

Conference Paper

Prediction of the Number of Foreign Tourist Arrival in Indonesia Halal Tourism Entrance using Simultaneously Fourier Series Estimator

M. Fariz Fadillah Mardianto¹, Eko Fajar Cahyono², Lailatus Syarifah³, and Putri Andriani³

¹Study Program of Statistics, Department of Mathematics, Airlangga University, Surabaya, Indonesia

²Department of Islamic Economic, Airlangga University, Surabaya, Indonesia

³Bachelor Student in Study Program of Statistics, Department of Mathematics, Airlangga University, Surabaya, Indonesia

Abstract

Indonesia is the largest Muslim country in terms of population and area. Thus, it is natural that the halal tourism develops in Indonesia. In 2015, the Indonesian Ministry of Tourism established three provinces in Indonesia as halal tourism destinations. The three provinces are Nangroe Aceh Darussalam, West Sumatra, and West Nusa Tenggara. Previously, Central Java, East Java, West Java and Banten provinces have developed Islamic tourism areas. Islamic tourism in Indonesia contributes to increasing national foreign exchange reserves, and empowering communities around tourism objects. Therefore, prediction the number of foreign tourist arrival in Indonesia halal tourism entrance is important to make a plan for determine next regulation and target satisfied. The entrance that be used is Kualanamu, Minangkabau, Soekarno Hatta, Adi Sucipto, Juanda, and Lombok Praya International Airport. Prediction that be done simultaneously because Statistically there are correlations from the data in each entrance. The research method used is nonparametric regression based on Fourier series estimator simultaneously, because the data does not form a certain pattern and it is periodic. The results obtained from this study indicate that estimator has satisfied the goodness of model, so it can be used to make prediction.

Corresponding Author:

M. Fariz Fadillah Mardianto

m.fariz.fadillah.m@fst.unair.ac.id

Received: 10 February 2019

Accepted: 14 March 2019

Published: 28 March 2019

Publishing services provided by
Knowledge E

© M. Fariz Fadillah Mardianto et al. This article is distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use and redistribution provided that the original author and source are credited.

Selection and Peer-review under the responsibility of the ICIEBP Conference Committee.

1. Introduction

The tourism sector plays an important role in the economy of a country, because this sector is a major contributor of foreign exchange and absorbent manpower in a country [1]. Indonesia is a country with great tourism potential. Some tourist destinations have become trademarks for international travelers. Some examples of tourism success such as Bali Island, Lake Toba, Mount Bromo in East Java, Raja Ampat beach in Papua, and Komodo Island in East Nusa Tenggara. Tourism's contribution to the economy is real.

 **OPEN ACCESS**

Based on the World Economic Forum (WEF) report, Indonesia's Travel and Tourism Competitiveness Index (TTCI) increases from 70th in 2013 to 50th in 2015, and 42th rank in 2017. Indonesia is among the 20 countries that have the fastest growing tourism growth. In fact, in the ASEAN region of Indonesia 4th rank after Singapore, Malaysia, and Thailand. Indonesia's tourism growth reached 25.68%, while in Southeast Asia only reached 7% [2]. According to the World Travel & Tourism Council (WTTC) data in the tourism sector which includes travel and tourists in Indonesia can produce 2.4 million new jobs. In addition, spending on foreign tourists contributes Rp.220 trillion to the economy. This sector also accounts for more than 55 percent in services [11]. Increasing tourism sector cannot be separated from the number of tourist arrivals both domestic and overseas which continues to grow. In 2017 the number of foreign tourist arrivals in the tourism sector is not in accordance with the prediction of the Tourism Ministry. This is because the number of foreign tourist arrivals in Indonesia has decreased. This decrease is caused by several factors, namely natural disasters, infrastructure damage, and others [3].

Halal tourism is a thing that has been discussed a lot lately. Halal tourism is tourism that follows the principles of sharia principles, namely tourism that adheres to all norms of Islamic law derived from the Al-Qur'an and As-Sunah as guidelines. Halal tourism is a tourist destination that is well done and made a choice according to sharia perspectives because in this tourism atmosphere is sought from any prohibited contamination [12]. Meanwhile the so-called halal tourist destination is a geographical area that is in one or more administrative areas in which there are tourist attractions, religious and public facilities, tourism facilities, accessibility, and communities that are interrelated and complement the realization of a statement in accordance with the principle sharia. Halal tourism practitioners in Indonesia need knowledge and skills on how to read trends or forecasting the number of arrivals in the future. By knowing the trend of the number of tourists, halal tourism actors will try to prepare themselves as best as possible. They can start planning programs to provide facilities with careful planning. Careful planning will produce a more rigorous and positive impact in the future.

In this research, to limit the research, the forecasting of the number of foreign tourist arrivals to Indonesia only based on some entrances are Kualanamu, Minangkabau, Soekarno Hatta, Adi Sucipto, Juanda, and Lombok Praya International Airport. All of the entrances are selected based on the halal tourism destinations in Indonesia. The data pattern of the number of foreign tourist arrivals to Indonesia does not form a specific pattern or spread and periodic because the data pattern is repeated. So, it can be analyzed by using Fourier series nonparametric regression [4]. Fourier series

is a nonparametric regression model used in estimating a data pattern in the form of trigonometry. Fourier series is a trigonometric polynomial function that has a high degree of flexibility [5].

