J. Pijar MIPA, Vol. 18 No. 4, July 2023: 632-637
DOI: 10.29303/jpm.v18i4.5469 ISSN 1907-1744 (Print)
ISSN 2460-1500 (Online)

COMPARISON OF THE ACCURACY OF SEASONAL DATA PREDICTION VALUES USING SARIMA AND WINTER EXPONENTIAL SMOOTHING ON THE NUMBER OF SHIP PASSENGERS IN BATAM CITY

Widya Reza
Department of Mathematics, Faculty of Information Technology, Institut Teknologi Batam, Kepulauan Riau, Indonesia
Email: Widya@iteba.ac.id

Received: June 21, 2023. Accepted: July 21, 2023. Published: July 31, 2023

Abstract: Sea transportation is one of the keys to the success of the tourism industry in a region. One of the problems faced in sea transportation is the increase in the number of passengers that needs to be balanced with the number of ships available. An accurate prediction model is very important in forecasting the number of passengers on a ship to prepare the number of ships according to the number of passengers. This study aims to compare the accuracy of prediction models on the number of ship passengers in Batam City, whose data fluctuates and contains seasonal patterns. This study used secondary data from the number of ship passengers in Batam City from January 2014 to December 2022. The methods used are SARIMA and Winter Exponential Smoothing. The results of this study show that the selected SARIMA model in forecasting the number of passengers in Batam City is SARIMA $(0,1,1)(0,1,1)^{12}$ with an MSD value of 0.0000021. In contrast, the selected Winter Exponential Smoothing model is Winter Exponential Smoothing at $\alpha = 0.9$, $\beta = 0.1$, and $\delta = 0.1$ with an MSD value of 0.00000. Based on these two models, the most accurate prediction model with the lowest MSD value in forecasting the number of ship passengers in Batam City is the Winter Exponential Smoothing model with values $\alpha = 0.9$, $\beta = 0.1$, and $\delta = 0.1$.

Keywords: *SARIMA, Winter Exponential Smoothing, Forecasting*

INTRODUCTION

The sea transportation system is very important in a region especially for provinces consisting of many islands [1]. Based on BPS data, Riau Islands Province always ranks first in the highest number of domestic ship passengers in Indonesia in 2019-2021. Based on BPS Batam data in 2021, the number of foreign cruise passengers entering Batam reached 20933 and those leaving were 6200 passengers. The number of domestic cruise passengers who came was 614892 passengers while the outgoing passengers reached 57588. Based on BPS data, the number of ship passengers in Batam City from 2017-2021 experienced a fluctuating pattern where the number of passengers increased from 2017-2019 but decreased in 2020 and 2021. In addition to experiencing fluctuating patterns, the number of ship passengers in Batam city also contains seasonal patterns so predictions model need to be made. Predicting the number of passengers is an important element in transportation services as the key to the success of transportation companies [2].

Prediction of the number of ship passengers has been done with various forecasting methods such as Autoregressive Integrated Moving Average (ARIMA), Seasonal Autoregressive Integrated Moving Average (SARIMA) and Exponential Smoothing [3–6]. ARIMA, SARIMA and Exponential Smoothing models can predict based on past data patterns containing seasonal patterns [7]. The SARIMA method is one of the modeling approaches in forecasting the tourism industry. Several studies have applied this method [8–10]. Forecasting the visits of foreign tourists entering Bandung airport from 2010 to 2017 using the SARIMA model and found that the model can be the best model to predict foreign tourist arrivals with good accuracy [8]. Forecasting international tourist arrivals to Sri Lanka used the SARIMA approach and found that the SARIMA model matched tourism data with the lowest MAPE values [9]. The SARIMA model was used in estimating monthly tourist arrivals from ASEAN countries from January 2000 to December 2014 [10]. Forecasting international tourism demand in Malaysia using the SARIMA model and concluded smoothing techniques should be incorporated to forecast tourism demand [11].

The winter exponential smoothing model has been used by many researchers for forecasting [10–12]. Forecasting tourist arrivals in Zambia using the Holt-Winter exponential model and compared to the ARIMA Model [13]. Hasil penelitian menunjukkan bahwa model Holt-Winter memberikan kesalahan terkecil. Perbandingan Holt-Winter dan ARIMA juga digunakan untuk meramalkan kedatangan wisatawan di India untuk periode 1981 hingga 2014 [11].

