

Forecasting Foreign Tourist Visits to West Nusa Tenggara Using ARIMA Method

Siti Soraya¹, Maulida Nurhidayati², Baiq Candra Herawati³, Anthony Anggrawan⁴, Lalu Ganda Rady Putra⁵, Didiharyono D⁶

^{1,3,4,5}Universitas Bumigora, Nusa Tenggara Barat, Indonesia

²IAIN Ponorogo, Jawa Timur, Indonesia

⁶Universitas Andi Djemma, Sulawesi Selatan, Indonesia

Article Info

Article history:

Received : 09-09-2021

Revised : 11-03-2021

Accepted : 11-05-2021

Keywords:

ARIMA;
Box-Jenkins;
Forecasting;
Pandemic Covid-19;
Tourism.

ABSTRACT

West Nusa Tenggara (NTB) is one of the provinces in Indonesia that has its own charm in the world of tourism and is known as a pioneer of halal tourism. In addition to domestic tourists, NTB tourism always has an attraction for foreign tourists. This is evidenced by the increasing number of foreign tourists visiting NTB from year to year before the Covid-19 pandemic. This condition certainly has a positive impact on increasing NTB's economic growth in the tourism sector and indirectly on the optimization of existing infrastructure. The purpose of the study is to predict the number of foreign tourist visits to NTB so that it can assist the government in making decisions in preparing adequate facilities and infrastructure in the event of a surge in tourist visits. The method used in this study is the Box-Jenkins-ARIMA model. The ARIMA method is based on 3 models that are formed from the results of plot data. The data used in this study is secondary data sourced from the Central Statistics Agency (BPS) of West Nusa Tenggara (NTB), from January 2010 to June 2019. The results show that the ARIMA (4,1,1) model is the most widely used model. This model is suitable for predicting the number of foreign tourists visiting NTB because this model produces the lowest SSE and MSE values compared to other models.

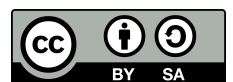


Accredited by Kemenristekdikti, Decree No: 200/M/KPT/2020
DOI: <https://doi.org/10.30812/varian.v5i1.1441>

Corresponding Author:

Baiq Candra Herawati,
Universitas Bumigora, Nusa Tenggara Barat.
Email: candrah@universitasbumigora.ac.id

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



A. INTRODUCTION

In this millennial era, tourism has become a sector that greatly contributes to the economic development of a region. One area in Indonesia that is of sufficient interest to foreign tourists is West Nusa Tenggara (NTB). With the unavailability of natural resources and large-scale manufacturing industries in NTB (Rizal et al., 2019), the tourism sector has indirectly assumed the role as one of the biggest economic contributors in NTB. With the tourism industry, certainly can open new jobs and ultimately can improve the economy of the community (Didiharyono and Syukri, 2020).

The development of the tourism sector must be carried out in a sustainable manner so as to provide direct benefits for the welfare of the community because the tourism sector is one component of economic development (Fahmeyzan et al., 2018). The policy direction in the development of the tourism sector includes marketing of national tourism by bringing in tourists and even foreign tourists.

As the increasing flow of tourist arrivals from foreign countries to NTB, the government needs to make good and sustainable planning so that it can optimally provide excellent service that ultimately can provide satisfaction to tourists. When tourists feel

satisfied, allow these tourists to return to visit NTB. Good and sustainable planning can be supported by the prediction of the number of tourist visits. The goal is that the direction of planning and policy in accordance with existing conditions. Among other things, it can be estimated the number of workers available, the availability of facilities and infrastructure which leads to benchmarks in improving the economy as well as the welfare of the community (Setiawan, Santi Puteri Rahayu, 2017), especially NTB. In order to support the national tourism target, NTB Provincial Tourism Office data from 2009 to 2017 show a significant increase in tourism trends. NTB's economic growth in 2017 reached 7.1 percent, exceeding national economic growth (5.6 percent). It is hoped that the next 5 years will also have a more positive impact on the welfare of the people of NTB.

