JTAM (Jurnal Teori dan Aplikasi Matematika)
http://journal.ummat.ac.id/index.php/jtam
p-ISSN 2597-7512 | e-ISSN 2614-1175
Vol. 6, No. 4, October 2022, pp. 1060-1080

Forecasting the Number of Ship Passengers with SARIMA Approach (A Case Study: Semayang Port, Balikpapan City)

Multiningsih[1], Emy Siswanah[2*], Minhayati Saleh[3]
[1,2,3]Department of Mathematics, Universitas Islam Negeri Walisongo Semarang, Indonesia
multiningsih24@gmail.com[1], emysiswanah@walisongo.ac.id[2], minhayati@walisongo.ac.id[3]

Article History:
Received : 29-07-2022
Revised : 26-09-2022
Accepted : 05-10-2022
Online : 08-10-2022

Keywords:
Forecasting;
SARIMA;
MAPE.

ABSTRACT

From year to year, the number of ship passengers at Semayang Port, Balikpapan city tends to fluctuate. It also doubles in certain months and repeats every year. Sea transportation companies need to make forecasts in order to implement policies related to predict the number and capacity of ships that need to be provided as well as the preparation of port facilities. The study aims at obtaining the best model, predicting and determining the accuracy of the forecasting results for the number of passengers arriving and departing at Semayang Port, Balikpapan city using SARIMA method. The SARIMA method is a time series data forecasting method that is able to identify seasonal patterns. The results showed that the best model for predicting the number of passengers departing at Semayang Port, Balikpapan city is the SARIMA $(4,1,0)(0,1,2)^{12}$ model with a MAPE of 14.05%. It means that the SARIMA model used produces good forecasting. Meanwhile, the best model to predict the number of passengers coming to Semayang Port Balikpapan city is the SARIMA $(0,1,1)(2,1,0)^{12}$ model with a MAPE value of 3.27% which exposes that the SARIMA model used succeed to provide accurate forecasting. The results of this forecast can be used as a reference for the government or port managers to anticipate a surge in passengers. The government or port management can prepare an adequate amount of transportation in certain months to avoid the accumulation of passengers and to make sea transportation more efficient.

Crossref
https://doi.org/10.31764/jtam.v6i4.10211

A. INTRODUCTION

Transportation is a way of fulfilling the needs of individual to be able to move from one location to another (Fatimah, 2019). It is also used to carry out activities and find the items needed. Advances in transportation can increase the mobility of people, marketed factors and the factors of production (Salim, 2016). Related to human mobility, transportation has an important role in various aspects, including social and cultural aspects, technical, economic and political and defense aspects (Nasution, 2015). Land and air transportations are not reliable means of transportation considering Indonesia's geography as an archipelagic country with more than 17,500 islands there (Putra et al., 2022). The existence of sea transportation plays an important role in facilitating inter-island relations which are an integral part of the archipelago (Nasution, 2015).

The sea port is one of sea transportation infrastructures that supports an important role in the acceleration transportation activities such as the entry and the exit of goods, the shift of passengers in shelters and as a place where passengers get into and get down from the ships.

In addition, the ports are the connecting infrastructures among islands and countries. Sea ports are supported by means of transportation in the form of ships that are useful for transporting cargo, both goods and passengers. The largest port in Balikpapan that serves inter-island routes is Semayang Port (BPS, 2019). It is ranked third for the density of passenger arrivals and departures after Makassar and Tanjung Perak ports.

The number of ship passengers arriving and departing from Semayang Port, Balikpapan city from year to year tends to fluctuate. The increase in the number of ship passengers has doubled in the months leading up to the Eid holiday and is repeated every year. The increase number of passengers occurred because of the desire of the nomads to return to their hometowns (BPS, 2019). The increase number of passengers causes the accumulation of the number of passengers and the flow of loading and unloading becomes clogged. As a result, there was a traffic jam for hours and passengers overflowed in the waiting room of the port. In the same time, the capacity of the port was limited and cramped. Sea transportation companies need to make forecasts to find out the estimated number of ship passengers so that the company is able to anticipate the increase number of passengers and can implement more prudent policies for the future. These policies relate to the estimation of the number and capacity of ships that need to be provided as well as the preparation of port facilities. Errors in planning the adequacy of the number and capacity of ships can be minimized by forecasting efforts (Nasution, 2015; Prabhadika et al., 2018). Forecasting is the process of predicting future events based on data from previous events (Andini & Sunyoto, 2018; Herjanto, 2007; Hyndman & Athanasopoulos, 2018; Montgomery et al., 2015).

The collected data for the number of arrivals and departures of ship passengers at Semayang Port, Balikpapan is in the form of time series data. Forecasting time series data is carried out by looking at the pattern of data in the past which is collected periodically based on the sequence, time either in daily, weekly, monthly, quarterly and yearly (Chatfield, 2000; Soelaeman, 2016). The form of time series data patterns is divided into four types, namely seasonal, horizontal, cyclical and trend patterns (Kokilavani & dkk, 2020; Makridakris et al., 1997). The forecasting method used to predict the number of ship passengers arriving and departing from Semayang Port, Balikpapan is SARIMA (Seasonal Autoregressive Integrated Moving Average) method.

The SARIMA method is able to identify seasonal and non-seasonal patterns in its forecasting model so that the identified data forms seasonal patterns that can be predicted using the SARIMA method (J. Liu et al., 2022; Tadesse & Dinka, 2017). A data is said to have a seasonal pattern if the data shows periodic behavior at certain intervals (Yusof & Kane, 2012). The SARIMA model is widely applied to predict seasonal time series data, such as to predict cases of tuberculosis (Mao et al., 2018), predict cases of malaria (Permanasari et al., 2013), predict the composition of iterations of coal Hardgrove grindability index (HGI) (Dindarloo et al., 2016), predict the number of covid-19 vaccines needed (Malki et al., 2022), predict consumer price index (Muthu et al., 2021), and others. SARIMA model is an accurate, precise, and suitable model to be applied in forecasting seasonal time series data (Bas et al., 2017; Dindarloo et al., 2016; Falatouri et al., 2022; Kumar Dubey et al., 2021; Malki et al., 2022; Mao et al., 2018; Muthu et al., 2021; Shen & Chen, 2017). SARIMA provides better forecasting results than other models (ArunKumar et al., 2021; He et al., 2021; Hu et al., 2007; H. Liu et al., 2020;

Xu et al., 2019). The predicted data using the SARIMA model has a small error value (Awang et al., 2022; Perone, 2022).

Based on the advantages that exist in the SARIMA method, in this study forecasting the number of passengers using the SARIMA method. Previous studies have used various methods to predict the number of passengers, such as ARIMA and ANFIS (Andalita & Irhamah, 2015), exponential smoothing (Oktaviarina, 2017), exponential smoothing event-based (Farida et al., 2021), Holt Winter's Exponential Smoothing (Baco et al., 2018; Sofiana et al., 2020), and Triple Exponential Smoothing (Darma et al., 2020; Fitria & Hartono, 2017). Research written by Nagara predicts the number of passengers using the SARIMA method and Winter's Exponential Smoothing (Negara, 2021). However, in Negara's research, forecasting is done on the total number of ship passengers, not the number of departing passengers and the number of arriving passengers. In previous studies, no one has predicted the number of ship passengers arriving and departing at the port using the SARIMA method.

This study aims to obtain the best model, obtain forecasting results and determine the accuracy of the forecasting results for the number of passengers arriving and departing at Semayang Port, Balikpapan with the SARIMA approach. With the models and methods that are able to predict the number of ship passengers, it is hoped that the marine transportation company can use the forecasting results as a reference and consideration for decision making to determine the best steps in dealing with the fluctuating number of ship passengers.

B. METHODS

The study used applied research with a quantitative approach. The data is in the form of time series data on the number of ship passengers arriving and departing from Semayang Port, Balikpapan from January 2006 to March 2020 with a monthly period. The data is a secondary data published on the official website of the Central Statistics Agency, www.bps.go.id. The forecasting carried out is the number of passengers at Semayang port coming and departing for the next 21 periods, starting from April 2020 to December 2021 using the SARIMA method. The forecasting steps using the SARIMA method are making a data plot, testing the stationarity of the data in the mean and variance, identifying the SARIMA model, parameter estimation, diagnostic checks, choosing the best model, predicting and determining the accuracy of forecasting results (Pangestu et al., 2020; Soelaeman, 2016).

C. RESULT AND DISCUSSION

1. Number of Passengers Departing at Semayang Port, Balikpapan City

a. Plot Data on The Number of Departing Passengers

The first step to predict time series data is to plot data from the original data on the number of passengers departing for the period January 2006 to March 2020 as many as 171 data. The data plot of the number of departing passengers can be seen in Figure 1.