Unlike the previous study, this study will compare the accuracy of prediction values using the SARIMA and Winter Exponential Smoothing methods against the number of ship passengers in Batam City which contains seasonal patterns and trends to get the most accurate prediction results. Seasonal patterns can occur in one year, one month, one week, or in one day [9–14]. Several studies have been conducted to predict. The results of these predictions can be used by the government and ship managers to determine policies and the number of ships that must be prepared so as not to suffer losses if the number of passengers and goods loaded does not meet the minimum operational costs of ships [15].

RESEARCH METHODS

This research is a quantitative research using time series data. The data used in this study is secondary data in the form of monthly ship passenger data obtained from the website of the Badan Pusat Statistik (BPS) Batam City, namely the period January 2017 to September 2021 consisting of 57 observations. The data used is in the form of data on the number of ship passengers for the arrival and departure of domestic.

The data analysis methods used in this study were SARIMA and winter exponential smoothing. The seasonal index can be calculated using a simple averaging method.

$$\text{Seasonal index} = \left(\frac{\overline{X}_i}{\overline{X}_j} x100\%\right) x12$$

$\overline{X}_i$ is the average of the first month of each year, $\overline{X}_j$ is the average of each month of the year.

Seasonal Autoregressive Integrated Moving Average (SARIMA)

SARIMA is time series data that shows repeated seasonal-periodic fluctuations with intensity each year [19]. Thus, the characteristics of seasonal time series are strong relationships at seasonal intervals (seasonal period). The general equation of SARIMA (p,d,q)(P,Q,S) is: [19–22].

$$X_t(1-B)^d(1-B^S) = (1-B\Phi_1)(1-\theta_1 B^S)e_t$$

$(1-B)^d$ is a non-seasonal difference, $(1-B^S)$ is seasonal difference, $\theta_1(B)$ a non-seasonal MA, B^S is a seasonal MA, e_t is a residual term.

Winter Exponential Smoothing

Winter Exponential Smoothing is a method used when data contains seasonal patterns. This method consists of three parameters, namely alpha (α), gamma (γ) and beta (β) whose value is between 0 and 1 with $0 < \alpha, \gamma, \beta < 1$. The equation of the Winter Exponential Smoothing Method is: [20–22]

Over all smoothing: $S_t = \alpha \frac{x_t}{l_{t-L}} + (1-\alpha)(S_{t-1} + b_{t-1}), 0 < \alpha < 1$

Trend smoothing : $b_t = \gamma(S_t - S_{t-1}) + (1-\gamma)b_{t-1}, 0 < \gamma < 1$

Seasonal smoothing : $l_t = \beta \frac{x_t}{S_t} + (1-\beta)l_{t-L}, \ 0 < \beta < 1$

Forecasting equation : $F_{t+m} = (S_t + b_t m)l_{t-L+m}$

And the trend $\beta = \frac{1}{L}\left[\left(\frac{x_{L+1}-x_1}{L}\right) + \left(\frac{x_{L+2}-x_2}{L}\right) + \cdots + \left(\frac{x_{L+L}-x_L}{L}\right)\right]$

L represents the length of the season, β is the component of the trend parameter, l is the seasonal adjustment factor and F_{t+m} is the prediction of m for the future period.

Accuracy is important for choosing a specific measure of the forecasting method. Some of the measurement methods used are Mean Absolute Deviation (MAD) and Mean Square Error (MSE). MAD is the overall forecasting error value for the model. The MAD value is calculated by taking the sum of the absolute values of the forecasting error divided by the number of data periods. Another measure of accuracy used is MSD, this is widely used in forecasting studies [15–18].

RESULTS AND DISCUSSION

SARIMA Model

The model identification process is carried out by looking at the actual data plot, as well as seeing if the stationary data is within variance and average.

