Prediction Model for Tourism Object Ticket Determination in Bangkalan, Madura, Indonesia

Fifin Ayu Mufarroha[1]*, Akhmad Tajuddin Tholaby[2], Devie Rosa Anamisa[3], and Achmad Jauhari[4]

[1-4] Informatics Engineering Study Program, Department of Informatics Engineering, Faculty of Engineering, University of Trunojoyo Madura
Jln. Raya Telang, Jawa Timur 69162, Indonesia
[1]fifin.mufarroha@trunojoyo.ac.id; [2]akhmadtajuddin87@gmail.com;
[3]devros_gress@trunojoyo.ac.id; [4]jauhari@trunojoyo.ac.id

Received: 20th December 2021/ **Revised:** 13th December 2022/ **Accepted:** 14th December 2022

How to Cite: Mufarroha, F. A., Tholaby, A. T., Anamisa, D. R., & Jauhari, A. (2023). Prediction Model for Tourism Object Ticket Determination in Bangkalan, Madura, Indonesia. *ComTech: Computer, Mathematics and Engineering Applications, 14*(2), 69−81. https://doi.org/10.21512/comtech.v14i2.7992

Abstract **-** One of the regencies in Madura, namely Bangkalan, with its local wisdom and beautiful landscapes has the potential to become a tourism center. However, there may be a decrease in the number of visits caused by some factors. The research used the time series method to build a prediction model for tourist attraction entrance tickets. The model development aimed to estimate the number of tourist attraction visits in the future. The right model was needed to get the best prediction results. Least square, Holt-Winter, Seasonal Autoregressive Integrated Moving Average (SARIMA), and Rolling were chosen as the models. Data collection related to the number of tourist objects was carried out directly at the Tourism Office to obtain valid data. Using data on visitors to tourist attractions in Bangkalan Regency from 2015 to 2019, the results of measuring errors using Mean Absolute Percentage Error (MAPE) and Root Mean Squared Error (RMSE) are obtained. The error measurement results show that the Holt-Winter model has the lowest error rate of 5% and RMSE of 307,1198. Based on these calculations, the Holt-Winter model is the best model for determining tourist attraction entrance tickets. The ranking of the error measurement results from the highest to the lowest are Holt-Winter, Rolling, SARIMA, and Least Square methods.

Keywords: prediction model, tourism object ticket, ticket determination

I. INTRODUCTION

Indonesia has a variety of tourist attractions scattered throughout the islands. The tourism sector is one of the fastest-growing sectors in this modernization era. In addition to functioning as a means of vacation for families or tourists, tourism is also used as a stimulus to grow the regional economy. Another advantage is that the development of new tourism can be a business area for the community (Damanik et al., 2021). As a result, it opens up new job opportunities. It contributes to improving the structure of the regional economy. It can increase independence and competition and significantly participate in regional income.

Madura Island is one of the islands in Indonesia with four regencies (Umam et al., 2020; Darmawan et al., 2020). One of the regencies in Madura is Bangkalan. Bangkalan is the district that has the closest distance to Surabaya City, connected by the Suramadu bridge. It is an advantage for Bangkalan and makes it a tourism center in Madura. These tourist objects are scattered throughout the Bangkalan area. The map of the Bangkalan Regency can be seen in Figure 1. Bangkalan Regency is located close to Surabaya City, with the distance between Surabaya City and Bangkalan Regency which is 44 km. It takes approximately an hour to drive.

With the existence of several tourist objects in Bangkalan, the government's efforts to develop the potential of tourist attractions really need to be improved to attract tourists. However, it does not demand the possibility that the number of tourists will decrease due to some factors, such as the condition of the tourist attraction, security, facilities, and other factors (Higgins-Desbiolles, 2018; Higgins-Desbiolles et al., 2019; Scheyvens & Biddulph, 2018). The increasing and decreasing number of tourists becomes a thought in the aspect of improving tourism infrastructure and increasing trade in tourist areas

*Corresponding Author

based on the experience of previous years. Uncertain visitor numbers must be predictable so that local governments can determine policies for changes in visitor numbers in the future. This situation is closely related to the availability of admission tickets. The generation of admission tickets needs to be estimated when the number of tourists increases or decreases. In issuing the number of entrance tickets, its function can be optimized according to use in the field. It is necessary to avoid excess entry tickets, which can cause losses due to higher expenses and unused tickets.

Prediction is the activity of projecting what will happen in the future (Mahmud et al., 2021). There are several models of forecasting methods to examine various problems (Dahiwade et al., 2019; Cui & Jing, 2019). There are assumptions for forecasting models, such as based on the nature of the arrangement or forecast time and data patterns (Assidiq et al., 2017). The prediction method is a way of quantitatively predicting what will happen in the future based on relevant data in the past (Shano et al., 2020). Thus, this prediction method is used in objective predictions referring to data based on time.

The time series method is a forecasting method that uses the average of the last period of data to predict the next time (Berlinditya & Noeryanti, 2019). The specialty of this method is the construction of predictions using the assumption that the future is a function of the past. This assumption means that events within a certain period and the use of past data in forecasting make it easier to determine patterns (Robial, 2018).

Based on the characteristics of the data pattern on the number of tourist attraction visitors in Bangkalan, the time series method is expected to produce forecasts with high accuracy to predict entrance tickets. Therefore, the researchers build a prediction model to get the right model for determining entrance tickets to optimize its expenses according to the ticket inventory that is available every month. Thus, it results in printing admission tickets according to needs. It can assist tourism object managers in aligning the plan for achieving the target number of visitors in the development and control of well-organized tourist objects.

II. METHODS

The research is a continuation of previous research (Mufarroha et al., 2023). Previous researchers only used one method in forecasting, but in the current research, the researchers develop it by looking for the best model by applying four models at once. The number of tourist objects registered with the Bangkalan Regency Tourism Office is 21 attractions. Data collection related to the number of tourist objects is carried out directly at the Tourism Office to obtain valid data. It is also beneficial for the Tourism Office

Figure 1 Map of Bangkalan Regency (Google Maps, n.d.)

regarding the appropriate results. The data are the number of tourist attraction visitors in Bangkalan Regency over five years, from 2015 to 2019. Various types of information are quite large and have been observed over a relatively long period to make robust predictions. The amount of data are also very influential on the prediction results. The more power is collected, the better the estimation or forecasting will be obtained, and vice versa.

The prediction model is constructing a model to find forecasts in the future (Dharma et al., 2020). Several models are applied in developing the model, especially in the time series method. A time series data can be seen as a representation of the realization of a random variable, which usually has the same time interval and is observed in a certain period (Madsen, 2007). If the analyzed time series data have been identified, and a pattern has been selected based on that identification, a mathematical model which is a representation of the process of forming the time series data can be determined or selected. It is presented in Equation (1). Meanwhile, if the data are stable, the average is very close to the existing observation values, and future data are expected to be around the average value. It can be seen in Equation (2). These equations have x_t as the observed random variable at time *t*, *A* as the constant level of the model and the random deviation that occurs at time *t*, and F_{t+1} as a function of some or all of the observed time series data up to time *t*.

$$x_t = A + e_t \tag{1}$$

$$F_{t+1} = \sum_{t=1}^{T} \frac{x_t}{T} \tag{2}$$

Four models are selected as representative models in forecasting the cases specified in the research. It is intended to obtain the best model for determining the number of entrance tickets to tourist attractions in Bangkalan Regency. The flowchart of the problem-solving process is shown in Figure 2.