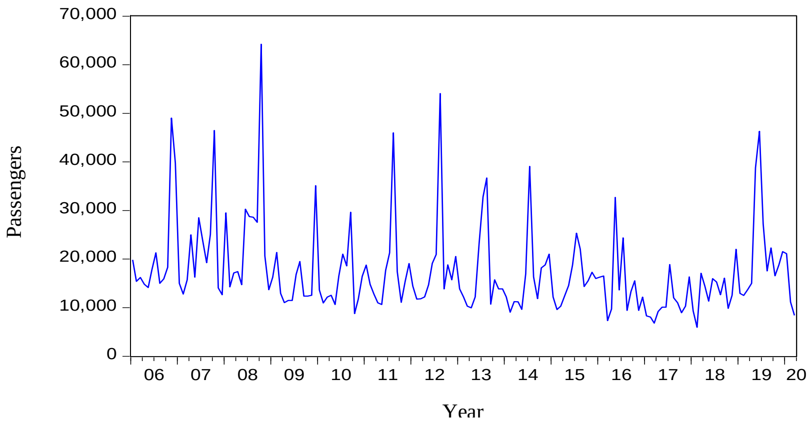


Figure 1. Plot of data on the number of departing passengers from January 2006 - March 2020

Figure 1 shows that visually the data on the number of departing passengers at Semayang Port, Balikpapan city forms a seasonal pattern because it experiences a drastic increase in certain months repeatedly with fixed time intervals. The average number of ship passengers departing from Semayang port is 17,578. Every year in certain months there is a surge in passengers departing from Semayang port. In 2006, the passenger surge occurred in November with 49,027 departing passengers. In 2007 and 2008, the passenger surge occurred in October with the number of passengers being 46,453 and 64,210 respectively. The passenger jumps in 2009 occurred in December with the number of passengers was 35,088. In September 2010 there were 29,640 passengers, which was the highest number of passengers in that year. In 2011, 2012, and 2013, a drastic increase in the number of passengers departing from Samayang port occurred in August with the number of passengers being 45,977, 54,068, and 36,688. The number of passengers was 39,084 and 25,325 in July 2014 and 2015 were the highest number of passengers. In that year. In May 2016, there was a surge in passengers with a total of 32,700 passengers. In 2017 there was no extreme increase in the number of passengers. The surge in passengers in 2018 occurred in December with the number of passengers as much as 22,027. In 2019, the extreme increase in the number of passengers occurred in June with the number of passengers 46,285. The highest increase in the number of passengers during January 2006 - March 2020 occurred in October 2008 while the lowest increase in passengers occurred in December 2018. In general, the increase in the number of passengers departing from Semayang port occurred during the Eid al-Fitr holiday.

b. Data Stationarity Test

The observation series is said to be stationary if the time series data from the past does not change due to changes in time and can be used to predict the future (Rosadi, 2012). Stationary data can be done using rounded values in the box-cox plot. If the result of rounded value = 1, then the data is stationary data in variance. Yet, if the result of the rounded value $\neq$ 1, the data shows that the data is not stationary in the variance. data that does not have a value of 1 must be transformed according to the box-cox transformation table.

The rounded value of the box-cox plot in Figure 2 is λ = 0.5. Based on the value of λ, it is known that the data on the number of departing passengers is data that is not stationary. Therefore, it needs to be transformed by $\frac{1}{\sqrt{Y_t}}$. The transformed data is shown in the box-cox plot of Figure 3. Based on that figure, it is obtained that the rounded value (λ)=1, then the data from the transformation of the number of departing passengers is stated to be stationary, as shown in Figure 2 and Figure 3.

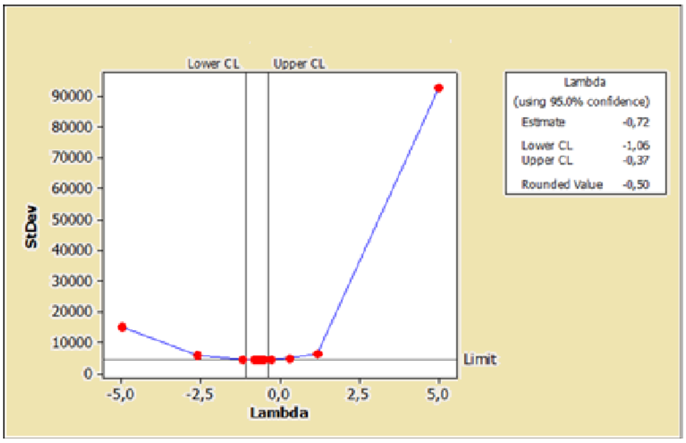


Figure 2. Box-cox Plot data on the number of departing passengers

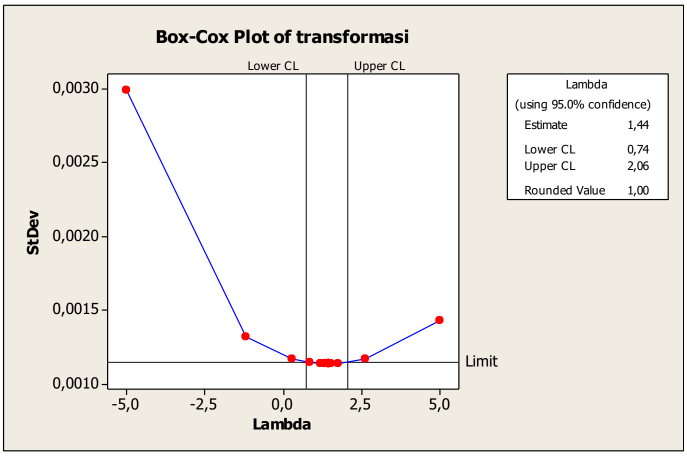


Figure 3. Box-cox plot data transformation of the number of departing passengers

When the stationarity of the data in the variance is met, then the data is evaluating for stationarity in the mean. To gain its stationary in the average, the ACF (Autocorrelation Function) correlogram and the ADF (Augmented Dickey Fuller) tests are used. The results of the ACF correlogram test are presented in Figure 4.

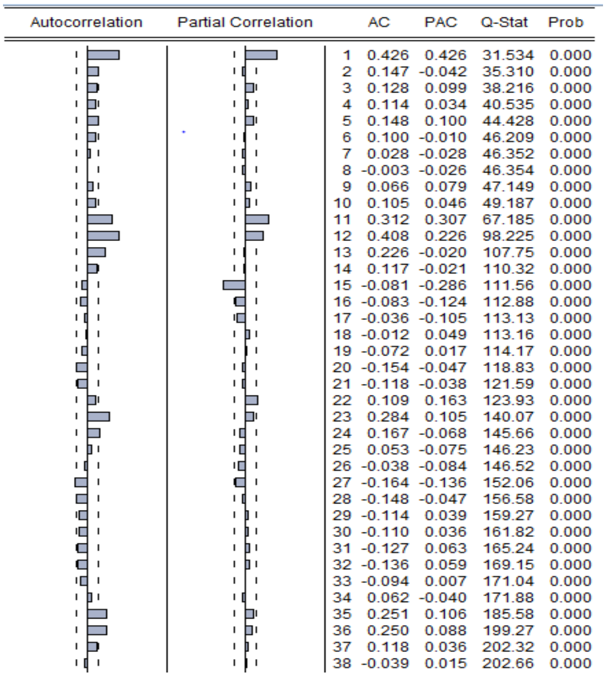


Figure 4. Correlogram of ACF and PACF data on the transformation of the number of departing passengers

Figure 4 of the ACF correlogram column shows that at lags 1, 2, 3, 4, 5, and 6 it decreases slowly, then at lag 12, lag 24 and lag 36 it has correlation values that exceed the significance line. This form of autocorrelation function states that the data has a seasonal pattern with a period of s, 2s, 3s. The transformation data on the number of departing passengers decreases slowly and contains a seasonal pattern which indicates that the data on the number of departing passengers is not stationary on average. Stationarity test of data with ACF correlogram is a visual test. The ADF (Augmented Dickey Fuller) test was carried out to ensure the stationarity of the data in the average. The following shows the results of the Augmented Dickey Fuller test, as shown in Table 1.

Table 1. Augmented Dickey Fuller test of transformed data

		t-Statistic	**Prob.***
ADF		0,018621	0,6872
Critical Value:	1%	-2,579587	
	5%	-1,942843	
	10%	-1,615376	

Based on table 1, the ADF value is 0.018621. This value is > compared to the critical value of the table with =5% which is -1.942843 which means that the transformation data of the number of departing passengers contains the unit root. This indicates that the data is not stationary in the mean. To gain the stationary data, non-seasonal differencing process is carried out with order d=1 and seasonal differencing with seasonal period = 12 (D=1). After differencing, the ADF value is -11.9688. This ADF value is < compared to the critical value of the table, which is -1.94291. This explains that the data is stationary in the mean.

c. Title Identification of the SARIMA model

After testing the stationary data, both variance and average stationaries, identification of orders in the forecasting model is carried out by looking at the results of the Autocorrelation Function (ACF) correlogram and Partial Autocorrelation Function (PACF) correlogram. The results of the ACF and PACF correlograms are presented in Figure 5.