NTB's poverty rate decreased from 23.08 percent in 2008 to 15.05 percent in 2017. The unemployment rate also decreased from 6.25 percent in 2009 to 3.32 percent in 2017. In the first quarter of 2018, the province of NTB occupies the lowest position in Indonesia in terms of economic growth. In 2019, through the Halal Tourism branding. The province of NTB is targeting 4 million foreign tourists, 2 million foreign tourists, and 2 million domestic tourists. To support this, in 2019, 18 events have been held, 4 of which have been included in the Top 100 National Event Calendar, namely the Enchantment of Bau Nyale, the Enchantment of Tambora Festival, the Enchantment of Moyo Festival and the Enchantment of the Khazanah Ramadhan. In the world of tourism, predicting the number of tourist visits is something that needs to be done (Annisa Fitri, Ika Purnamasari, 2019). What's more, NTB is one of the destinations most frequently visited by tourists after Bali, for the Nusa Tenggara region. Besides that, one of the determinants or influences on the economy is from the tourism sector (Nguyen, 2020).

However, due to the Covid-19 Pandemic in 2020, the number of tourism visits has decreased, even in a few months there were zero visits because the NTB government implemented the Large-Scale Social Restrictions (PSBB) policy and tourism sites were closed as an effort to cut the spread of Covid-19 in NTB. Therefore, in forecasting it is necessary to consider the Pandemic Covid-19 variable which has an impact on the decrease in the number of foreign tourists to NTB.

Several previous studies related to forecasting the number of tourist visits have been carried out, among others: (Waluyo, 2019) in the Greater Bandung Tourism Area for the period June 2017 to December 2017 using the ARIMA model and the results obtained are 78.895% able to predict the number of tourist visits. Next (Hendayanti and Nurhidayati, 2020) in Bali for the period January 2019 to December 2019 the ARIMA model is 90% able to predict the number of tourist visits visiting these tourist attractions. However, this article involves a new element, namely intervention variables such as earthquakes and the COVID-19 pandemic which resulted in the low number of foreign tourist visits to West Nusa Tenggara.

Based on the background, forecasting on the number of foreign tourist visits is very necessary to do. Therefore, the ARIMA model is very appropriate to be used to forecast foreign tourists visiting, knowing the nature of the data to come. This process is done by dividing the data into in-sample data and out of sample data. Out of sample data is used to predict or validate the accuracy of the data in an deepening process. So the emerging model is the best model from in-sample data (Hendayanti and Nurhidayati, 2020). In this study, the selection of the best model is done using the SSE and MSE values in the existing model then forecasting based on the best in-sample model.

B. LITERATURE REVIEW

The Box-Jenkins method is a method of applying time series data popularized by Box and Jenkins in 1970, by combining the moving average and autoregressive approaches (City et al., 2019). Box-Jenkins recommends that whether or not a stationary time series of data is stationary, in the process one or more process differentiations can be carried out using the ARIMA Model approach (Eni and Adeyeye, 2015). A method that is able to solve various problems in forecasting time series data. Included in predicting the number of tourists visiting NTB, by looking at visiting patterns of visits in the past (Didiharyono and Bakhtiar, 2018).

The ARIMA model is formed from 3 models, namely Autoregressive (AR), Moving Average (MA), and Autoregressive and Moving Average (ARMA), which are preceded by checking stationary data (Chuang and Wei, 1991).

If there is order d ($d \geq 1$) in a Z_t forecasting process then the homogeneous non-stationary model is known as the ARIMA (p, d, q) Model (Annisa Fitri, Ika Purnamasari, 2019). In general, the ARIMA Model can be seen in equation (1) (Chuang and Wei, 1991):

$$\phi_p(B)(1-B)^d Z_t = \theta_0 + \theta_q(B)a_t \quad (1)$$

with: p is the number of autoregressive terms; d is the number of differences; q is the number of moving averages; B is the backshift operator; and a_t is the number of white noise.

Where AR is represented by:

$$\phi_p(B) = (1 - \phi_1 B - \dots - \phi_p B^p) \quad (2)$$

and the MA is represented by:

$$\theta_q(B) = (1 - \theta_1 B - \dots - \theta_q B^q) \quad (3)$$

In the next section, the parameter θ_0 shows different roles between $d = 0$ and $d \geq 0$. When $d = 0$ implies that the data is stationary, such as: $\theta_0 = \mu(1 - \theta_1 - \dots - \theta_q)$.