The Least Square method is one of the methods in the form of time series data, which requires data from the past to make forecasts (Liengaard et al., 2021; Yang et al., 2014; Dengen et al., 2018; Sianturi et al., 2020). This method performs analytical techniques in forecasting the future. The form of the equation for the linear trend value is presented in Equation (3). It has $\hat{Y}$ as the variable to find the trend and *X* as the time variable (day, month, year). Next, it looks for the values of constants a and b with Equation (4). The values of the variables *XY* and X^2 are needed to calculate the constants *a* and *b*.

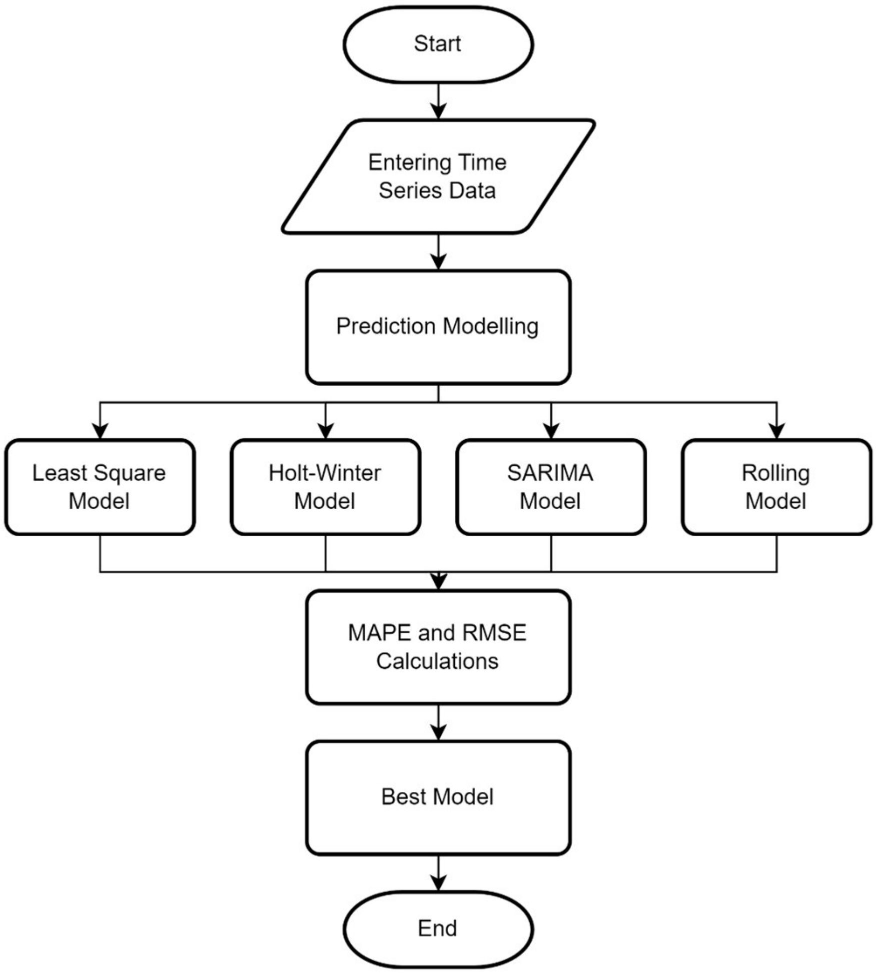


Figure 2 Research Flowchart

$$\hat{Y} = a + bX \quad (3)$$

$$a = \frac{\Sigma Y}{N}, b = \frac{\Sigma XY}{X^2} \quad (4)$$

The second model is the Holt-Winter. This method is an advanced exponential smoothing model. It is used when the data exhibit seasonal trends and behavior. It has been developed with three main equations called the Holt-Winter method, after the inventor's name, to deal with this problem. It has three main equations that determine the overall smoothing value, trend smoothing value, and seasonal smoothing value (Pertiwi, 2020).

The Holt-Winter method is also often called the exponential smoothing method, which can improve forecasting by increasing the chances of capturing different patterns in the data (Liu & Wu, 2022). This method is divided into two parts: the Seasonal Multiplication method for seasonal variations in data that have increased/decreased (fluctuating) and the Additive Seasonal method for constant seasonal variations. In the multiplicative Holt-Winter method, the seasonal component is expressed relative, and the time series can be seasonally adjusted by dividing by the seasonal component (Almazrouee et al., 2020).

Next, Seasonal Method Autoregressive Integrated Moving Average (SARIMA) is a time series forecasting method for stochastic model data with seasonal data patterns. This model is the development of the ARIMA model to fit seasonal time series specifically. It is designed to take into account the seasonal nature of the series to be modeled. Seasonality, in this case, refers to the regularly recurring changes in patterns over a certain period. In general, the SARIMA model can be formulated in Equation (5). It shows (*p, d, q*) as the non-seasonal part of the model and (*P, D, Q*) as the seasonal portion of AR. Meanwhile, the *P* variable shows the seasonal order for AR, while *Q* is the *MA* seasonal order. Then, the number of seasonal differences is represented as *D*. Next, *s* defines the number of periods until the pattern repeats again (Tadesse & Dinka, 2017).

$$ARIMA(p, d, q)(P, D, Q)s \quad (5)$$

The general principle of the Rolling model is to make forecasts every three months (quarterly), or if it has sufficient and competent resources, it can be made every month (Ton-Nu, 2014). Applying this model enables organizations to see the future by proposing a logical process. The main purpose of the rolling forecasting process is to facilitate more dynamic and proactive decision-making, unlike scenario planning, which only provides an estimate of a single point in the future (Henttu-Aho, 2018). Forecasting is not a simple task. It just extrapolates the budget prepared by Blab and updates the remaining time of the month. Instead, it involves a bottom-up process from budget holders to readjust assumptions. Assumptions used are the previous budget and the actual situation.

The test is carried out by testing errors using several methods of measuring the accuracy of the forecasting results. It consists of Mean Absolute Percentage Error (MAPE) and Root Mean Squared Error (RMSE). These techniques are employed to ascertain the precision of forecasted outcomes and pinpoint inaccuracies within the forecasting procedure. They may enhance their forecasting techniques and make better selections by gauging the accuracy of the forecasted results.