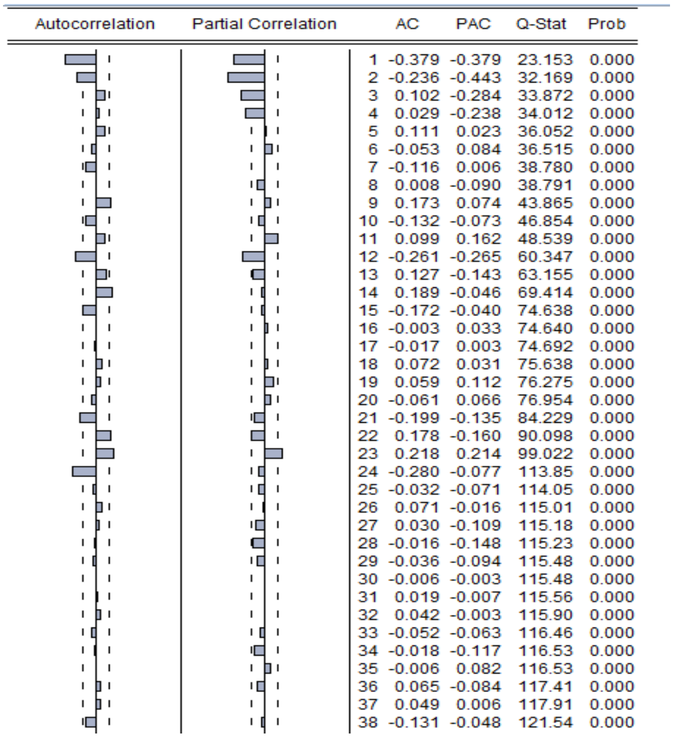


Figure 5. Correlogram of ACF and PACF data transformations and differences in the number of departing passengers

In Figure 5 the ACF correlogram column shows that after lag 2 there is a cut off, so that q = 2 is obtained to estimate the orde of the non-seasonal MA model. Then on the PACF correlogram it is known that after lag 4 there is a cut off. As a result, p = 4 is obtained to estimate the orde of the non-seasonal AR model. For the seasonal pattern on the ACF correlogram, Figure 5 shows that after lag 24 there is a cut off, as of Q = 2 is obtained to estimate the orde of the seasonal MA model. Next, for the seasonal pattern on the PACF correlogram Figure 5, it is known that after lag 12 there is a cut off, so that P = 1 is obtained to estimate the orde of the seasonal AR model. The tentative model SARIMA (p,d,q) $(P,D,Q)^s$ is gained by increasing or decreasing the seasonal and non-seasonal ordes respectively until 70 tentative models are gained.

d. Parameter estimation

A proper model that can be implemented to predict is a model that meets the parameter significance test. In other words, these parameters affect the model so that models with parameters that are not significant must be eliminated. A model is declared to meet the parameter significance test if all the parameters in the model have a probability value less than the value of α= 0.05. Of the 70 tentative models, there are 24 models with significant parameters.

e. Diagnostic check

Diagnostic checks are carried out to test whether a tentative model with significant parameters is feasible for the forecasting process. There are two conditions that must be met so that a model is feasible for forecasting. Those are if the residual is independent and the residual is a normal distribution. The following shows the results of the normality of the SARIMA $(4,1,0)(1,1,1)^{12}$ model, as shown in Figure 6.

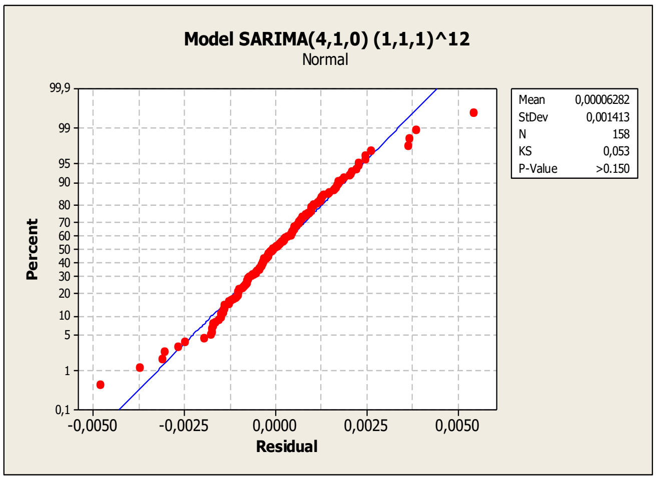
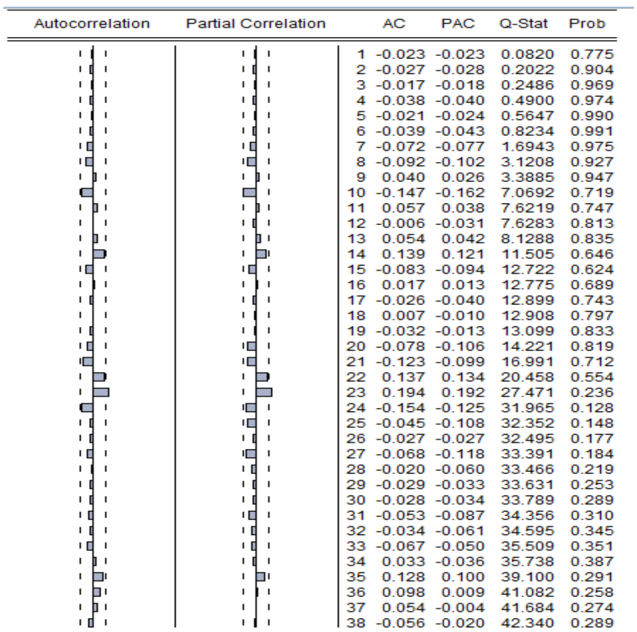


Figure 6. Normality of the residual probability plot of the SARIMA model$(4,1,0)(1,1,1)^{12}$

In Figure 6 it can be seen that the p_{value} is 0.15. Because the $p_{value} > \alpha$ (0.15>0.05) then the residual $SARIMA$ model $(4,1,0)(1,1,1)^{12}$ is normally distributed, as shown in Figure 7.



Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.023	-0.023	0.0820	0.775
		2	-0.027	-0.028	0.2022	0.904
		3	-0.017	-0.018	0.2486	0.969
		4	-0.038	-0.040	0.4900	0.974
		5	-0.021	-0.024	0.5647	0.990
		6	-0.039	-0.043	0.8234	0.991
		7	-0.072	-0.077	1.6943	0.975
		8	-0.092	-0.102	3.1208	0.927
		9	0.040	0.026	3.3885	0.947
		10	-0.147	-0.162	7.0692	0.719
		11	0.057	0.038	7.6219	0.747
		12	-0.006	-0.031	7.6283	0.813
		13	0.054	0.042	8.1288	0.835
		14	0.139	0.121	11.505	0.646
		15	-0.083	-0.094	12.722	0.624
		16	0.017	0.013	12.775	0.689
		17	-0.026	-0.040	12.899	0.743
		18	0.007	-0.010	12.908	0.797
		19	-0.032	-0.013	13.099	0.833
		20	-0.078	-0.106	14.221	0.819
		21	-0.123	-0.099	16.991	0.712
		22	0.137	0.134	20.458	0.554
		23	0.194	0.192	27.471	0.236
		24	-0.154	-0.125	31.965	0.128
		25	-0.045	-0.108	32.352	0.148
		26	-0.027	-0.027	32.495	0.177
		27	-0.068	-0.118	33.391	0.184
		28	-0.020	-0.060	33.466	0.219
		29	-0.029	-0.033	33.631	0.253
		30	-0.028	-0.034	33.789	0.289
		31	-0.053	-0.087	34.356	0.310
		32	-0.034	-0.061	34.595	0.345
		33	-0.067	-0.050	35.509	0.351
		34	0.033	-0.036	35.738	0.387
		35	0.128	0.100	39.100	0.291
		36	0.098	0.009	41.082	0.258
		37	0.054	-0.004	41.684	0.274
		38	-0.056	-0.020	42.340	0.289

Figure 7. Autocorrelation of residuals with Ljung Box test statistics on the model of $SARIMA(4,1,0)(1,1,1)^{12}$

The Ljung Box test is carried out to see the residual assumption meets the independent nature. The results of the Ljung Box test are as shown in Figure 7. Based on that figure, it can be seen that the probability value is greater than α = 5% ($p_{value} > 0{,}05$) in all lags, namely lag 1 to lag 38. It indicates that the residuals are not autocorrelated and the assumption of independent residuals are met in the $SARIMA(4{,}1{,}0)(1{,}1{,}1)^{12}$ model. Diagnostic checks were carried out on 24 models using significant parameters. It was so the results obtained in Table 2.