Data plots and identification of the formation of the ARIMA Model are things that need to be done (Zhang et al., 2019). This is related to the stationary in the data is a condition of the ARIMA Model, both stationary in mean and variance (Arini and Nawangsih, 2015). If the data is not stationary, transformation needs to be done (Didiharyono and Syukri, 2020). When the stationary requirements have been met then the identification of the model formed can be seen in the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots also need to be done (Li and Zhang, 2009; Series, 2020). So as to ensure that the model that reflects the forecasting is good, the estimation of the best model that aims to ensure that residuals are white noise or not is very necessary.

C. RESEARCH METHOD

In this study the Box-Jenkins Method is used with the ARIMA Model. The ARIMA model is used to predict the number of foreign tourist arrivals to the Province of West Nusa Tenggara in January 2021 to December 2023. Data used as forecasting is the monthly data of foreign tourist visits from January 2010 to June 2019. The data in this study were sourced from the Central Statistics Agency of Nusa Tenggara West through the page <https://ntb.bps.go.id/>. The stages of the research carried out are as follows (Hendayanti and Nurhidayati, 2020):

- a. Identifying data through data plots and checking data validation, both in mean and variance
- b. Identifying ARIMA models
- c. Determination of the parameters p, d and q in ARIMA
- d. Determination of the appropriate ARIMA model equation is used for the application of data forecasting.
- e. Validate the time series data used for forecasting.
- f. Forecasting.

D. RESULT AND DISCUSSION

1. Data Plot and Identification of Stationary Models

At this stage stationarity testing is done both in mean and variance. Stationary testing in variants is carried out by Box-Cox testing. The Box-Cox test results are shown in Figure 1.

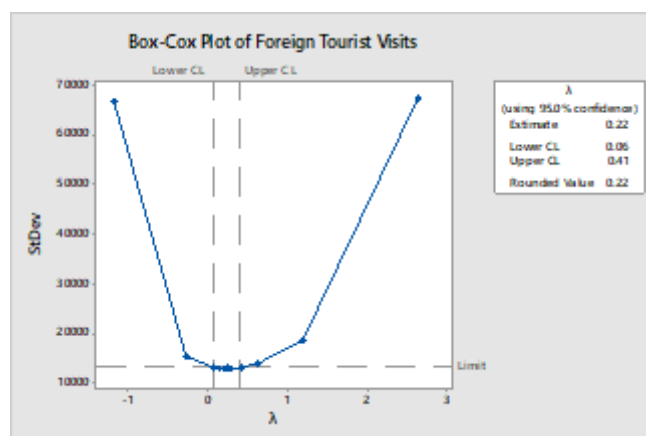


Figure 1. Box-Cox Test Results

The Box-Cox test results in Figure 1 show that the value of λ obtained is 0.22, which indicates that the data is not stationary in the variance, so the transformation is based on the value of λ obtained. The next step is testing stationarity in the mean by using ACF and PACF transformed tourist visit data as shown in Figure 2 and Figure 3.