The average percentage difference between the expected and actual values is determined by MAPE. To compute it, divide the absolute difference between the actual and projected numbers by the actual value, and then multiply the result by 100. When evaluating the accuracy of several forecasting models, the MAPE approach is helpful. MAPE's formula can be seen in Equation (6). The square root of the average of the squared discrepancies between the predicted and actual values is determined by RMSE. It is helpful in determining the extent of forecasting mistakes. The RMSE formula is explained in Equation (7). It shows that A_t is actual value at time *t*. Meanwhile, F_t is forecasted value at time *t*.

$$MAPE = \frac{1}{n} \times \Sigma(|A_t - F_t| \times 100) \quad (6)$$

$$RMSE = \sqrt{\left(\frac{1}{n}\right) \times \Sigma(|A_t - F_t|)^2} \quad (7)$$

III. RESULTS AND DISCUSSIONS

The overall data on the number of tourist attraction visitors in Bangkalan Regency are described in Table 1 and Figure 3 (see Appendices). In 2017, there was a decrease in tourist arrivals around 915. Meanwhile, in 2016, 2018, and 2019, the decline in tourist visits was in the range of 850, 625, and 814. Based on the data, it needs to make the right prediction to overcome the determination of optimal entry tickets according to the available ticket inventory each month. Thus, the printing of entrance tickets is produced as needed.

Data are divided into two: learning and testing data. Learning data are determined in the range of January 2015 to June 2019. So, it has 42 learning data. Meanwhile, for the testing data, the data are selected in the range of July 2019 to December 2019. It has 6 testing data.

The first model built is the Least Square method. The Least Square method requires sales data in the past to forecast future sales so that the results can be determined. Least Square method sees trends from time series data. Figure 4 (see Appendices) shows the results of the calculation of the Least Square method. The implementation of this method results in relatively rising data.

The second model is Holt-Winter. The idea behind the Holt-Winter method is to apply exponential smoothing to seasonal components other than levels and trends. Figure 5 (see Appendices) shows the results of the Holt-Winter model. The forecasting results are shown in Figure 5 (see Appendices). The predicted values are in close proximity with the test data values. The two graphical charts have near proximity to one another.

Table 1 Bangkalan Tourism Monthly Visit Data

Year-Month	Total	Year-Month	Total
15-Jan	2693	17-Jul	3485
15-Feb	2761	17-Aug	3974
15-Mar	2698	17-Sep	3451
15-Apr	2727	17-Oct	3401
15-May	2670	17-Nov	3885
15-Jun	2691	17-Dec	3977
15-Jul	3644	18-Jan	4632
15-Aug	4506	18-Feb	4336
15-Sep	4384	18-Mar	4134
15-Oct	3897	18-Apr	3528
15-Nov	3779	18-May	4422
15-Dec	3687	18-Jun	4503
16-Jan	4524	18-Jul	4992
16-Feb	4540	18-Aug	4077
16-Mar	3354	18-Sep	4466
16-Apr	3276	18-Oct	4473
16-May	3389	18-Nov	4119
16-Jun	4214	18-Dec	4988
16-Jul	3383	19-Jan	4937
16-Aug	4001	19-Feb	4379
16-Sep	3359	19-Mar	4291
16-Oct	3580	19-Apr	4818
16-Nov	4189	19-May	4132
16-Dec	4481	19-Jun	4883
17-Jan	4123	19-Jul	5117
17-Feb	4748	19-Aug	4940
17-Mar	4137	19-Sep	4902
17-Apr	4603	19-Oct	4126
17-May	4157	19-Nov	4495
17-Jun	4122	19-Dec	5000

In the third and fourth models, SARIMA and Rolling models are conducted. It is necessary to identify the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) and calculate the difference. First, the Dickey-Fuller (DF) test is calculated and gets a higher-than-expected p-value (< 0,05). According to the DF test, the null hypothesis is a non-stationary time series. A high p-value supports the null hypothesis. So, it proves that the time series is not stationary. It needs to create a stationary time series by eliminating the trend first. It is done using a differencing technique. This technique takes the difference between the values in the time step and the previous time step (Jain & Mallick, 2017). For example, first-order differencing subtracts the value in the time step (*t*) and the value in the time step (*t-1*).

By removing trends, seasonality, and other non-stationary elements from the data, differencing improves the model's ability to identify underlying patterns and stabilizes time series variance. The differentiation results shown in Figure 6 (see Appendices) reflect the fluctuation of changes between months after implementing the first differencing. The trend is eliminated, resulting in a more stationary time series in future time series analysis. After the first order differencing process, the p-value is 0,000120. The significantly less than p-values indicate that the different time series are stationary. So, ACF and PACF plots can be carried out on the first difference series.

Based on Figure 7 (see Appendices), an advanced model can be built using the SARIMA algorithm. First, the researchers find the optimum *p, q* and *P, Q*. Then, the researchers set *d* as 1 (first differencing), *D* also as 1 and *s* as 12. As a general rule, the researchers set the model parameters that *D* never exceeds one, and the total difference '*d* + *D*' never exceeds 2.

The sequence of the model's autoregressive, differencing, and moving average components is shown by the numbers in parenthesis, respectively, on Figure 8 (see Appendices). The model's ordering in this instance is 1 for autoregressive, 1 for differencing, and 0 for moving average. The seasonal autoregressive, seasonal differencing, and seasonal moving average components of the model are denoted by the numbers enclosed in square brackets, respectively. In this instance, the model's seasonal period is 12 with its seasonal autoregressive order of 1, its seasonal differencing order of 1, and its seasonal moving average order of 0. The time required to fit the model to the data is known as the total fit time. The ARIMA model is useful for forecasting time series data, and the output provides information about the best model to use for the data.

Next, the researchers perform calculations to obtain the best ARIMA model. The values are assigned to the variables of *p, q, P*, and *Q*. The program runs calculations with the formula in Equation (4), accompanied by the results of the computation time in each model calculation. So, the best ARIMA model (1,1,0)(1,1,0) is obtained with following details: p=1, d=1, q=0, P=1, D=1, Q=0, and m=12. Based on the best model ARIMA calculation, the researchers set it as the value used in the forecast process. Figure 9 (see Appendices) shows the residual results. It verifies the residuals' mean and variance. Relatively consistent variance over time and a mean around zero are signs of well-behaved residuals. This implies that the model is accurately representing the general behavior of the data. In Figure 10 (see Appendices), the forecasting results demonstrate how this algorithm generates forecast estimates that are rather distant. Even though the predicting has been able to approach the test data in the final three months, the first three months' findings demonstrate that the forecasting has extremely far-off outcomes.

The last forecasting model is the application of the Rolling model. The residual in the Rolling model comes from the SARIMA model, so it can be depicted in Figure 11 (see Appendices). Explanation about residual plot is same as SARIMA residual plot. It aims to verify residuals' mean and variance for well-behaved residuals. It shows consistent variance over time and a mean around zero, indicating accurate data representation. Then, the forecasting results of this model are shown in Figure 12 (see Appendices). The Rolling forecasting model results show that this model produces forecasting figures that are close to the test data figures, as well as the Holt-Winter model. Likewise, it is with the graphic flow between the two numbers. However, the forecast numbers for the first month have a very large distance between the test data and the forecast results.