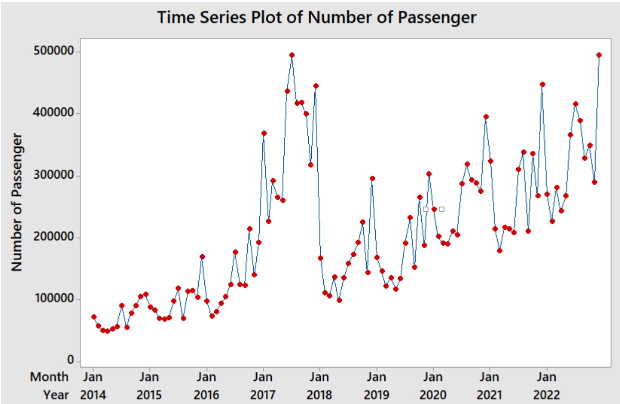


Figure 1. Passenger Data Plot

Based on the data plot in Figure 1, it can be seen that the data non stationer which is characterized by an upward trend pattern. In addition to the upward trend, the data also forms a seasonal pattern that can be seen from the increase in the number of passengers at each certain period regularly.

Testing the stationarity of data on variance is done by looking at rounded values. Data is said to be stationer in variance if the rounded value of the fault is one. If the data is non stationer in variance, then a Box-Cox transformation is performed. After testing, the data reached stationer in variance after performing two transformations which can be seen in Figure 2.

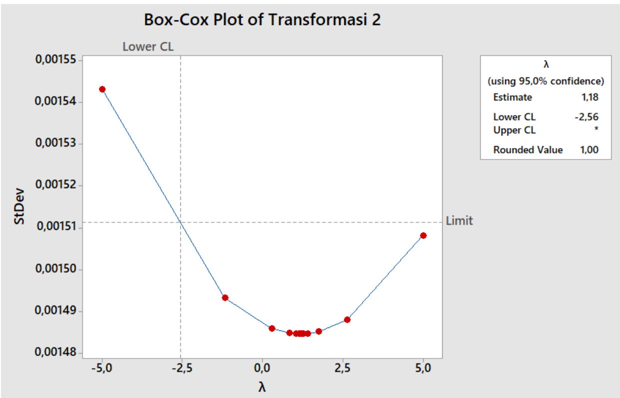


Figure 2. Stationary Test on Variance

J. Pijar MIPA, Vol. 18 No. 4, July 2023: 632-637
DOI: 10.29303/jpm.v18i4.5469

ISSN 1907-1744 (Print)
ISSN 2460-1500 (Online)

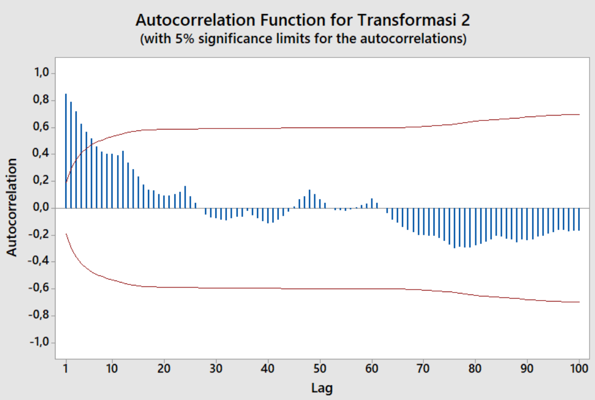


Figure 3. Stationary Test on Average

Stationary testing of the mean is performed using plots of Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF). The results of stationary testing against the average can be seen in Figure 3.