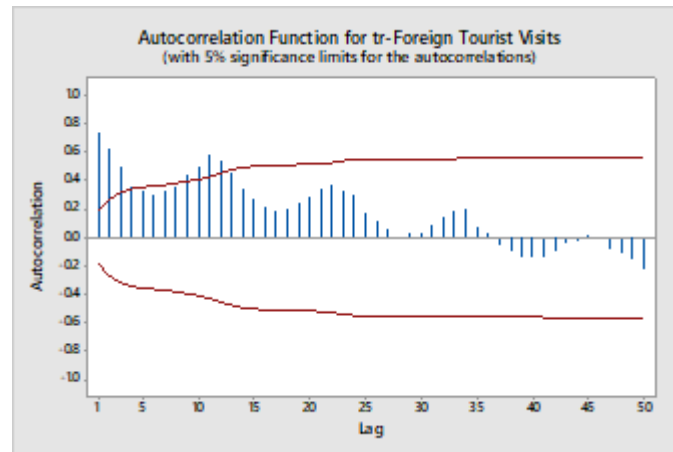


Figure 2. ACF Data Plot Transformation

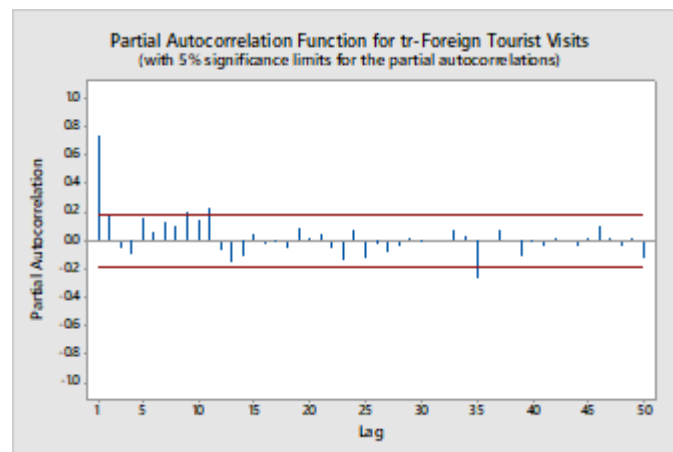


Figure 3. PACF Data Plot Transformation

Figure 2 and 3 show the results of the ACF plot falling slowly and the PACF plot cutting off in lag 1 means that the transformation data of the number of foreign tourist arrivals is not stationary in the mean so it needs to be done differencing to stationate the data. ACF and PACF plot results for differencing data are shown in Figure 4 and 5.

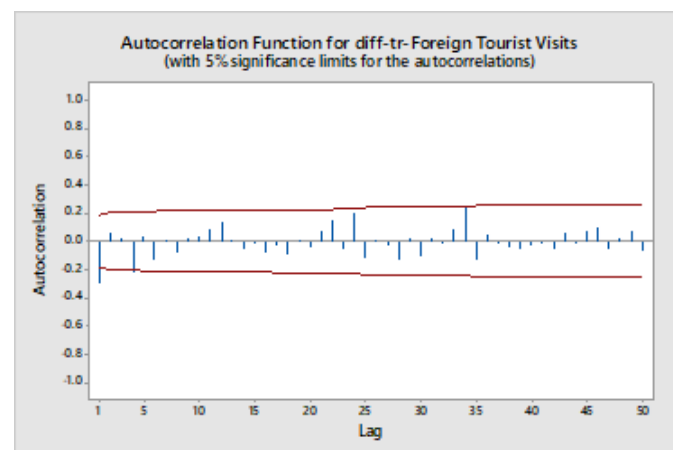


Figure 4. PACF Data Plot Differencing

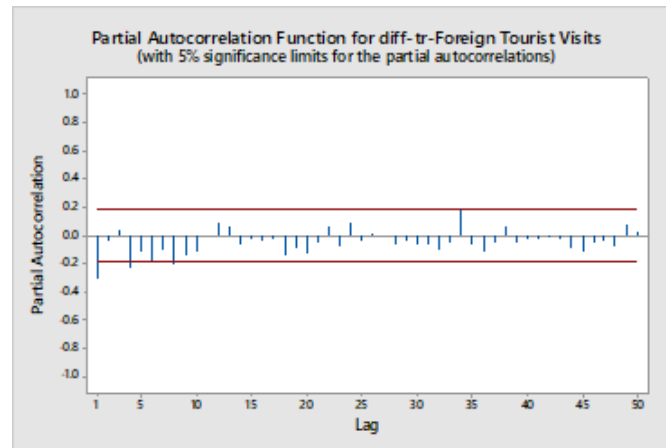


Figure 5. PACF Data Plot Differencing

Figure 4 and 5 show that the ACF plot cuts off in lag 1 while the PACF plot cuts off in lag 1 and 4 and decays towards zero for other lags. It can be concluded that differencing data follows the ARIMA (1,1,1) and ARIMA (4,1,1) models and the overfitting models of the two models are the ARIMA (1,1,4) and ARIMA (4,1,4) models. The next model is named first model for ARIMA (1,1,1), second model for ARIMA (4,1,1), third model for ARIMA (1,1,4) and four model for ARIMA (4,1, 4).