Based on the model development that has been done, the test is carried out by measuring the level of accuracy. The accuracy test is calculated to describe the data by using an error test to indicate how much the forecasting error is compared to the actual value. The level of error testing is MAPE and RMSE (Haq & Ni, 2019; Yang et al., 2020). The error measurement expresses the error percentage in forecasting the actual results over a certain period. It provides information on whether the error percentage is too high or too low. It means that MAPE is the absolute average over a given period, which is multiplied by 100% to get a percentage result (Tofallis, 2015). In terms of RMSE, the smaller the RMSE is, the better the model's capacity to forecast reliably is. A higher RMSE indicates a larger gap between projected and actual results. Each error value is obtained by applying the four models to the ticket sales time series data at a predetermined time, as shown in Table 2. Meanwhile, the comparison of the results of the data predictions from each model can be seen in Table 3.

Table 2 provides information that from the four models that have been worked on, the Holt-Winter model is the best model. It is proven by the MAPE error in this model being the lowest among the other models. Then, Table 3 shows the comparison in the period from July to December 2019. It shows the actual value of the data compared with the prediction results of the Least Square, Holt-Winter, SARIMA, and Rolling models. If Table 2 shows the error value, then Table 3 explains the value of the forecasting results for each model. It validates the measured error findings, making it simple to make connections between all the outcomes produced using the shown data.

Figure 13 (see Appendices) represents Table 3, which illustrates the results of the actual data with the predicted results of the four models from July to December 2019. The prediction results of the Least Square and SARIMA models have a big difference from the actual data. It also causes the error size of the two models to be very high compared to the other two models (Holt-Winter and Rolling models).

It can be summarized that based on Tables 2 and 3 and Figure 13 (see Appendices), the Least Square method occupies the first position with the highest error rate with MAPE of 10,54% and RMSE of

Table 2 Comparison of Accuracy Values

Model	MAPE	RMSE
Least Square Method	10,54	474,9828
Holt-Winter	5	307,1198
SARIMA	8,2	491,9056
Rolling	7	442,6829

Note: Seasonal Method Autoregressive Integrated Moving Average (SARIMA), Mean Absolute Percentage Error (MAPE), and Root Mean Squared Error (RMSE).

Table 3 Forecasting Testing Data

Period	Actual Data	Least Square	Holt-Winter	SARIMA	Rolling
Jul-19	5117	4698,88	4643,67	4325,06	4325,06
Aug-19	4940	4725,24	4967,66	4454,01	4873,47
Sep-19	4902	4751,60	4677,94	4195,13	4822,66
Oct-19	4126	4777,96	4583,56	4190,64	4826,84
Nov-19	4495	4804,32	4780,47	4366,90	4586,45
Dec-19	5000	4830,69	5008,09	4738,82	4804,08

Note: Seasonal Method Autoregressive Integrated Moving Average (SARIMA)

474,9828. The prediction results of this model provide an increasing value, which can be seen in Figure 13 (see Appendices). It can be said that this model is the worst model for predicting ticket sales.

After the Least Square Method model, the SARIMA model follows it with MAPE of 8,2% and RMSE of 491,9056. This model is also considered unsuitable for ticket prediction even though it has implemented the best ARIMA model calculations. Excavating this model can be done if people focus on determining the values in the formula for the best ARIMA model. It is when experiments are carried out one by one with the value of the variable until the error value is reduced. However, the point of the research is to compare SARIMA with other models. So, further and in-depth research can be carried out regarding the prediction of the SARIMA model on ticket prediction.

The second-best ranking is the Rolling model, with an error value with MAPE of 7% and RMSE of 442,6829. Although this model is better than the two previous models (Least Square Method and SARIMA), it is not the first choice. It is because Holt-Winter gets the smallest error values with MAPE of 5% and RMSE of 307,1198. It can also be seen in Figure 13 (see Appendices) that the Holt-Winter prediction line is closest to the actual data line compared to the Least Square, SARIMA, and Rolling models.

IV. CONCLUSIONS

The development of the forecasting model is intended to determine the number of entrance tickets to tourist attractions in Bangkalan Regency in Madura, Indonesia. Based on the data for the five years (2015 to 2019), the data are divided into learning and testing data. In modeling, the method is chosen based on time series because it can represent random variables. Finally, observations can be made over a certain period. Least Square, Holt-Winter, SARIMA, and Rolling models are selected. Then, error measurement is determined using MAPE and RMSE. The error measurement results show that the Holt-Winter model has the lowest MAPE error rate of 5% and RMSE of 307,1198. The ranking of the error measurement results from the highest to the lowest are Holt-Winter, Rolling, SARIMA, and Least Square. Based on the calculation of errors, the best model for predicting the number of entrance tickets to a tourist attraction is the Holt-Winter model.

The limitation of the research is time series data as the data used. The researchers do not take into account factors that can cause the number of tourist visits to change. Future research can consider the determining factors and use appropriate models with the influence therein. Moreover, the research aims to get the best model for predicting tourist object ticket determination with data from the last five years. Based on this aim, future research can implement the obtained model by adding time series data and providing prediction results to the manager of tourist attractions. The predicted results are intended to help the manager supply tickets as needed. The addition of time series data will further improve the prediction results.

REFERENCES

Almazrouee, A. I., Almeshal, A. M., Almutairi, A. S., Alenezi, M. R., & Alhajeri, S. N. (2020). Long-term forecasting of electrical loads in Kuwait using Prophet and Holt–Winters Models. *Applied Sciences*, *10*(16), 1–17. https://doi.org/10.3390/app10165627

Assidiq, A., Hendikawati, P., & Dwidayati, N. (2017). Perbandingan metode Weighted Fuzzy Time Series, Seasonal ARIMA, dan Holt-Winter's Exponential Smoothing untuk meramalkan data musiman. *Unnes Journal of Mathematics*, *6*(2), 129–142.

Berlinditya, B., & Noeryanti. (2019). Pemodelan time series dalam peramalan jumlah pengunjung objek wisata di Kabupaten Gunungkidul menggunakan metode ARIMAX Efek Variasi Kalender. *Jurnal Statistika Industri dan Komputasi*, *4*(01), 81–88.