Table 2. Diagnostic examination results

Model	Residual	
	independent	**Normal**
$\mathbf{(4,1,0)(1,1,1)^{12}}$	✓	✓
$(4{,}1{,}0)(1{,}1{,}0)^{12}$	×	×
$\mathbf{(4,1,0)(0,1,2)^{12}}$	✓	✓
$(4{,}1{,}0)(0{,}1{,}1)^{12}$	✓	×
$(3{,}1{,}2)(1{,}1{,}0)^{12}$	×	✓
$(3{,}1{,}1)(1{,}1{,}1)^{12}$	×	✓
$(3{,}1{,}1)(1{,}1{,}0)^{12}$	×	✓
$(3{,}1{,}1)(0{,}1{,}1)^{12}$	×	✓
$(3{,}1{,}0)(1{,}1{,}1)^{12}$	×	✓
$(3{,}1{,}0)(1{,}1{,}0)^{12}$	×	✓
$(3{,}1{,}0)(0{,}1{,}2)^{12}$	×	✓
$(3{,}1{,}0)(0{,}1{,}1)^{12}$	×	×
$(2{,}1{,}1)(1{,}1{,}0)^{12}$	×	✓
$(2{,}1{,}1)(0{,}1{,}2)^{12}$	×	✓
$(2{,}1{,}0)(1{,}1{,}1)^{12}$	×	✓
$(2{,}1{,}0)(1{,}1{,}0)^{12}$	×	✓
$(2{,}1{,}0)(0{,}1{,}2)^{12}$	×	✓
$(2{,}1{,}0)(0{,}1{,}1)^{12}$	×	✓
$(1{,}1{,}0)(1{,}1{,}0)^{12}$	×	×
$(1{,}1{,}0)(0{,}1{,}2)^{12}$	×	×
$(1{,}1{,}0)(0{,}1{,}1)^{12}$	×	×
$(0{,}1{,}1)(1{,}1{,}0)^{12}$	×	✓
$(0{,}1{,}1)(0{,}1{,}2)^{12}$	×	✓
$(0{,}1{,}1)(0{,}1{,}1)^{12}$	×	×

Based on Table 2, two models are obtained that match the assumptions of the diagnostic examination, namely the model of $SARIMA\,(4{,}1{,}0)\,(1{,}1{,}1)^{12}$ and model of $SARIMA\,(4{,}1{,}0)(0{,}1{,}2)^{12}$.

f. Selection of the best model

The next step is to choose the best model from the two SARIMA models obtained by looking at the Akaike Information Criterion (AIC) and Schwartzt Bayesian Criterion (SBC) values. The model that has the smallest AIC and SBC values is declared the best model, as shown in Table 3.

Table 3. Comparison of AIC and SBC values

Model	AIC	SBC
$(4,1,0)(1,1,1)^{12}$	-2062.38	-2044
$(4,1,0)(0,1,2)^{12}$	-2065.38	-2047

Table 3 explains that the smallest AIC and SBC values are in model of $SARIMA(4,1,0)(0,1,2)^{12}$.

g. Forecasting with The Best Model

The Model of $SARIMA(4,1,0)(0,1,2)^{12}$ expressed as the best model, so that the general equation is gained as follows:

$$Y_t = (1+\phi_1)Y_{t-1} - (\phi_1-\phi_2)Y_{t-2} - (\phi_2-\phi_3)Y_{t-3} - (\phi_3-\phi_4)Y_{t-4} - \phi_4 Y_{t-5} + Y_{t-12} - (1+\phi_1)Y_{t-13} + (\phi_1-\phi_2)Y_{t-14} + (\phi_2-\phi_3)Y_{t-15} + (\phi_3-\phi_4)Y_{t-16} + \phi_4 Y_{t-17} + e_t - \Theta_1 e_{t-12} - \Theta_2 e_{t-24} \quad (1)$$

Forecasting for the next 21 periods starting from April 2020 to December 2021 is done by substituting the estimated value of the AR (autoregressive) parameter which is denoted by ϕ that is $\phi_1 = -0.6997$, $\phi_2 = -0.675, \phi_3 = -0.4184, \phi_4 = -0.2287$ and the estimated value of the SMA (Seasonal Moving Average) parameter which is denoted by Θ that is $\Theta_1 = 0.3788$ and $\Theta_2 = 0.2987$. Therefore the equation is obtained (2).

$$Y_t = 0.3003Y_{t-1} + 0.0245Y_{t-2} + 0.2568Y_{t-3} + 0.1897Y_{t-4} + 0.2287Y_{t-5} + Y_{t-12} - 0.3003Y_{t-13} - 0.0245Y_{t-14} - 0.2568Y_{t-15} - 0.1897Y_{t-16} - 0.2287Y_{t-17} + e_t - 0.3788e_{t-12} - 0.2987e_{t-24} \quad (2)$$

Equation (2) is used to forecast the next 21 periods. Then the results are transformed back to the original data scale $\frac{1}{{Y_t}^2}$. Forecasting results are as shown in Table 4.

Table 4. Results of forecasting the number of departing passengers in April 2020 - December 2021

t	Year	Month	Forecasting Result
172	2020	April	14389
173		May	27369
174		June	28571
175		July	24837
176		August	16747
177		September	18504
178		October	15802
179		November	16494
180		December	19263
181	2021	January	16639
182		February	10244
183		March	9526
184		April	13436
185		May	19116
186		June	21428
187		July	23042
188		August	16178
189		September	16612
190		October	14394
191		November	15066
192		December	19098

The following shows a comparison of the plot of the original data on the number of departing passengers and the data from the forecasting model of $SARIMA(4,1,0)(0,1,2)^{12}$ as the best model, as shown in Figure 8.

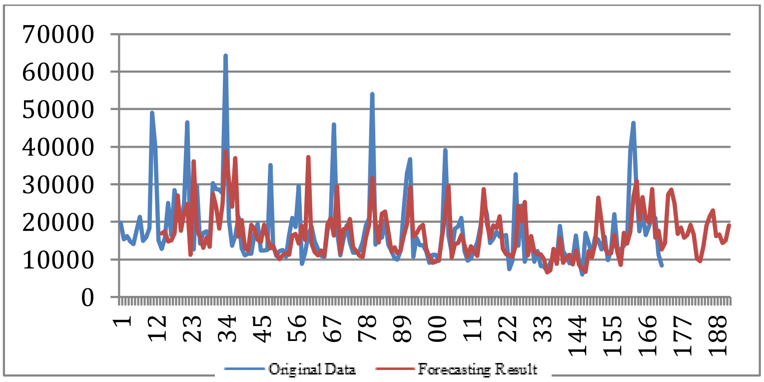


Figure 8. Plot of Original Data and Result of Forecasting Model of $SARIMA(4,1,0)(0,1,2)^{12}$

It can be seen in The Figure 8 that the forecasting results with the best model are close to the actual value in the original data.

h. Accuracy of Forecasting Results

A measure of the accuracy of forecasting results model of $SARIMA(4,1,0)(0,1,2)^{12}$ using *Mean Absolute Percentage Error* (MAPE) value. Forecasting using the best model, namely model of $SARIMA(4,1,0)(0,1,2)^{12}$ produces a MAPE value of 14.05%. Based on the criteria for the accuracy of the MAPE value of 14.05%, it interprets that the SARIMA model used produces good forecasts in predicting the number of departing passengers at Semayang Port, Balikpapan city. Forecasting results show that during the period April 2020 - December 2021, the highest number of passengers departing from Semayang port is in June 2020 with 28,571 passengers and in July 2021 with 23,042 passengers.

2. Number of Passengers Arrival at Semayang Port, Balikpapan City

a. Plot data on the number of arriving passengers

The following is a data plot on the number of passengers arriving from January 2006 - March 2020, as shown in Figure 9.

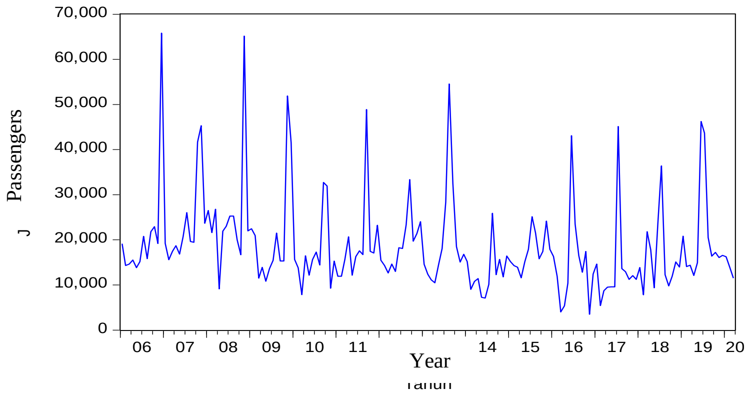


Figure 9. Plot of data on the number of passengers arriving from January 2006 - March 2020

The average number of ship passengers arriving at Semayang port is 18,715. Every year there is a drastic increase in the number of passengers in certain months. During the period from January 2006 to March 2020, the largest increase in the number of passengers occurred in December 2006 with a total of 65,817 passengers. The lowest increase in the number of passengers occurred in July 2015 with the number of passengers arriving at 25,133.