2. Estimation of The Model and Selection of The Best Model.

The ARIMA Model is estimated using Minitab 15 with the results of the estimated ARIMA model parameters shown in Table 1.

Table 1. Estimation Results for The ARIMA Model

Model	Type	Coef	P
ARIMA(1,1,1)	AR 1	-0.1649	0.590
	MA 1	0.1528	0.618
	Constant	0.0561	0.598
ARIMA(4,1,1)	AR 1	0.4951	0.000
	AR 2	0.2174	0.043
	AR 3	-0.01	0.925
	AR 4	-0.2239	0.021
	MA 1	0.9746	0.000
	Constant	0.019998	0.000
ARIMA(1,1,4)	AR 1	0.0747	0.791
	MA 1	0.5381	0.052
	MA 2	0.0152	0.930
	MA 3	0.1048	0.330
	MA 4	0.3008	0.016
	Constant	0.03566	0.003
ARIMA(4,1,4)	AR 1	0.789	0.000
	AR 2	0.5219	0.015
	AR 3	-0.8731	0.000
	AR 4	-0.0108	0.934
	MA 1	1.3312	0.000
	MA 2	0.2103	0.294
	MA 3	-1.2458	0.000
	MA 4	0.5655	0.000
	Constant	0.0144	0.394

Table 1 shows that all parameters of model 1 were not significant because the resulting P value was more than 0.05. Model 2 has 1 parameter which is not significant, that is on AR 3 with a value of $P = 0.925$. Model 3 only has 1 significant parameter, which is MA 4. And Model 4 has 2 insignificant parameters. Because model 1 has all the parameters are insignificant, the model is excluded from the analysis and there are only 3 models which are then subjected to diagnostic check testing to ensure the model obtained from the estimation results meets the white noise assumption using Ljung-Box Test in Table 2.

Table 2. White Noise Test Results

Model	Lag	Chi-Square	DF	P-Value
ARIMA(4,1,1)	12	5.4	6	0.50
	24	12.5	18	0.82
	36	29.9	30	0.47
	48	35.1	42	0.77
ARIMA(1,1,4)	12	7.7	6	0.26
	24	18.8	18	0.41
	36	38.4	30	0.14
	48	46.6	42	0.29
ARIMA(4,1,4)	12	9.5	3	0.02
	24	13.8	15	0.54
	36	27.5	27	0.44
	48	32.5	39	0.76

Based on Table 2 it is known that all Q statistics for lags 12, 24, 36, and 48 in models 2 and 3 are not significant, meaning that model 2 and model 3 are white noise. Model 4 has a value of $P = 0.02$ at lag 12 which indicates that lag 12 is significant, meaning that model 4 is not white noise. Next to determine the best model is done by looking at the performance value of each model using SSE and MSE. The results of the ARIMA model performance tests are shown in Table 3.

Table 3. ARIMA Model Performance Test Results

Model Performance Test	Model ARIMA(4,1,1)	Model ARIMA(1,1,4)
SSE	160.740	162.117
MSE	1.502	1.515
df	107	107

Based on Table 3, it is known that the df values of model 2 and model 3 are the same, namely 107 so that to determine the best model is done by looking at SSE and MSE. SSE and MSE results show that model 2 has smaller SSE and MSE values compared to model 3 so that the ARIMA model chosen is the ARIMA Model (4,1,1).