Cui, K., & Jing, X. (2019). Research on prediction model of geotechnical parameters based on BP neural network. *Neural Computing and Applications*, *31*, 8205–8215. https://doi.org/10.1007/s00521-018-3902-6

Dahiwade, D., Patle, G., & Meshram, E. (2019). Designing disease prediction model using machine learning approach. In *2019 3rd International Conference on Computing Methodologies and Communication (ICCMC)* (pp. 1211–1215). IEEE. https://doi.org/10.1109/ICCMC.2019.8819782

Damanik, E. L., Simanjuntak, D. H., & Daud, D. (2021). Cultural heritage buildings for urban tourism destinations: Portraits of Siantar, Indonesia, in the past. *F1000Research 2021*, 10, 554–562. https://doi.org/10.12688/f1000research.48027.1

Darmawan, A. K., Siahaan, D., Susanto, T. D., Hoiriyah, H., Umam, B., Hidayanto, A. N., ... & Santosa, I. (2020). Hien's framework for examining information system quality of mobile-based smart regency service in Madura Island Districts. In *2020 4th International Conference on Informatics and Computational Sciences (ICICoS)* (pp. 1–5). IEEE. https://doi.org/10.1109/ICICoS51170.2020.9299015

Dengen, N., Haviluddin, Andriyani, L., Wati, M., Budiman, E., & Alameka, F. (2018). Medicine stock forecasting using Least Square method. In *2018 2nd East Indonesia Conference on Computer and Information Technology (EIConCIT)* (pp. 100–103). IEEE. https://doi.org/10.1109/EIConCIT.2018.8878563

Dharma, F., Shabrina, S., Noviana, A., Tahir, M., Hendrastuty, N., & Wahyono, W. (2020). Prediction of Indonesian inflation rate using regression model based on Genetic algorithms. *Jurnal Online Informatika, 5*(1), 45–52.

Google Maps. (n.d.). *Pulau Madura*. Retrieved from https://www.google.co.id/maps/place/Madura+Island/@-7.0569576,112.8385688,199132m/data=!3m2!1e3!4b1!4m5!3m4!1s0x2dd9d3445c8704d1:0x5a2751be1dfcce84!8m2!3d-7.0777326!4d113.287085

Haq, M. R., & Ni, Z. (2019). A new hybrid model for short-term electricity load forecasting. *IEEE Access*, *7*, 125413–125423. https://doi.org/10.1109/ACCESS.2019.2937222

Henttu-Aho, T. (2018). The role of Rolling forecasting in budgetary control systems: Reactive and proactive types of planning. *Journal of Management Control, 29*(3-4), 327–360.

Higgins-Desbiolles, F. (2018). Sustainable tourism: Sustaining tourism or something more? *Tourism Management Perspectives*, *25*(January), 157–160. https://doi.org/10.1016/j.tmp.2017.11.017

Higgins-Desbiolles, F., Carnicelli, S., Krolikowski, C., Wijesinghe, G., & Boluk, K. (2019). Degrowing tourism: Rethinking tourism. *Journal of Sustainable Tourism*, *27*(12), 1926–1944. https://doi.org/10.1080/09669582.2019.1601732

Jain, G., & Mallick, B. (2017). *A study of time series models ARIMA and ETS*. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2898968

Liengaard, B. D., Sharma, P. N., Hult, G. T. M., Jensen, M. B., Sarstedt, M., Hair, J. F., & Ringle, C. M. (2021). Prediction: Coveted, yet forsaken? Introducing a cross-validated predictive ability test in Partial Least Squares path modeling. *Decision Sciences*, *52*(2), 362–392. https://doi.org/10.1111/deci.12445

Liu, L., & Wu, L. (2022). Holt–Winters model with grey generating operator and its application. *Communications in Statistics - Theory and Methods*, *51,* 3542–3555. https://doi.org/10.1080/03610926.2020.1797804

Madsen, H. (2007). *Time series analysis*. CRC Press.

Mahmud, T., Billah, M., Hasan, M., & Roy-Chowdhury, A. K. (2021). Prediction and description of near-future activities in video. *Computer Vision and Image Understanding*, *210*(September), 1–12. https://doi.org/10.1016/j.cviu.2021.103230

Mufarroha, F. A., Tholaby, A. T., Anamisa, D. R., & Jauhari, A. (2023). The design of the Least Square method on sales of admission tickets to Madura tourism in forecasting cases. In *AIP Conference Proceedings* (Vol. 2679, No. 1). AIP Publishing.

Pertiwi, D. D. (2020). Applied exponential smoothing Holt-Winter method for predict rainfall in Mataram City. *Journal of Intelligent Computing and Health Informatics*, *1*(2), 46–49. https://doi.org/10.26714/jichi.v1i2.6330

Robial, S. M. (2018). Perbandingan model statistik pada analisis metode peramalan time series: (Studi kasus: PT. Telekomunikasi Indonesia, Tbk Kandatel Sukabumi). *Jurnal Ilmiah SANTIKA*, *8*(2), 1–17.

Scheyvens, R., & Biddulph, R. (2018). Inclusive tourism development. *Tourism Geographies*, *20*(4), 589–609. https://doi.org/10.1080/14616688.2017.1381985

Shano, L., Raghuvanshi, T. K., & Meten, M. (2020). Landslide susceptibility evaluation and hazard zonation techniques–A review. *Geoenvironmental Disasters, 7*(1), 1–19.

Sianturi, C. J. M., Ardini, E., & Sembiring, N. S. B. (2020). Sales forecasting information system using the Least Square method in Windi Mebel. *Jurnal Inovasi Penelitian*, *1*(2), 75–82. https://doi.org/10.47492/jip.v1i2.52

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. https://doi.org/10.1515/jwld-2017-0088

Tofallis, C. (2015). A better measure of relative prediction accuracy for model selection and model estimation. *Journal of the Operational Research Society, 66*(8), 1352–1362. https://doi.org/10.1057/jors.2014.103

Ton-Nu, V. (2014). *Rolling forecasts in a Beyond Budgeting environment: A case study on the use of Rolling forecasts as a management tool*. Retrieved from http://hdl.handle.net/11250/226581

Umam, B., Darmawan, A. K., Anwari, A., Santosa, I., Walid, M., & Hidayanto, A. N. (2020). Mobile-based smart regency adoption with TOE framework: An empirical inquiry from Madura Island Districts. In *2020 4th International Conference on Informatics and Computational Sciences (ICICoS)* (pp. 1–6). https://doi.org/10.1109/ICICoS51170.2020.9299025

Yang, C. H., Wu, C. H., & Hsieh, C. M. (2020). Long short-term memory recurrent neural network for tidal level Forecasting. *IEEE Access*, *8*, 159389–159401. https://doi.org/10.1109/ACCESS.2020.3017089

Yang, F., Li, M., Huang, A., & Li, J. (2014). Forecasting time series with genetic programming based on Least Square method. *Journal of Systems Science and Complexity*, *27*, 117–129. https://doi.org/10.1007/s11424-014-3295-2

APPENDICES

Figure 3 Time Series Plot of Observed Average Monthly Visitor Numbers from 2015 to 2019