Figure 9 indicates visually that data on the number of ship passengers arriving at Semayang Harbor in Balikpapan City forms a seasonal pattern because it experiences a drastic increase in certain months repeatedly with fixed time intervals. The data does not fluctuate around a constant mean value and seems that the data variance is not constant in each observation. Therefore, visually the data does not meet the assumption of stationarity, both in the average and in the variance.

b. Data stationarity test

Stationary data can be done using rounded values in the box-cox plot. The results of the rounded value box-cox Plot of passenger data arriving as shown in Figure 10 and Figure 11.

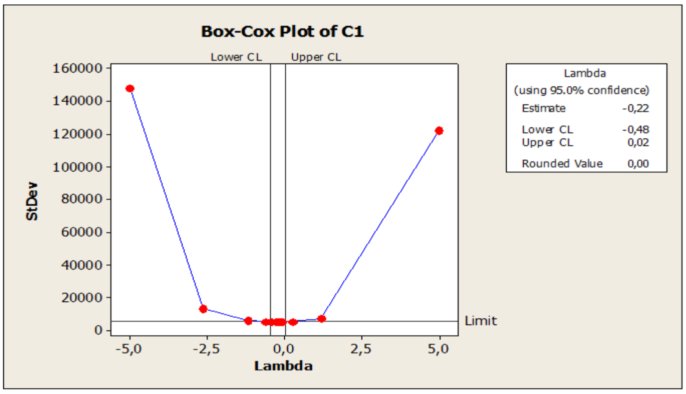


Figure 10. Box-cox Plot data on the number of arriving passengers

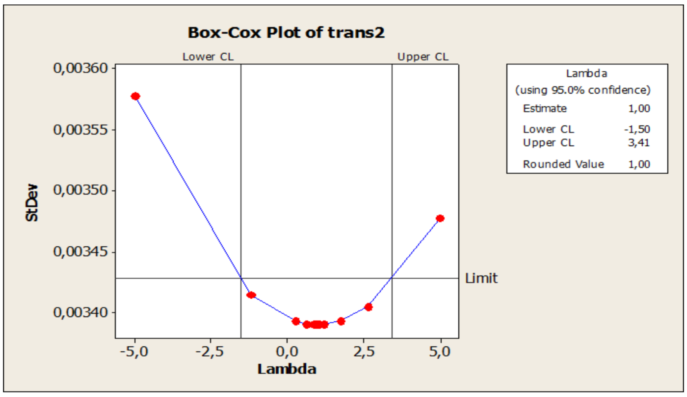


Figure 11. Box-cox plot of the data transformation of the two numbers of passengers arriving

Figure 10 shows the rounded value $(\lambda) = -0.00$. The value is ≠ 1 which indicates that the data on the number of arriving passengers is not stationary in the variance, then it is transformed into $\ln X_t$. Transformation of $\ln X_t$ results value of $\lambda = -1$. The λ value also still describes that the data is not stationary in the variance. Next, another transformation is carried out on $\frac{1}{X_t}$. The results are presented in Figure 11. Based on Figure 11 it is obtained that rounded value $(\lambda) = 1$. Then the data from the transformation of the number of arriving passengers is stated to be stationary in the variance. The data resulting from the Box-cox transformation that satisfies stationarity in variance is then tested for stationarity in average using the ACF and PACF Correlograms. The results are presented in Figure 12.

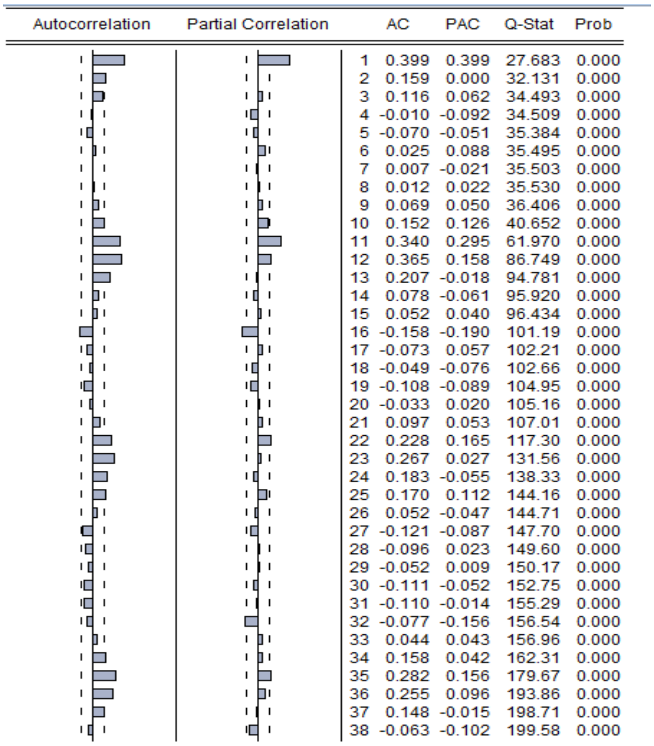


Figure 12. Correlogram of ACF and PACF data on the transformation of the number of arriving passengers

Figure 12 in the ACF (Autocorrelation Function) correlogram column shows that lag 12, lag 24 and lag 36 produce correlation values that exceed the significance line. This form of autocorrelation function states that the data has a seasonal pattern with a period of s, 2s, 3s. The data on the transformation of the number of arriving passengers contains a seasonal pattern which indicates that the data on the number of arriving passengers is not stationary on average. Stationarity test of data using ACF correlogram is a visual test. Then it was confirmed again using the ADF (Augmented Dickey Fuller) test to see the stationarity of the data in the average. The following shows the results of Augmented Dickey Fuller, as shown in Table 5.

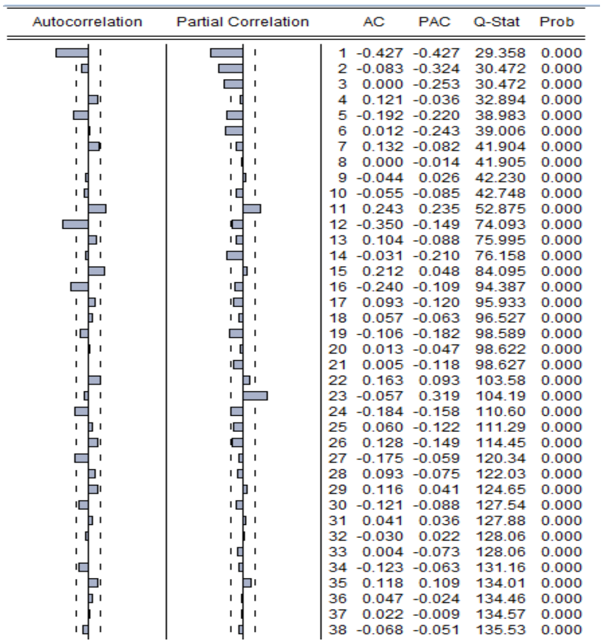
Table 5. Augmented Dickey Fuller test of transformed data

		t-Statistic	**Prob.***
ADF		0.167549	0.7335
Critical Value:	1%	-2.579587	
	5%	-1.942843	
	10%	-1.615376	

Based on Table 5, it is obtained that the ADF value > when compared to the critical value of the table with $\alpha = 5\%$ $(0.167594 > -1.942843)$ then the transformation data of the number of arriving passengers contains the unit root. The ADF value means that the data is not stationary in average. For so, it is necessary to carry out a differencing process, both non-seasonal and seasonal. Non-seasonal differentiation is carried out with the ordo of $d = 1$ dan *differencing* seasonal use seasonal period $= 12$ $(D = 1)$. After differencing, the value of ADF -10.53332 is obtained. ADF value is < compared to the critical value on the table that is -1.943012. This indicates that the data is stationary in the mean.

c. SARIMA Model Identification

SARIMA model identification is conducted by identifying the ordos of the model and looking at the correlogram ACF (Autocorrelation Function) and correlogram PACF (Partial Autocorrelation Function), as shown in Figure 13.



Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.427	-0.427	29.358	0.000
		2	-0.083	-0.324	30.472	0.000
		3	0.000	-0.253	30.472	0.000
		4	0.121	-0.036	32.894	0.000
		5	-0.192	-0.220	38.983	0.000
		6	0.012	-0.243	39.006	0.000
		7	0.132	-0.082	41.904	0.000
		8	0.000	-0.014	41.905	0.000
		9	-0.044	0.026	42.230	0.000
		10	-0.055	-0.085	42.748	0.000
		11	0.243	0.235	52.875	0.000
		12	-0.350	-0.149	74.093	0.000
		13	0.104	-0.088	75.995	0.000
		14	-0.031	-0.210	76.158	0.000
		15	0.212	0.048	84.095	0.000
		16	-0.240	-0.109	94.387	0.000
		17	0.093	-0.120	95.933	0.000
		18	0.057	-0.063	96.527	0.000
		19	-0.106	-0.182	98.589	0.000
		20	0.013	-0.047	98.622	0.000
		21	0.005	-0.118	98.627	0.000
		22	0.163	0.093	103.58	0.000
		23	-0.057	0.319	104.19	0.000
		24	-0.184	-0.158	110.60	0.000
		25	0.060	-0.122	111.29	0.000
		26	0.128	-0.149	114.45	0.000
		27	-0.175	-0.059	120.34	0.000
		28	0.093	-0.075	122.03	0.000
		29	0.116	0.041	124.65	0.000
		30	-0.121	-0.088	127.54	0.000
		31	0.041	0.036	127.88	0.000
		32	-0.030	0.022	128.06	0.000
		33	0.004	-0.073	128.06	0.000
		34	-0.123	-0.063	131.16	0.000
		35	0.118	0.109	134.01	0.000
		36	0.047	-0.024	134.46	0.000
		37	0.022	-0.009	134.57	0.000
		38	-0.068	-0.051	135.53	0.000

Figure 13. Correlogram of ACF and PACF data transformation and differencing of the number of arriving passengers

In Figure 13 the ACF data correlogram column shows that after *lag* 1 there is a *cut off*. That is so $q = 1$ is obtained to estimate the order of the non-seasonal MA model. After that, on the PACF correlogram it is known that after *lag* 3 *cut off* occurs. Therefore, $p = 3$ is gained to estimate that orde model AR is non seasonal. For the seasonal pattern on the ACF correlogram, Figure 13 shows that after *lag* 24 there is a *cut off* then it is predicted that orde model MA seasonal is $Q = 2$. Next, for the seasonal pattern on the PACF correlogram Figure 13, it is known that after *lag* 24 a *cut off* occurs and gets $P = 2$ to forecast that orde model AR is seasonal. SARIMA tentative model $(p, d, q)(P, D, Q)^S$ are obtained by increasing or decreasing each seasonal and non-seasonal order so that 56 tentative models are obtained.

d. Parameter estimation

Parameter significance test was conducted to select a suitable model for forecasting. In other words, these parameters affect the model so that models unsignificant parameters are eliminated. A model is declared to meet the significance test if all the parameters in the model have a probability < 0.05. From 56 tentative models, 17 models have significant parameters.

e. Diagnostic Checks

Diagnostic checks are carried out to test whether a tentative model with significant parameters is feasible for the forecasting process. There are two conditions that must be fulfilled so that a model is feasible for forecasting. Those are if the residual is independent and the residual is a normal distribution. The results of the normality test on the residuals of the SARIMA model is (0,1,1) $(2,1,0)^{12}$ as presented in Figure 14.

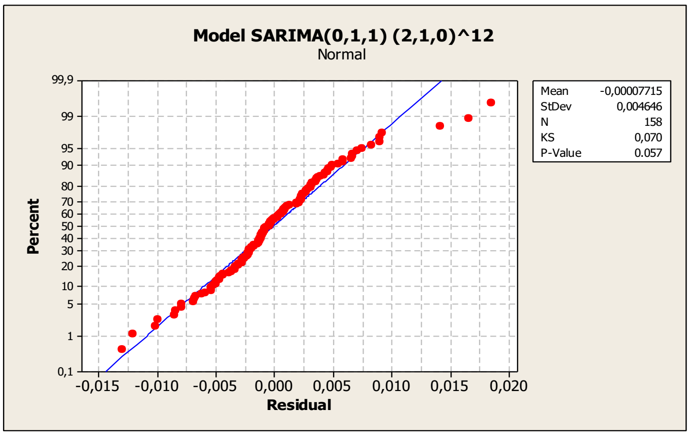
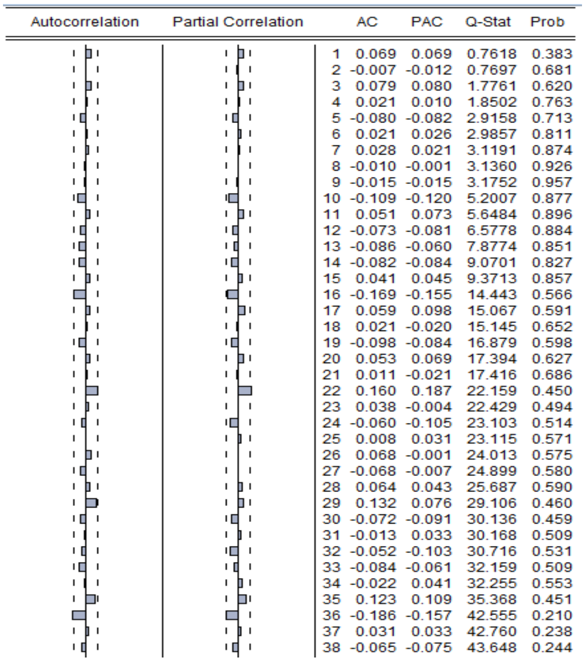


Figure 14. Normality of probability plot residual model of $SARIMA(0,1,1)(2,1,0)^{12}$

Figure 14 indicates p_{value} is 0,057. Because the value is $p_{value} > \alpha$ $(0.057 > 0.05)$ so the residual model of $SARIMA$ $(0,1,1)(2,1,0)^{12}$ normally distributed, as shown in Figure 15.



Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.069	0.069	0.7618	0.383
		2	-0.007	-0.012	0.7697	0.681
		3	0.079	0.080	1.7761	0.620
		4	0.021	0.010	1.8502	0.763
		5	-0.080	-0.082	2.9158	0.713
		6	0.021	0.026	2.9857	0.811
		7	0.028	0.021	3.1191	0.874
		8	-0.010	-0.001	3.1360	0.926
		9	-0.015	-0.015	3.1752	0.957
		10	-0.109	-0.120	5.2007	0.877
		11	0.051	0.073	5.6484	0.896
		12	-0.073	-0.081	6.5778	0.884
		13	-0.086	-0.060	7.8774	0.851
		14	-0.082	-0.084	9.0701	0.827
		15	0.041	0.045	9.3713	0.857
		16	-0.169	-0.155	14.443	0.566
		17	0.059	0.098	15.067	0.591
		18	0.021	-0.020	15.145	0.652
		19	-0.098	-0.084	16.879	0.598
		20	0.053	0.069	17.394	0.627
		21	0.011	-0.021	17.416	0.686
		22	0.160	0.187	22.159	0.450
		23	0.038	-0.004	22.429	0.494
		24	-0.060	-0.105	23.103	0.514
		25	0.008	0.031	23.115	0.571
		26	0.068	-0.001	24.013	0.575
		27	-0.068	-0.007	24.899	0.580
		28	0.064	0.043	25.687	0.590
		29	0.132	0.076	29.106	0.460
		30	-0.072	-0.091	30.136	0.459
		31	-0.013	0.033	30.168	0.509
		32	-0.052	-0.103	30.716	0.531
		33	-0.084	-0.061	32.159	0.509
		34	-0.022	0.041	32.255	0.553
		35	0.123	0.109	35.368	0.451
		36	-0.186	-0.157	42.555	0.210
		37	0.031	0.033	42.760	0.238
		38	-0.065	-0.075	43.648	0.244

Figure 15. Autocorrelation of residuals with test statistics $Ljung\ Box$ in the model of $SARIMA(0,1,1)(2,1,0)^{12}$

The residual assumption test is independent or not shown in Figure 15. Based on the figure, the probability value is greater than $\alpha = 5\%$ in all lag , those are lag 1 up to lag 38. It means that the residuals are not autocorrelated and the assumption of independence of the residuals is met at model of $SARIMA(0,1,1)(2,1,0)^{12}$. Diagnostic examinations were carried out on 17 models with significant parameters. The results of the diagnostic checks on the 17 models are shown in Table 6.