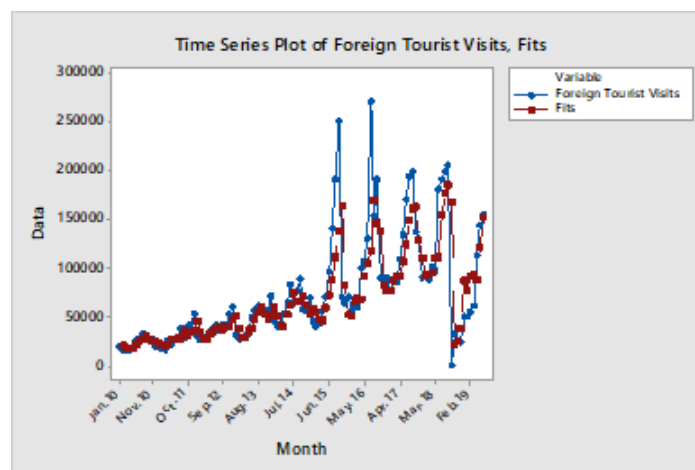


Figure 6. Data Plot of International Tourist Visits and ARIMA Models (4,1,1)

Figure 6 shows that the results of forecasting conducted with the ARIMA Model are able to follow the pattern of data possessed by data on the number of foreign tourist arrivals.

3. Forecasting Overseas Tourist Visits

The ARIMA model (4,1,1) is the best model because it has the smallest SSE and MSE values and meets the white noise assumption. This model is then used to forecast foreign tourists visiting for the next 12 months for 3 years, from January 2021 to December 2023. Forecasting results are shown in Table 4.

Table 4. Forecasting Overseas Tourist Visits 2021-2023

Month	2019	2020	2021	2022	2023
Jan		124919	151905	174656	199645
Feb		127402	153659	176635	201845
Mar		131132	155509	178636	204063
Apr		134886	157414	180657	206301
May		138331	159336	182697	208558
Jun		141140	161256	184755	210835
Jul	163752	143261	163164	186830	213130
Aug	152438	144904	165061	188922	215445
Sep	141860	146249	166955	191031	217779
Oct	133267	147507	168855	193157	220132
Nov	126429	148830	170768	195301	222505
Dec	124402	150288	172701	197464	224898

Based on Table 4 it is known that the largest number of tourist visits occurred in December with the number of each year that is 172701 visitors in 2021, 197464 in 2022, and continues to increase in 2023 by 224898 visitors. The results of this forecasting indicate that the number of tourist visits will always show an increase every year. This means that the development of tourism in NTB is quite good. This is supported by several tourism facilities as a support as well as adequate facilities and infrastructure that can provide comfort for foreign tourists visiting NTB. Increase in foreign tourist arrivals are shown in Figure 7.

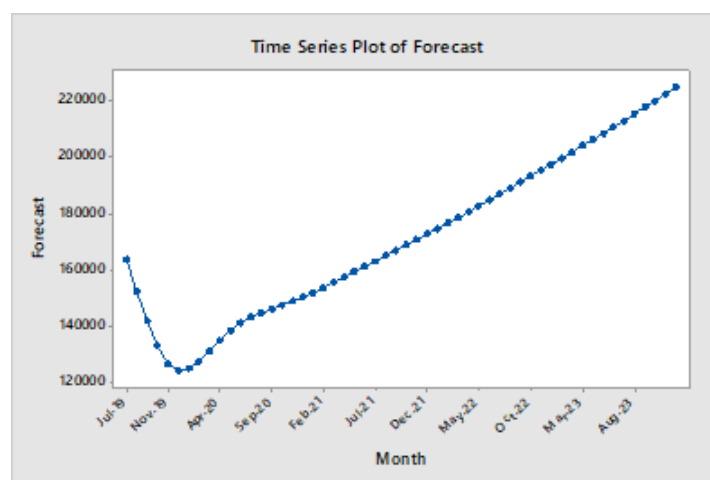


Figure 7. Data Plot Forecasting Foreign Tourist Visits in 2019-2023

E. CONCLUSION AND SUGGESTION

Based on the results of the research that has been done it can be concluded that in forecasting the number of foreign tourist visits to NTB using the ARIMA method, the results of the study are indicated by the value of the SSE and MSE so that the ARIMA model chosen is the ARIMA model (4,1,1).

From the forecasting process the number of foreign tourist arrivals to NTB for the next four periods, from January 2021 to December 2023, shows that the number of tourist arrivals shows a positive trend and is increasing every month. The highest achievement in each year occurred in December with 172701 visits in 2021, 197464 in 2022 and also continued to increase in 2023 by 224898. The earthquake disaster and the COVID-19 pandemic are not over yet. However, the NTB Regional Government must continue to operate, especially adding accommodation and other supporting facilities. In addition, the NTB Regional Government can also hold several events with the theme of regional culture or other promotions to increase tourist interest in visiting NTB.