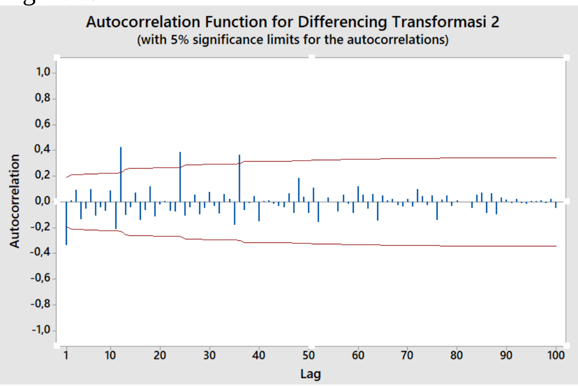


Figure 4. Stationary Test on Average with first Difference data

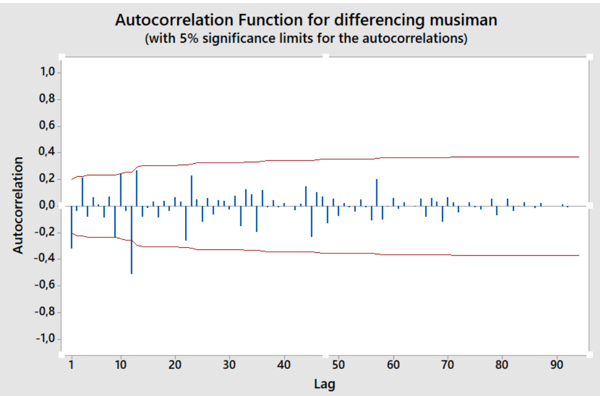
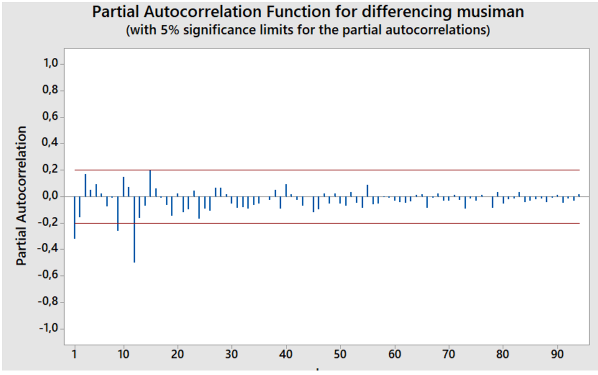


Figure 5. Seasonal pattern ACF plot

In Figure 3 it can be seen that the first three lags are out of the confidence interval. This indicates that the data is not stationary in the average. Furthermore, differencing of passenger data was carried out and ACF plots were carried out again with the results in Figure 4.

The test results in Figure 4 after the first difference showed that the data pattern was stationer because the first three lags did not cross the confidence interval. Although the data is already stationer in the average, the seasonal pattern still shows a pattern that is non stationer because the first three lags in the seasonal pattern still pass the confidence interval. To overcome this problem, differencing was carried out on the seasonal pattern and stationer tests were carried out again using the ACF plot shown in Figure 5.

In figure 5 it can be seen that the data is already stationer on seasonal patterns because only one first seasonal lag is out of the confidence interval. Furthermore, the order of SARIMA will be determined based on the ACF and PACF plots.



Gambar 6. Plot PACF pola musiman

Table 1. SARIMA Model

No	SARIMA Model	MSD	Description
1	$(1,1,1)(1,1,1)^{12}$	0.0000028	AR(1), SAR(12), MA(1) not Significant
2	$(0,1,1)(1,1,1)^{12}$	0.0000028	SAR(12) not Significant
3	$(1,1,0)(1,1,1)^{12}$	0.0000028	SAR(12) not Significant
4	$(1,1,1)(0,1,1)^{12}$	0.0000028	AR(1), MA(1) not Significant
5	$(1,1,1)(1,1,0)^{12}$	0.0000030	AR(1), MA(1) not Significant
6	$(0,1,1)(0,1,1)^{12}$	0.0000021	All parameters are significant
7	$(1,1,0)(1,1,0)^{12}$	0.0000030	All parameters are significant

The ACF and PACF plots in Figures 4 and 5 show that one of the first three lags passes the confidence interval. This shows that the order AR and MA are one. For the seasonal order in the ACF and PACF plots there is also one seasonal lag that comes out of the confidence interval so that the seasonal order for AR and MA is also one. Here is the tentative SARIMA model that will be selected based on the best model.

Of the several SARIMA models above, the best model is the SARIMA $(0,1,1)(0,1,1)^{12}$. The model has the smallest MSD value and all parameters of the model are significant at alpha 0.05.

DOI: 10.29303/jpm.v18i4.5469
ISSN 1907-1744 (Print)
ISSN 2460-1500 (Online)

Model Winter Exponential Smoothing

The formation of a prediction model using the Winter Exponential Smoothing method uses a multiplicative model because seasonal pattern fluctuations in ship passenger data are more variable, where additive models are used if seasonal pattern fluctuations are stable. The Winter Exponential Smoothing method uses parameters α (level), β (trend), and δ (musiman). Model selection is done by a try and error process with a value scale of 0.1 to 0.9 and combine to find the smallest error value. Before modeling using winter exponential smoothing, a transformation process was carried out on the original data first, because in the SARIMA method the data on the number of ship passengers had been transformed. This needs to be done to make a comparison of the error rate or MSD value. Here are some α, β, dan δ parameter tests.