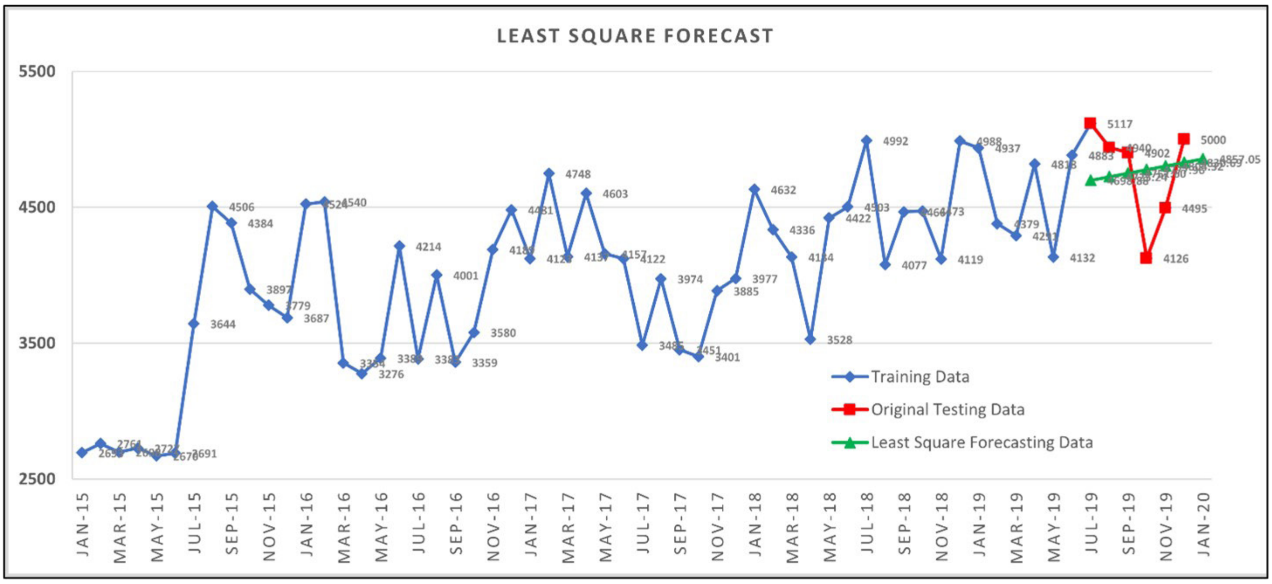


Figure 4 Least Square Forecasting Plot

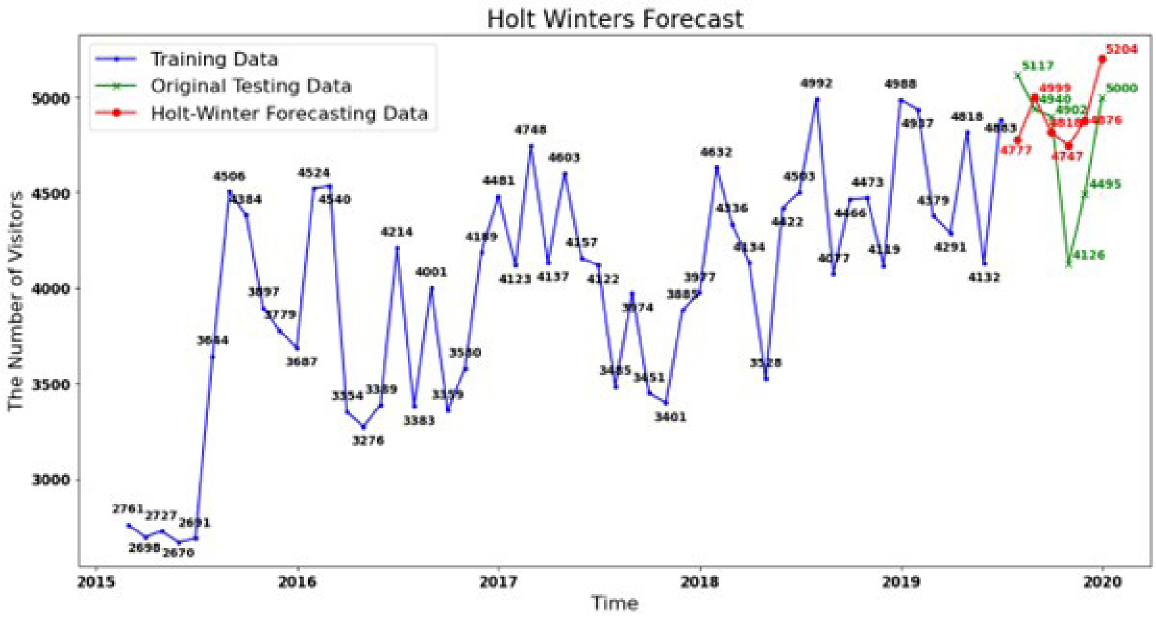


Figure 5 Holt-Winter Model Forecasting Plot

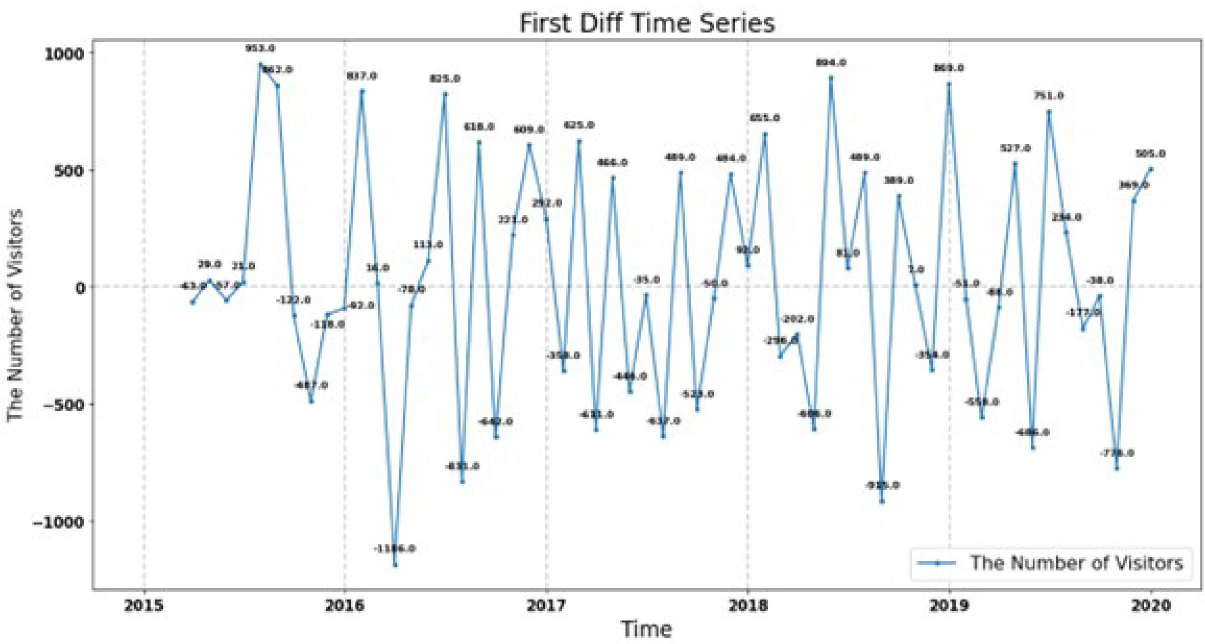
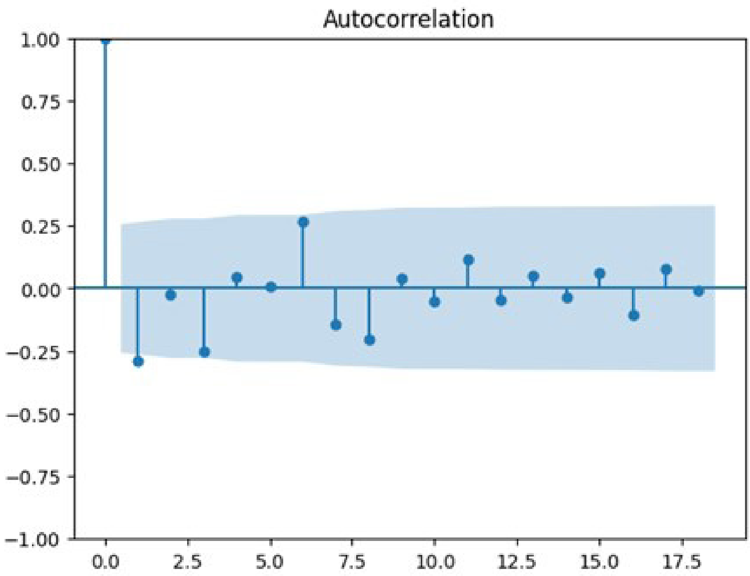
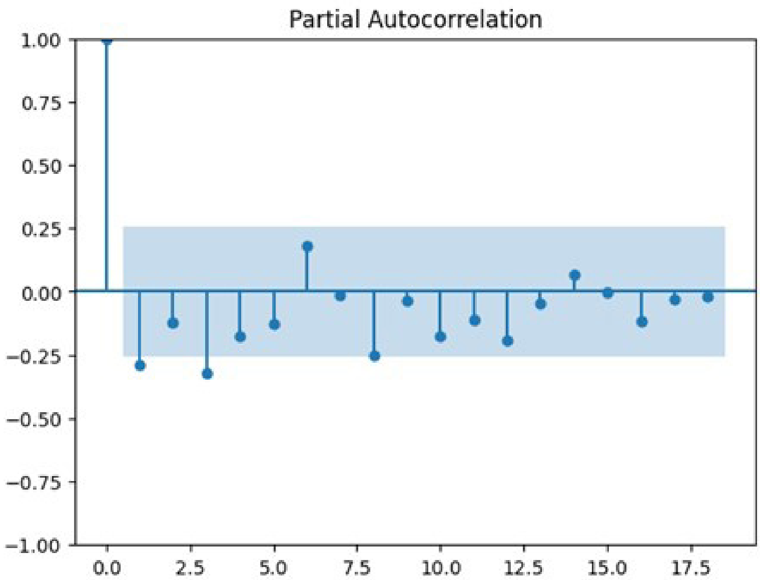


