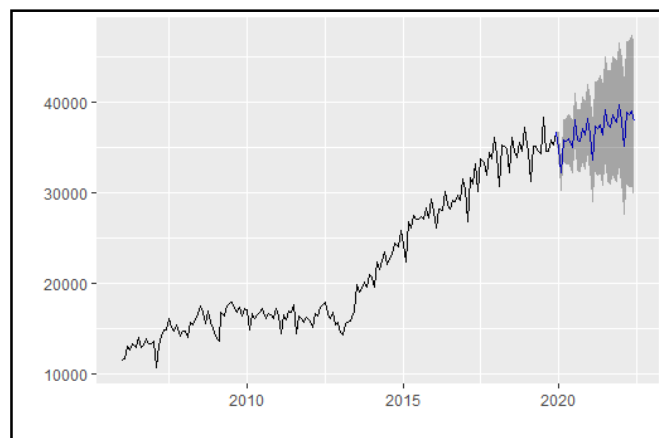
Conclusion: Assumptions are met, except that the remainder is not normally distributed.

#### Forecast

Prediction test is done by determining the number of days to be predicted. In this case, the test is carried out to predict the number of train passengers for the next 30 months. The prediction results using SARIMA can be seen in Figure 11.

```
> forecastAP <- forecast(arimaAP, level = c(95), h = 30)
> forecastAP
      Point Forecast      Lo 95      Hi 95
Jan 2020 34922.34 33116.89 36727.78
Feb 2020 32166.64 30097.92 34235.36
Mar 2020 35797.68 33525.73 38069.64
Apr 2020 35644.68 33095.54 38193.81
May 2020 35924.69 33157.39 38691.99
Jun 2020 35013.33 32049.52 37977.14
Jul 2020 37985.59 34828.87 41142.31
Aug 2020 35860.65 32524.59 39196.71
Sep 2020 35655.45 32150.01 39160.88
Oct 2020 37008.72 33340.86 40676.57
Nov 2020 36321.81 32498.62 40145.00
Dec 2020 38108.95 34136.59 42081.30
Jan 2021 36331.42 31891.35 40771.48
Feb 2021 33603.72 28889.30 38318.14
Mar 2021 37285.91 32323.44 42248.38
Apr 2021 37079.90 31847.79 42312.02
May 2021 37505.52 32028.83 42982.21
Jun 2021 36412.80 30704.23 42121.36
Jul 2021 39183.95 33249.09 45118.80
Aug 2021 37414.84 31263.01 43566.68
Sep 2021 37141.97 30780.89 43503.05
Oct 2021 38532.47 31968.45 45096.49
Nov 2021 37814.23 31053.43 44575.02
Dec 2021 39691.08 32739.13 46643.03
Jan 2022 37881.33 30556.62 45206.04
Feb 2022 35146.90 27553.04 42740.76
Mar 2022 38839.57 30992.61 46686.54
Apr 2022 38630.52 30517.80 46743.24
May 2022 39055.04 30692.34 47417.73
Jun 2022 37963.51 29359.52 46567.50
```

**Figure 11.** Predictions Obtained from the Model



**Figure 12.** Plot Prediction

## Conclusion

Based on the results and discussion above, the best model to predict the number of train passengers on the island of Java is to use the ARIMA model  $(2,1,0)(0,1,2)$  [12] with an AIC value of 2379,265. Although it doesn't use too much data, 2006-2020, good results are obtained. The test is carried out to predict the number of train passengers for the next 30 months.