Table 6. Diagnostic Check Results

Model	Residual	
	Independent	**normal**
$(3,1,0)(2,1,0)^{12}$	×	✓
$(3,1,0)(1,1,0)^{12}$	×	✓
$(3,1,0)(0,1,2)^{12}$	×	✓
$(3,1,0)(0,1,1)^{12}$	×	✓
$(2,1,1)(2,1,0)^{12}$	×	×
$(2,1,1)(1,1,0)^{12}$	×	×
$(2,1,0)(2,1,0)^{12}$	×	✓
$(2,1,0)(1,1,0)^{12}$	×	✓
$(2,1,0)(0,1,1)^{12}$	×	✓
$(1,1,1)(2,1,1)^{12}$	×	×
$(1,1,0)(2,1,0)^{12}$	×	×
$(1,1,0)(1,1,0)^{12}$	×	×
$(1,1,0)(0,1,2)^{12}$	×	✓
$(1,1,0)(0,1,1)^{12}$	×	×
$(0,1,1)(2,1,0)^{12}$	✓	✓
$(0,1,1)(1,1,0)^{12}$	×	×
$(0,1,1)(0,1,1)^{12}$	✓	×

Based on Table 6, there is only one model that satisfies the assumption of a diagnostic examination, namely the model of $SARIMA(0,1,1)(2,1,0)^{12}$. Consequently, the model is model the best model that forecasts data on the number of passengers arriving at Semayang port.

f. Forecasting with the best model

The model of $SARIMA(0,1,1)(2,1,0)^{12}$ expressed as the best model, so that the general equation is obtained in equation (3):

$$X_t = X_{t-1} + (1 + \Phi_1)X_{t-12} - (1 + \Phi_1)X_{t-13} - (\Phi_1 - \Phi_2)X_{t-24} - (-\Phi_1 + \Phi_2)X_{t-25} - \Phi_2 X_{t-36} + \Phi_2 X_{t-37} + e_t - \theta_1 e_{t-1} \quad (3)$$

Forecasting for the next 21 periods starting from April 2020 to December 2021 is conducted by substituting the estimated value of the SAR (Seasonal Autoregressive) parameter which is denoted by Φ viz $\Phi_1 = -0.5780$ and $\Phi_2 = -0.4628$ then the estimated value of the MA (Moving Average) parameter which is denoted by θ that is $\theta_1 = 0.8829$, until the equation (4).

$$X_t = X_{t-1} + 0.422X_{t-12} - 0.422X_{t-13} + 0.1152X_{t-24} - 0.1152X_{t-25} + 0.4628X_{t-36} - 0.4628X_{t-37} + e_t - 0.8829e_{t-1} \quad (4)$$

Equation (4) is used for forecasting in the next 21 periods. Then the results are transformed back to the original data scale, as shown in Table 7.

Table 7. Results of forecasting the number of arriving passengers from April 2020 - December 2021

t	Year	Month	Forecasting Result
172	2020	April	14595
173		May	14876
174		June	26125
175		July	61037
176		August	21050
177		September	18108
178		October	17658
179		November	18311
180		December	17731
181	2021	January	20462
182		February	13652
183		March	20787
184		April	20639
185		May	15543
186		June	36088
187		July	64911
188		August	21484
189		September	17472
190		October	19241
191		November	21834
192		December	20798

Comparison of the plot of the original data on the number of arriving passengers and the data from the forecasting results on model of $SARIMA(0,1,1)(2,1,0)^{12}$ as the best model is presented in Figure 16.

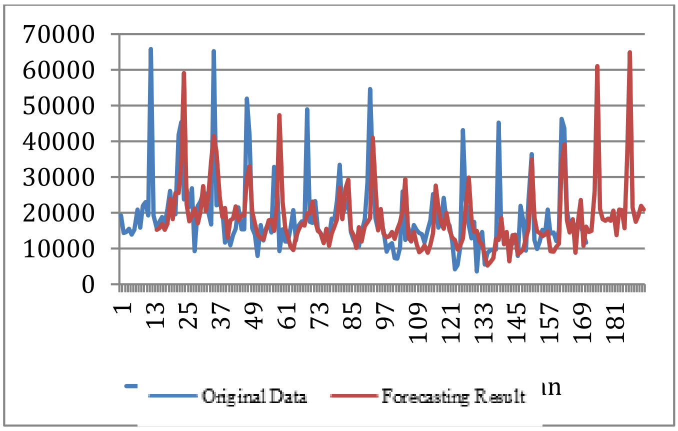


Figure 16. Plot of Original Data and Results of Forecasting Model of $SARIMA(0,1,1)(2,1,0)^{12}$

The figure indicates that the forecasting results with the best model are close to the actual values in the original data.

g. Accuracy of forecasting results

A measure of the accuracy of forecasting results with model of $SARIMA(0,1,1)(2,1,0)^{12}$ using the MAPE (Mean Absolute Percentage Error) value. Forecasting using the best model, namely model of $SARIMA(0,1,1)(2,1,0)^{12}$ results to MAPE value as 3.27%. This value means that the SARIMA model used delivers good forecasts in predicting the number of passengers arriving at Semayang Port, Balikpapan City. Forecasting results show that during the period April 2020 - December 2021, the highest number of passengers arriving at Semayang port was in July 2020 and July 2021 with 61,037 and 64,911 passengers.

The results showed that the MAPE value for forecasting the number of departing ship passengers was 14.05% and 3.27% for forecasting the number of arriving ship passengers. This error value is small so that the SARIMA method is the right method for forecasting seasonal data. This is as stated by Awang et al. and Perone. Awang et al. (2022) dan Perone (2022)stated that SARIMA is a time series method that has a small error.

In previous studies, the SARIMA method was used to predict the number of train passengers (Milenković et al., 2018), the number of airplane passengers (Li et al., 2017), the number of ship passengers (Negara, 2021). The results of this study are in accordance with three previous studies which state that the SARIMA method is the right method to be used on seasonal time series data. The results of this study are also in accordance with the research of Falatouri et al. (2022), Kumar Dubey et al. (2021), Malki et al. (2022), and Muthu et al. (2021) which states that the SARIMA method is an accurate, precise, and suitable model to be applied in seasonal forecasting. This study provides results that have never been done in previous research, namely forecasting the number of ship passengers departing and arriving at Semayang Harbor. Separate forecasting between the number of departing and arriving ship passengers will provide clear data so that the competent authorities can make the right decisions. The government or port manager can predict how many ships need to be added in certain months to carry passengers arriving and departing. The government or port manager can provide good and efficient service so that there is no accumulation of passengers and can anticipate a surge in passengers departing and arriving at Semayang Port.

D. CONCLUSION AND SUGGESTIONS

Based on the results of data analysis, it can be concluded that the best forecasting model that appropriates to predict the number of departing passengers is the model of $SARIMA(4,1,0)(0,1,2)^{12}$ and MAPE value gained is as much as 14.05%. It means that the SARIMA model used produces good forecasts. Forecasting results show that during the period April 2020 - December 2021, the highest number of passengers departing from Semayang port is in June 2020 and July 2021. The best forecasting model that can be applied to predict the number of arriving passengers is the model of $SARIMA(0,1,1)(2,1,0)^{12}$ with the resulting MAPE value of 3.27%. The interpretation of the MAPE value is the SARIMA model which is used to produce accurate forecasts. The highest number of passengers arriving at Semayang port based on forecasting results during the April 2020 - December 2021 period occurred in July 2020 and July 2021. The results of this forecast can be used as a reference for the government or port managers to anticipate a surge in passengers. The government or port management can prepare an adequate amount of transportation in certain months to avoid the accumulation of passengers and to make sea transportation more efficient.

REFERENCES

Andalita, I., & Irhamah. (2015). Peramalan Jumlah Penumpang Kereta Api Kelas Ekonomi Kertajaya Menggunakan ARIMA dan ANFIS. *Jurnal Sains Dan Seni ITS*, *4*(2), 311–316.

Andini, T. D., & Sunyoto, R. M. (2018). Sistem Peramalan Jumlah Penumpang Kapal Laut di Pelabuhan Tanjung Perak Surabaya Menggunakan Triple Eksponensial Smoothing Berbasis Android. *Jurnal Sistem Dan Teknologi Informasi*, *4*(2), 113–124.

ArunKumar, K. E., Kalaga, D. V., Sai Kumar, C. M., Chilkoor, G., Kawaji, M., & Brenza, T. M. (2021). Forecasting the dynamics of cumulative COVID-19 cases (confirmed, recovered and deaths) for top-16 countries using statistical machine learning models: Auto-Regressive Integrated Moving Average (ARIMA) and Seasonal Auto-Regressive Integrated Moving Averag. *Applied Soft Computing*, *103*(December 2019), 107161. https://doi.org/10.1016/j.asoc.2021.107161

Awang, N., Samy, G. A. N., Hassan, N. H., Maarop, N., & Perumal, S. (2022). Implementation of SARIMA Algorithm in Understanding Cybersecurity Threats in University Network. *Turkish Online Journal of Qualitative Inquiry (TOJQI)*, *6*(3), 8442–8451.

Baco, E., Lamusa, F., & Nurfadhilah, K. (2018). Peramalan Jumlah Penumpang Pada PT. Angkasa Pura I (Persero) Kantor Cabang Bandar Udara Internasional Sultan Hasanuddin Makassar. *Jurnal Matematika Dan Statistika Serta Aplikasinya*, *6*(24–29).

Bas, M. del C., Ortiz, J., Ballesteros, L., & Martorell, S. (2017). Evaluation of a multiple linear regression model and SARIMA model in forecasting 7Be air concentrations. *Chemosphere*, *177*, 326–333. https://doi.org/10.1016/j.chemosphere.2017.03.029

BPS. (2019). *Kota Balikpapan dalam Angka 2019*. Badan Pusat Statistik (BPS) Kota Balikpapan.

Chatfield, C. (2000). *Time Series Forecasting*. Chapman & Hall/CRC.