Table 2. Model Parameter Testing

No	α	β	δ	MAPE	MAD	MSD
1	0.1	0.1	0.1	2.1474	0.0018	0.00001
2	0.5	0.1	0.1	1.6879	0.0014	0.00000
3	0.9	0.1	0.1	1.4545	0.0012	0.00000
4	0.1	0.5	0.1	2.3613	0.0019	0.00001
5	0.5	0.5	0.1	1.7593	0.0015	0.00000
6	0.9	0.5	0.1	1.5131	0.0013	0.00000
7	0.1	0.9	0.1	2.5880	0.0021	0.00001
8	0.5	0.9	0.1	1.4947	0.0012	0.00000
9	0.9	0.9	0.1	2.0445	0.0017	0.00000
10	0.1	0.1	0.5	2.1797	0.0018	0.00001
11	0.1	0.5	0.5	2.5359	0.0021	0.00001
12	0.1	0.9	0.5	3.7836	0.0031	0.00002
13	0.5	0.1	0.9	1.5979	0.0013	0.00000
14	0.5	0.5	0.9	1.8758	0.0015	0.00000
15	0.9	0.9	0.9	2.2537	0.0019	0.00001

Based on Table 2, the best model with the smallest error values is obtained at α = 0.9, β = 0.1, dan δ = 0.1. The model has an error rate with MAPE of 1.4545, MAD of 0.0012 and MSD of 0.0000. This model will be used for predictions on the number of passengers on the ship.

In Figure 3 it can be seen that the first three lags are outside the confidence interval. This shows that the data is non stationer against the average. Furthermore, the passenger data difference was carried out and the ACF plot was carried out again with the results in Figure 4.

Table 3. Comparison of error

No	Model	MSD
1	SARIMA	0.0000021
2	Winter Exponential Smoothing	0.0000000

Based on the error rate comparison, the Winter Exponential Smoothing model has a minimum MSE error value of 0.0000 which is smaller than the SARIMA model, so that the Winter Exponential Smoothing model with α = 0.9, β = 0.1, dan δ = 0.1 is feasible to be used as a prediction model for the number of ship passengers in Batam city. The following are the results of the prediction of the number of ship passengers in Batam using the Winter Exponential Smoothing model with α = 0.9, β = 0.1, dan δ = 0.1.

The results of this study are in line with a study on the comparison of winter Exponential Smoothing and SARIMA. This study found that the winter Exponential Smoothing model is the best model in forecasting the number of international flight passengers at Soekarno Hatta International Airpor [23]. In some studies it was also found that the Exponential Smoothing model has a better level of accuracy than SARIMA in forecasting seasonal time series data [24-28].

Table 4. Predicted Results of Ship Passengers in 2023

No	Month	Predicted Number of Passengers
1	January	307369
2	February	229254
3	March	221567
4	April	232625
5	May	229097
6	June	280853
7	July	379040
8	August	325314
9	September	321488
10	October	402449
11	November	324393
12	December	502070

The model plot and the results of the prediction of the number of ship passengers in Batam City can be seen in Figure 7.

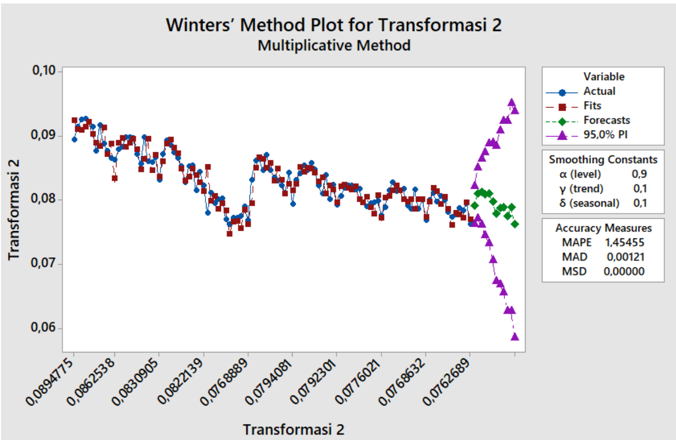


Figure 7. Plot Model Winter Exponential Smoothing

Figure 7 shows a graph of the predicted number of passengers in 2022. The blue line is the original data and the red line is the smoothing of the forecasting model's fits. The forecasting results can be seen on the

ISSN 1907-1744 (Print)
ISSN 2460-1500 (Online)

green line. The purple line represents the lower border and upper border. The predicted results are seasonal although there is a not too significant increase. In December the number of passengers experienced a significant increase. So, by making predictions, it can be known whether the number of passengers each year has increased or decreased.