## References

- [1] P. D. W. Oktama, "Nowcasting Jumlah Penumpang Kereta Api di Indonesia Menggunakan Indeks Google Trends," *Semin. Nas. Off. Stat.*, vol. 2021, no. 1, pp. 958–967, 2021, doi: 10.34123/semnasoffstat.v2021i1.820.
- [2] M. A. Rizaty, "Berapa Jumlah Stasiun Kereta Api Penumpang di Jawa dan Sumatera?," *databooks*, 2021. <https://databoks.katadata.co.id/datapublish/2021/11/04/berapa-jumlah-stasiun-kereta-api-penumpang-di-jawa-dan-sumatera> (accessed May 30, 2022).
- [3] Suryadi, "Kinerja Dan Peramalan Pertumbuhan Angkutan Kereta Api Menggunakan Model Sarima," *War. Penelit. Perhub.*, vol. 26, no. 7, p. 381, 2019, doi: 10.25104/warlit.v26i7.922.
- [4] Y. I. Katabba, "Metode Seasonal Autoregressive Integrated Moving Average (SARIMA) untuk Memprediksi Jumlah Penumpang Kereta Api di Pulau Sumatera," Jambi University, 2021.
- [5] B. W. Arianto, "Peramalan Jumlah Penumpang Kereta Api Di Pulau Jawa Dan Sumatera Menggunakan Arima Box-Jenkins," Institut Teknologi Sepuluh November, 2017. [Online]. Available: <http://repository.its.ac.id/43384/>
- [6] T. Widiyaningtyas, Muladi, and A. Qonita, "Use of ARIMA Method to Predict the Number of Train Passenger in Malang City," *Proceeding - 2019 Int. Conf. Artif. Intell. Inf. Technol. ICAIIT 2019*, pp. 359–364, 2019, doi: 10.1109/ICAIIIT.2019.8834663.
- [7] A. Qonita, A. G. Pertiwi, and T. Widiyaningtyas, "Prediction of rupiah against us dollar by using ARIMA," *Int. Conf. Electr. Eng. Comput. Sci. Informatics*, vol. 2017-Decem, no. September, pp. 19–21, 2017, doi: 10.1109/EECSI.2017.8239205.
- [8] C. F. Chen, Y. H. Chang, and Y. W. Chang, "Seasonal ARIMA forecasting of inbound air travel arrivals to Taiwan," *Transportmetrica*, vol. 5, no. 2, pp. 125–140, 2009, doi: 10.1080/18128600802591210.
- [9] N. AISHAH, D. DEVIANTO, and M. MAIYASTRI, "Pemodelan Jumlah Kunjungan Wisatawan Mancanegara Ke Indonesia Melalui Bandara Ngurah Rai Bali Dengan Model Sarima-Arch," *J. Mat. UNAND*, vol. 10, no. 3, p. 248, 2021, doi: 10.25077/jmu.10.3.248-259.2021.
- [10] E. Padang, G. Tarigan, and U. Sinulingga, "Peramalan Jumlah Penumpang Kereta Api Medan-Rantau Prapat Dengan Metode Pemulusan Eksponensial Holt-Winters," *Saintia Mat.*, vol. 1, no. 2, pp. 161–174, 2013.
- [11] S. Salmon, N. Nainggolan, and D. Hatidja, "Pemodelan ARIMA Dalam Prediksi Penumpang Pesawat Terbang Pada Bandara Internasional Sam Ratulangi Manado," *d'CARTESIAN*, vol. 4, no. 1, p. 59, 2015, doi: 10.35799/dc.4.1.2015.8099.
- [12] J. Zhu, W. Xu, H. Jin, and H. Sun, "Prediction of Urban Rail Traffic Flow Based on Multiply Wavelet-ARIMA Model Wavelet analysis Threshold processing," pp. 1–12, 2018, doi: 10.1007/978-981-10-3551-7.
- [13] J. He and B. Si, "The application of ARIMA-RBF model in urban rail traffic volume forecast," *Proc. 2nd Int. Conf. Comput. Sci. Electron. Eng. (ICCSEE 2013)*, vol. 1, no. March 2013, 2013, doi: 10.2991/iccsee.2013.416.
- [14] B. M. Williams and L. A. Hoel, "Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: Theoretical basis and empirical results," *J. Transp. Eng.*, vol. 129, no. 6, pp. 664–672, 2003, doi: 10.1061/(ASCE)0733-947X(2003)129:6(664).
- [15] W. Rahmalina, "Pemodelan Seasonal Autoregressive Integrated Moving Average Untuk Memprediksi Jumlah Kasus Covid-19 di Padang," *J. Mat. Integr.*, vol. 17, no. 1, p. 23, 2021, doi: 10.24198/jmi.v17.n1.32024.23-31.
- [16] D. Ruhiat and A. Effendi, "Pengaruh Faktor Musiman Pada Pemodelan Deret Waktu Untuk Peramalan Debit Sungai Dengan Metode Sarima," *Teorema*, vol. 2, no. 2, pp. 117–128, 2018, doi: 10.25157/.v2i2.1075.
- [17] R. Ayu, R. Gernowo, D. Fisika, F. Sains, U. Diponegoro, and S. E-, "Metode Autoregressive Integrated Movingaverage (Arima) Dan Metode Adaptive Neuro Fuzzy Inference System (Anfis) Dalam Analisis Curah Hujan," *Berk. Fis.*, vol. 22, no. 1, pp. 41–48, 2019.
- [18] M. Al Aradi and N. Hewahi, "Prediction of Stock Price and Direction Using Neural Networks: Datasets Hybrid Modeling Approach," *2020 Int. Conf. Data Anal. Bus. Ind. W. Towar. a Sustain. Econ. ICDABI 2020*, 2020, doi: 10.1109/ICDABI51230.2020.9325697.
- [19] E. Erdem and J. Shi, "ARMA based approaches for forecasting the tuple of wind speed and direction," *Appl. Energy*, vol. 88, no. 4, pp. 1405–1414, 2011, doi: 10.1016/j.apenergy.2010.10.031.