Figure 6 First Differencing Time Series Plot



(a)



(b)

Figure 7 Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) Plots

```
 ARIMA(0,1,0)(1,1,1)[12]             : AIC=inf, Time=0.22 sec
 ARIMA(0,1,0)(0,1,0)[12]             : AIC=650.130, Time=0.03 sec
 ARIMA(1,1,0)(1,1,0)[12]             : AIC=634.712, Time=0.17 sec
 ARIMA(0,1,1)(0,1,1)[12]             : AIC=inf, Time=0.30 sec
 ARIMA(1,1,0)(0,1,0)[12]             : AIC=646.133, Time=0.06 sec
 ARIMA(1,1,0)(2,1,0)[12]             : AIC=634.802, Time=0.77 sec
 ARIMA(1,1,0)(1,1,1)[12]             : AIC=inf, Time=1.31 sec
 ARIMA(1,1,0)(0,1,1)[12]             : AIC=inf, Time=0.57 sec
 ARIMA(1,1,0)(2,1,1)[12]             : AIC=inf, Time=1.36 sec
 ARIMA(0,1,0)(1,1,0)[12]             : AIC=640.349, Time=0.13 sec
 ARIMA(2,1,0)(1,1,0)[12]             : AIC=636.707, Time=0.24 sec
 ARIMA(1,1,1)(1,1,0)[12]             : AIC=636.699, Time=0.25 sec
 ARIMA(0,1,1)(1,1,0)[12]             : AIC=635.157, Time=0.16 sec
 ARIMA(2,1,1)(1,1,0)[12]             : AIC=inf, Time=0.49 sec
 ARIMA(1,1,0)(1,1,0)[12] intercept   : AIC=636.212, Time=0.19 sec

Best model:  ARIMA(1,1,0)(1,1,0)[12]
Total fit time: 6.265 seconds
```