CONCLUSION

In this study, the best model to predict the number of ship passengers in Batam City is the Winter Exponential Smoothing model with α = 0.9, β = 0.1, $dan\ \delta$ = 0.1. The Winter Exponential Smoothing model has the most precise level of accuracy with the smallest error value compared to the Sarima model. In future research, Winter Exponential Smoothing can be used to compare the predicted value between the number of domestic and foreign ship passengers.

ACKNOWLEDGEMENTS

Thank you to LPPM Batam Institute of Technology for providing support in the form of internal research funds for this research.

REFERENCES

[1] Ahmed Shokeralla, A., Al Rahman Idress Al Sameeh, F., Gibreel Mohammed Musa, A., ELrhman Elsmih, F., & Shokeralla, A. A. (2020). Prediction the daily number of confirmed cases of covid-19 in Sudan with arima and holt winter exponential smoothing. Article in International *Journal of Development Research*, August.
[2] Arifin, B., Tanaya, G. L. P., & Usman, A. (2021). Peramalan Harga Kelapa Sawit Dunia Pada Tahun 2020-2024. *Prosiding Saintek LPPM Universitas Mataram*, 3, 349–368.
[3] ArunKumar, K. E., Kalaga, D. V., Mohan Sai Kumar, C., Kawaji, M., & Brenza, T. M. (2022). Comparative analysis of Gated Recurrent Units (GRU), long Short-Term memory (LSTM) cells, autoregressive Integrated moving average (ARIMA), seasonal autoregressive Integrated moving average (SARIMA) for forecasting COVID-19 trends. *Alexandria Engineering Journal*, 61(10), 7585–7603.
[4] 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.
[5] Ayunda, N., Ningsih, L., & Sari, A. N. (2022). Pengujian Model Multiplicative Holt Winter ' s Exponential Smoothing dalam Peramalan Data Time-Series Terdampak Covid- Holt Winter ' s Exponential Smoothing Multiplicative Model Testing for Time-Series Data Forcasting that Impacted by Covid-19. 12(1), 41–50.
[6] Azmi, A. N., Ruhiat, D., & Kamilah, W. N. (2023). Pemodelan Forecasting Tingkat Inflasi di Indonesia Menggunakan Metode Triple Exponential Smoothing dan Seasonal Arima. 3, 19–29.
[7] Fahik, D. S., & Jatipaningrum, M. T. (2021). Peramalan Jumlah Penumpang Penerbangan Internasional Di Bandar Udara Internasional Soekarno Hatta Dengan Metode Holt-Winters Exponential Smoothing Dan Seasonal Arima. *Statistika Industri Dan Komputasi*, 6(1), 77–87.
[8] Fahrudin, R. (2018). Forecasting Tourist Visits Using Seasonal Autoregressive Integrated Moving Average Method. *IOP Conference Series*: Materials Science and Engineering, 407(1).
[9] Golding, N., Hewitt, C., Zhang, P., Liu, M., Zhang, J., & Bett, P. (2019). Co-development of a seasonal rainfall forecast service: Supporting flood risk management for the Yangtze River basin. *Climate Risk Management*, 23(November 2018), 43–49.
[10] Gupta, R., & Pal, S. K. (2020). Trend Analysis and Forecasting of COVID-19 outbreak in India. MedRxiv, 1–19.
[11] Harizahayu, Harahap, A., Fathoni, M., & Sumardi, H. (2023). Prediction Farmer Exchange Rate Comparative Method of Analysis Holth-Winters Smoothing and Seasonal ARIMA (Vol. 1). *Atlantis Press* SARL.
[12] Jere, S., Banda, A., Kasense, B., Siluyele, I., & Moyo, E. (2019). Forecasting Annual International Tourist Arrivals in Zambia Using Holt-Winters Exponential Smoothing. Open *Journal of Statistics*, 09(02), 258–267.
[13] Kodippili, A., & Senaratne, D. (2017). Forecasting Tourist Arrivals to Sri Lanka Using Seasonal ARIMA. *An International Peer-Reviewed Journal*, 29(December), 2312–5179. www.iiste.org
[14] Laksono, J., Kristiantoro, H., Tjahjadi, B., & Soewarno, N. (2018). Penentuan Lokasi Pembangunan Pusat Logistik Berikat Di Provinsi Jawa Timur Berdasar Aspek Sustainability Dengan Menggunakan Metode Analytic Hierarchy Process. EKUITAS (*Jurnal Ekonomi Dan Keuangan*), 1(3), 342–360.
[15] Negara, R. I. P. (2021). Peramalan Jumlah Penumpang Kapal di Pelabuhan Pantai Baru dengan Metode Sarima Dan Winter ' s Exponential Smoothing. *Jstar*, 1(1), 63–78.
[16] Nurhamidah, N., Nusyirwan, N., & Faisol, A. (2020). Forecasting Seasonal Time Series Data Using the Holt-Winters Exponential Smoothing Method of Additive Models. *Jurnal Matematika Integratif*, 16(2), 151.
[17] Nurjanah, I. S., Ruhiat, D., & Andiani, D. (2018). Implementasi Model Autoregressive Integrated