REFERENCES

- Annisa Fitri, Ika Purnamasari, M. S. (2019). Peramalan Jumlah Wisatawan Mancanegara Menggunakan Model Arima. *Statistika*, 7(1).
- Arini, P. S. and Nawangsih, E. (2015). Peramalan Jumlah Kunjungan Wisatawan Mancanegara (Wisman) ke Bali Tahun 2019: Metode ARIMA. *Jurnal Ekonomi Kuantitatif Terapan*, 8(2):136–141.
- Chuang, A. and Wei, W. W. S. (1991). Time Series Analysis: Univariate and Multivariate Methods.
- City, S., Tahir, H. A., Ahmed, R. A., and Mhamad, A. J. (2019). Forecasting for the Imported Weight of Equipment to Cargo of Sulaimani International Airport. *sjcus*, 3(1):62–82.
- Didiharyono, D. and Bakhtiar, B. (2018). Forecasting Model With Box-Jenkins Method To Predict Tourists Who Visit Tourism Place In Toraja. *Jurnal of Economic, Management and Accounting*, 1(1):34–41.
- Didiharyono, D. and Syukri, M. (2020). Forecasting With Arima Model in Anticipating Open Unemployment Rates in South Sulawesi. *International Journal of Scientific and Technology Research*, 9(3):3838–3841.
- Eni, D. and Adeyeye, F. J. (2015). Seasonal ARIMA Modeling and Forecasting of Rainfall in Warri Town, Nigeria. *Journal of Geoscience and Environment Protection*, 03(06):91–98.
- Fahmeyzan, D., Soraya, S., Etmay, D., and Mataram, S. B. (2018). Uji Normalitas Data Omzet Bulanan Pelaku Ekonomi Mikro Desa Senggigi dengan Menggunakan Skewness dan Kurtosis. *Jurnal Varian*, 2(1):31–36.
- Hendayanti, N. P. N. and Nurhidayati, M. (2020). Perbandingan Metode Seasonal Autoregressive Integrated Moving Average (SARIMA) dengan Support Vector Regression (SVR) dalam Memprediksi Jumlah Kunjungan Wisatawan Mancanegara ke Bali. *Jurnal Varian*, 3(2):149–162.
- Li, W. and Zhang, Z. G. (2009). Based on Time Sequence of ARIMA Model in The Application of Short-Term Electricity Load Forecasting. *ICRCCS 2009 - 2009 International Conference on Research Challenges in Computer Science*, pages 11–14.
- Nguyen, A. T. (2020). The Relationship between Tourism and Economic Growth: Evidence from Oceania. *Journal of Tourism Management Research*, 7(1):32–41.
- Rizal, A. A., Soraya, S., and Tajuddin, M. (2019). Sequence to Sequence Analysis with Long Short Term Memory for Tourist Arrivals Prediction. *Journal of Physics: Conference Series*, 1211(1).
- Series, C. (2020). Self-Identification Deep Learning ARIMA Self-Identification Deep Learning ARIMA. *Journal of Physics: Conference Series*, 1564.
- Setiawan, Santi Puteri Rahayu, S. S. (2017). Economic Growth Modelling In East Java Using Bayesian. *Teknomatika*, 07(02):57–69.
- Waluyo, J. E. (2019). Peramalan Kedatangan Wisatawan Manca Negara Melalui Bandara Husein Sastra Negara Bandung Dengan Menggunakan Metode Arima (Autoregressive Integreted Moving Average). *Jurnal Kepariwisata: Destinasi, Hospitalitas dan Perjalanan*, 3(1):18–26.
- Zhang, Y., Luo, L., Yang, J., Liu, D., Kong, R., and Feng, Y. (2019). A Hybrid ARIMA-SVR Approach for Forecasting Emergency Patient Flow. *Journal of Ambient Intelligence and Humanized Computing*, 10(8):3315–3323.