Figure 8 ARIMA Best Model

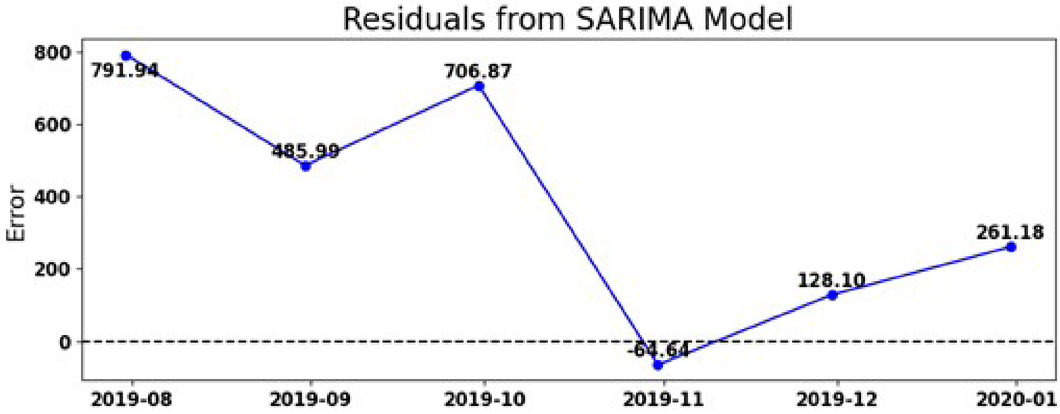


Figure 9 SARIMA Residual Plot

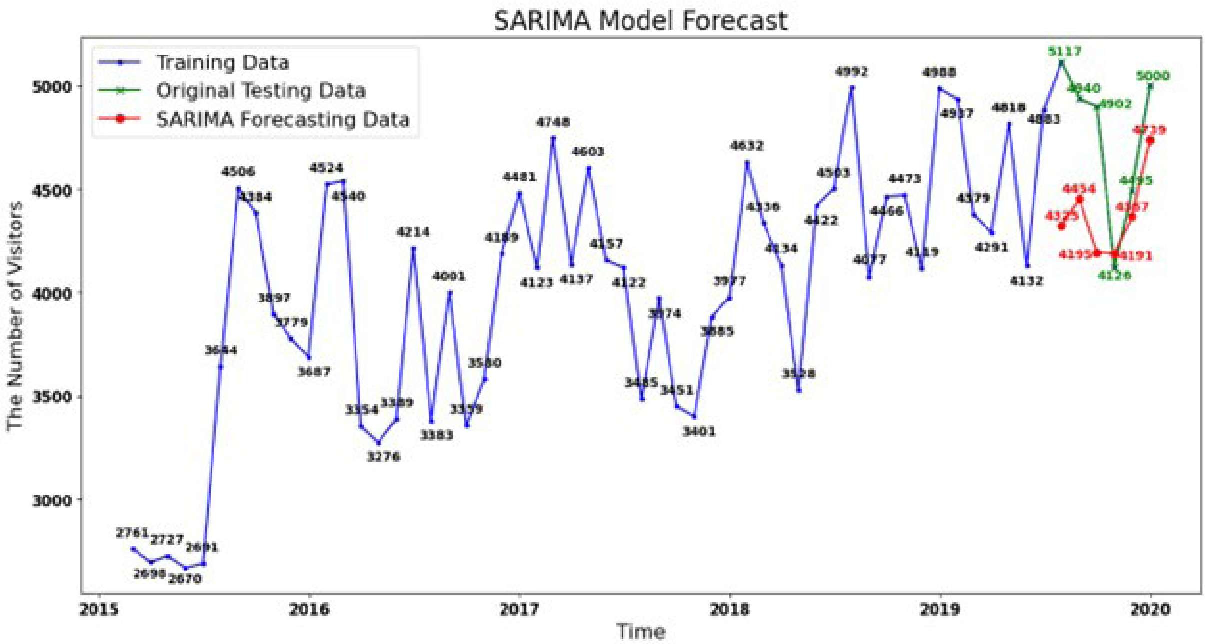


Figure 10 SARIMA Model Forecasting Plot

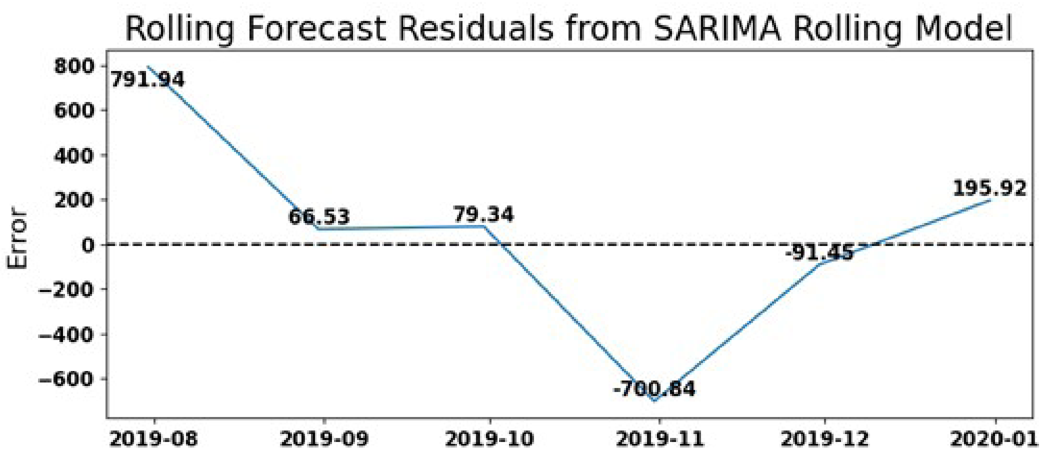


Figure 11 Rolling Residual Plot

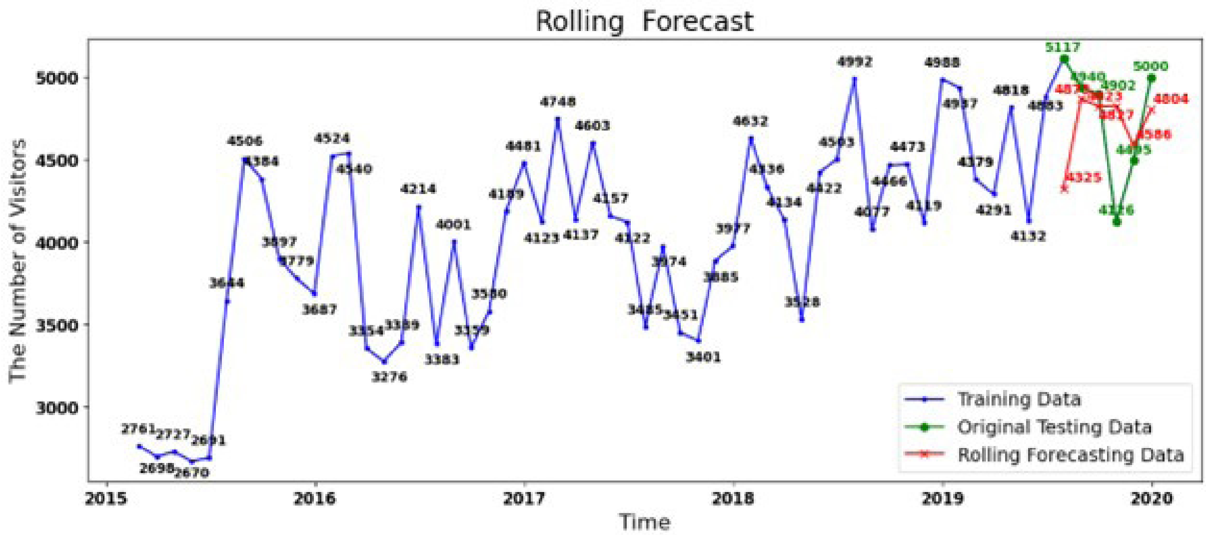


Figure 12 Rolling Model Forecasting Plot

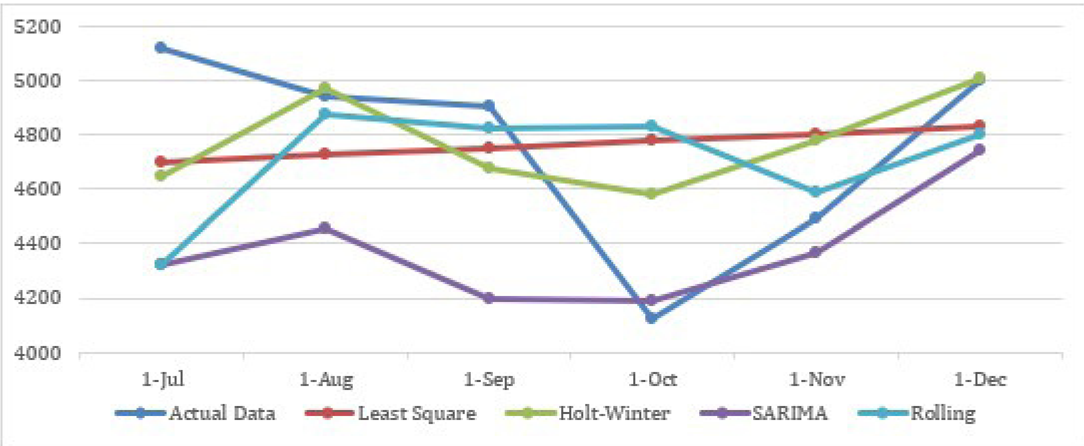


Figure 13 Comparison of Actual Data with Forecasting Data