Moving Average (Arima) Untuk Peramalan Jumlah Penumpang Kereta Api Di Pulau Sumatera. TEOREMA : *Teori Dan Riset Matematika*, 3(2), 145.

[18] Pongdatu, G. A. N., & Putra, Y. H. (2018). Seasonal Time Series Forecasting using SARIMA and Holt Winter's Exponential Smoothing. *IOP Conference Series*: Materials Science and Engineering, 407(1).

[19] Putri, S., & Sofro, A. (2022). Peramalan Jumlah Keberangkatan Penumpang Pelayaran Dalam Negeri di Pelabuhan Tanjung Perak Menggunakan Metode ARIMA dan SARIMA. MATHunesa: *Jurnal Ilmiah Matematik*a, 10(1), 61–67.

[20] Rabbani, M. B. A., Musarat, M. A., Alaloul, W. S., Rabbani, M. S., Maqsoom, A., Ayub, S., Bukhari, H., & Altaf, M. (2021). A Comparison Between Seasonal Autoregressive Integrated Moving Average (SARIMA) and Exponential Smoothing (ES) Based on Time Series Model for Forecasting Road Accidents. *Arabian Journal for Science and Engineering*, 46(11), 11113–11138.

[21] Rashed, Y., Meersman, H., Van De Voorde, E., & Vanelslander, T. (2017). Short-term forecast of container throughout: An ARIMA-intervention model for the port of Antwerp oa. *Maritime Economics and Logistics*, 19(4), 749–764.

[22] Rezky, M., & Payu, F. (2019). Metode Exponential Smoothing Event Based (Eseb) Dan Metode Winter ' S Exponential Smoothing (Wes) Untuk Peramalan Jumlah Penumpang Tiba di Pelabuhan Penyeberangan Gorontalo Exponential Smoothing Event Based (ESEB) Method And Winter ' s Exponential S. 13(April), 199–204.

[23] Rufino, C. C. (2018). Forecasting monthly tourist arrivals from ASEAN + 3 countries to the Philippines for 2015-2016 Using SARIMA Noise Modeling Forecasting. Research Gate, March.

[24] Soeryawinata, J., Novianus, H., & Santoso, L. W. (2022). Sales Forecasting pada Dealer Motor X Dengan LSTM, ARIMA dan Holt-Winters Exponential Smoothing. Jurnal Infra.

[25] Sugishita, Y., Kurita, J., Sugawara, T., & Ohkusa, Y. (2020). Forecast of the COVID-19 outbreak, collapse of medical facilities, and lockdown effects in Tokyo, Japan. MedRxiv, 2020.04.02.20051490.

[26] Surabaya-jayapura, R., Ainul, E., Statistika, D., Matematika, F., & Data, S. (2019). Peramalan Penumpang Angkutan Laut di PT Pelayaran Nasional Indonesia (Persero) Cabang Surabaya. 8(2).

[27] Tambuwun, P. F. A., Nainggolan, N., & Langi, Y. A. R. (2023). d ' CartesiaN Jurnal Matematika dan Aplikasi Peramalan Banyaknya Penumpang Bandar Udara Internasional Sam Ratulangi Manado Dengan Metode Winter ' s Exponential Smoothing dan Seasonal ARIMA.

[28] Trull, Ó., García-Díaz, J. C., & Troncoso, A. (2020). Stability of multiple seasonal holt-winters models applied to hourly electricity demand in Spain. Applied Sciences (Switzerland), 10(7), 1–16.