Darma, I. W. A. S., Gunawan, I. P. E. G., & Sutramiani, N. P. (2020). Peramalan Jumlah Kunjungan Wisatawan Menggunakan Triple Exponential Smoothing. *Jurnal Ilmiah Merpati*, *8*(3), 2011–2021.

Dindarloo, S., Hower, J. C., Trimble, A. S., Bagherieh, A., & Trimble, A. S. (2016). Fundamental evaluation of petrographic effects on coal grindability by seasonal autoregressive integrated moving average (SARIMA). *International Journal of Mineral Processing*, *154*, 94–99. https://doi.org/10.1016/j.minpro.2016.07.005

Falatouri, T., Darbanian, F., Brandtner, P., & Udokwu, C. (2022). Predictive Analytics for Demand Forecasting - A Comparison of SARIMA and LSTM in Retail SCM. *Procedia Computer Science*, *200*(2019), 993–1003. https://doi.org/10.1016/j.procs.2022.01.298

Farida, Y., Yusi, S., & Yuliati, D. (2021). Peramalan Jumlah Penumpang Pesawat di Bandar Udara Internasional Juanda Menggunakan Metode Exponential Smoothing Event-Based. *Barekeng Jurnal Ilmu Matematika Dan Terapan*, *15*(4).

Fatimah, S. (2019). *Pengantar Transportasi*. Myria Publisher.

Fitria, V. A., & Hartono, R. (2017). Peramalan Jumlah Penumpang Pada Siluet Tour And Travel Kota Malang Menggunakan Metode Triple Exponential Smoothing. *Jurnal Ilmiah Teknologi Informasi Asia*, *11*(1), 15–20.

He, K., Ji, L., Wu, C. W. D., & Tso, K. F. G. (2021). Using SARIMA–CNN–LSTM approach to forecast daily tourism demand. *Journal of Hospitality and Tourism Management*, *49*(August), 25–33. https://doi.org/10.1016/j.jhtm.2021.08.022

Herjanto, E. (2007). *Manajemen Operasi. Edisi Ketiga*. Grasindo.

Hu, W., Tong, S., Mengersen, K., & Connell, D. (2007). Weather Variability and the Incidence of Cryptosporidiosis: Comparison of Time Series Poisson Regression and SARIMA Models. *Annals of Epidemiology*, *17*(9), 679–688. https://doi.org/10.1016/j.annepidem.2007.03.020

Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting Principles and Practice* (Second). Otexts.

Kokilavani, S., & dkk. (2020). Sarima Modelling and Forecasting of Monthly Rainfall Pattern for Coimbatore, Tamil Nadu, India. *Journal of Applied Science and Technology*, *39*(8), 69–76.

Kumar Dubey, A., Kumar, A., García-Díaz, V., Kumar Sharma, A., & Kanhaiya, K. (2021). Study and analysis of SARIMA and LSTM in forecasting time series data. *Sustainable Energy Technologies and Assessments*, *47*(May), 101474. https://doi.org/10.1016/j.seta.2021.101474

Liu, H., Li, C., Shao, Y., Zhang, X., Zhai, Z., Wang, X., Qi, X., Wang, J., Hao, Y., Wu, Q., & Jiao, M. (2020). Forecast of the trend in incidence of acute hemorrhagic conjunctivitis in China from 2011–2019 using the Seasonal Autoregressive Integrated Moving Average (SARIMA) and Exponential Smoothing (ETS) models. *Journal of Infection and Public Health*, *13*(2), 287–294. https://doi.org/10.1016/j.jiph.2019.12.008

Liu, J., Zhao, Z., Zhong, Y., Zhao, C., & Zhang, G. (2022). Prediction of the dissolved gas concentration in power transformer oil based on SARIMA model. *Energy Reports*, *8*, 1360–1367. https://doi.org/10.1016/J.EGYR.2022.03.020

Makridakris, S., Wheelwright., S. C., & Hynmand, R. J. (1997). *Forecaseting Method and Application* (Third). John Wiley & Sons. Inc.

Malki, A., Atlam, E.-S., Hassanien, A. E., Ewis, A., Dagnew, G., & Gad, I. (2022). SARIMA model-based forecasting required number of COVID-19 vaccines globally and empirical analysis of peoples' view towards the vaccines. *Alexandria Engineering Journal*, *61*(12), 12091–12110. https://doi.org/10.1016/j.aej.2022.05.051

Mao, Q., Zhang, K., Yan, W., & Cheng, C. (2018). Forecasting the incidence of tuberculosis in China using the seasonal auto-regressive integrated moving average (SARIMA) model. *Journal of Infection and Public Health*, *11*(5), 707–712. https://doi.org/10.1016/j.jiph.2018.04.009

Montgomery, D. C., Jennings, C. L., & Kulahci, M. (2015). *Introduction to Time Series Analysis and Forecasting* (Second). John Wiley & Sons, Inc.

Muthu, N. S., Kannan, K. S., Deneshkumar, V., & Thangasamy, P. (2021). SARIMA Model for Forecasting Price Indices Fluctuations. *European Journal of Mathematics and Statistics*, *2*(6), 1–6. https://doi.org/10.24018/ejmath.2021.2.6.67

Nasution, M. N. (2015). *Manajemen Transportasi. Edisi keempat*. Ghalia Indonesia.

Negara, R. I. P. (2021). Peramalan Jumlah Penumpang Kapal di Pelabuhan Pantai Baru dengan Metode Sarima dan Winter's Exponential Smoothing. *Jurnal Statistika Terapan*, *1*(1).

Oktaviarina, A. (2017). Peramalan Jumlah Penumpang Kereta Api di Indonesia Menggunakan Metode Exponential Smoothing. *Buana Matematika Jurnal Ilmiah Matematika Dan Pendidikan Matematika*, *7*(2).

Pangestu, C. J., Piantari, E., & Munir. (2020). Prediction of Diarrhea Sufferers in Bandung with Seasonal Autoregressive Integrated Moving Average (SARIMA). *Journal of Computers for Society*, *1*(1), 61–79.

Permanasari, A. E., Hidayah, I., & Bustoni, I. A. (2013). SARIMA (Seasonal ARIMA) implementation on time series to forecast the number of Malaria incidence. *Proceedings - 2013 International*

Conference on Information Technology and Electrical Engineering: "Intelligent and Green Technologies for Sustainable Development", ICITEE 2013, *2*, 203–207. https://doi.org/10.1109/ICITEED.2013.6676239

Perone, G. (2022). Using the SARIMA Model to Forecast the Fourth Global Wave of Cumulative Deaths from COVID-19: Evidence from 12 Hard-Hit Big Countries. *Econometrics*, *10*(2). https://doi.org/10.3390/econometrics10020018

Prabhadika, I. P. Y., Tastrawati, N. K. T., & Harini, L. P. I. (2018). Peramalan Persediaan Infus Menggunakan Metode Autoregressive Integrated Moving Average (ARIMA) pada Rumah Sakit Umum Pusat Sanglah. *Jurnal Matematika*, *7*(2), 129–133.

Putra, A. P., Sitompul, M. K., Mardalena, T., & Romadani. (2022). Analisis Keterkaitan Transportasi Laut dengan Pelabuhan Kargo Teluk Dalam untuk Menunjang Kegiatan Perekonomian Masyarakat Penyalai Kecamatan Kuala Kampar Provinsi Riau. *Jurnal Jalasena*, *3*(2).

Rosadi, D. (2012). *Ekonometrika dan Analisis Runtun Waktu Terapan dengan Eviews*. CV Andi Offset.

Salim, A. A. (2016). *Manajemen Transportasi*. PT. Raja Grafindo Persada.

Shen, S., & Chen, S. (2017). Application of <i>SARIMA</i> Model on Money Supply. *Open Journal of Statistics*, *07*(01), 112–121. https://doi.org/10.4236/ojs.2017.71009

Soelaeman, I. (2016). *Analisis Runtun Waktu*. Universitas Terbuka.

Sofiana, Suparti, Hakim, A. R., & Triutami, I. (2020). Peramalan Jumlah Penumpang Pesawat di Bandara Internasional Ahmad Yani dengan Metode Holt Winter's Exponential Smoothing dan Metode Exponential Smoothing Event Based. *JURNAL GAUSSIAN*, *9*(4), 535–545.

Tadesse, K. B., & Dinka, M. O. (2017). Application of SARIMA model to forecasting monthly flows in Waterval River South Africa. *Journal of Water and Land Development*, *35*(X–XII), 229–236.

Xu, S., Chan, H. K., & Zhang, T. (2019). Forecasting the demand of the aviation industry using hybrid time series SARIMA-SVR approach. *Transportation Research Part E: Logistics and Transportation Review*, *122*(December 2018), 169–180. https://doi.org/10.1016/j.tre.2018.12.005

Yusof, F., & Kane, I. L. (2012). Modeling monthly rainfall time series using ETS state space and SARIMA models. *International Journal of Current Research*, *4*(09), 1–6.