2. Literature Review

Regression analysis is an analytical technique in statistics that can be used to explain the pattern of functional relationship between response variable (y) and predictor variable (t) [5]. If the scatter plot between the variables y and t shows the form of a clear relationship pattern such as linear, then the parametric regression approach can be used but if the relationship pattern cannot be known clearly the form can be used regression approach nonparametric [6]. In the parametric regression approach the form of a regression function is assumed to have a certain form such as linear, quadratic, exponential, etc., whereas in a nonparametric regression approach it does not assume a certain form of the regression function. Nonparametric regression functions are only assumed to be smooth or contained in a particular function space, so that it has high flexibility in determining the form of its regression function [5]. There are several smoothing techniques in nonparametric regression including kernel, linear local, local polynomial, spline, Fourier series, and others. The nonparametric regression model:

$$y_i = g(t_i) + \epsilon_i \quad (1)$$

with ϵ_i is a random error with mean 0 and variance σ^2 . Here, i represents the number of observations, $i = 1, 2, \dots, n$.

Fourier series is a trigonometric polynomial function that has a high degree of flexibility. Fourier series is a curve that shows the functions of sine and cosine. With the expansion into the Fourier series, a periodic function can be expressed as the sum of several harmonic functions, the functions of the sine and cosine, including the sinusoidal functions. Fourier series is often used as an estimator for nonparametric regression curves because it is periodic and can approach data patterns. By adjusting the Fourier series formulation, the Fourier series estimator is built by adding the trend function as follows [7]:

$$g(t_i) = \frac{a_0}{2} + \omega t_i + \sum_{\lambda=1}^K a_{\lambda} \cos \lambda t_i + b_{\lambda} \sin \lambda t_i \quad (2)$$

Thus, the nonparametric regression approach with the Fourier series estimator for paired data (t_i, y_i) is as follows:

$$y_i = \frac{a_0}{2} + \omega t_i + \sum_{\lambda=1}^K a_{\lambda} \cos \lambda t_i + b_{\lambda} \sin \lambda t_i + \varepsilon_i \tag{3}$$

with $a_0, \omega, a_{\lambda}, b_{\lambda}$ is a regression parameter whose value is estimated, $\lambda = 1, 2, \dots, K$ is an oscillation parameter. Estimator for (3) is as follows:

$$\hat{y}_i = \frac{\hat{a}_0}{2} + \hat{\omega} t_i + \sum_{\lambda=1}^K \hat{a}_{\lambda} \cos \lambda t_i + \hat{b}_{\lambda} \sin \lambda t_i \tag{4}$$

Regression parameter estimator values in vector form can be obtained based on the optimization method with the Least Square (LS) approach. The results of the optimization of the LS estimator are as follows:

$$\hat{\beta} = (T^T T)^{-1} T^T y \tag{5}$$

with the description of each vector as follows,

$$T = \begin{bmatrix} 1 & t_{11} & \cos t_{11} & \dots & \cos \lambda t_{11} & \sin t_{11} & \dots & \sin \lambda t_{11} \\ 1 & t_{12} & \cos t_{12} & \dots & \cos \lambda t_{12} & \sin t_{12} & \dots & \sin \lambda t_{12} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & t_{1n} & \cos t_{1n} & \dots & \cos \lambda t_{1n} & \sin t_{1n} & \dots & \sin \lambda t_{1n} \end{bmatrix}$$

and $y = \begin{bmatrix} y_1 & y_2 & \dots & y_n \end{bmatrix}^T$

The goodness indicators that often be used in nonparametric regression is Mean Square Error (MSE), Generalized Cross Validation (GCV), and determination coefficient (R^2). All of goodness indicators can be applied to Fourier series estimator in nonparametric regression. In nonparametric regression based on Fourier series estimator, an optimal oscillation parameter (λ) is determined. In determining optimal oscillation parameter can be used GCV formula. GCV often be used because have asymptotically optimal properties [8]. For determining an optimal oscillation parameter can be seen based on the smallest GCV value. The formula of GCV given as follows:

$$GCV(\lambda) = \frac{MSE(\lambda)}{(n^{-1} \text{trace}(I - A(\lambda)))^2} \tag{6}$$

with $MSE(\lambda) = n^{-1} y^T (I - A(\lambda))^T (I - A(\lambda)) y$ and $A(\lambda)$ is a hat matrix which consists of an estimator of a nonparametric regression model, y is a vector consisting of response variable data, I is an identity matrix [9]. By choosing of an optimal oscillation parameter

will give impact to produce a determination coefficient with high value, or approximate to 100%. The determination coefficient formula given as follows:

$$R^2 = \frac{(\hat{y} - \bar{y})^T (\hat{y} - \bar{y})}{(y - \bar{y})^T (y - \bar{y})} \quad 0 \leq R^2 \leq 1 \quad (7)$$

with \hat{y} is a vector that include of estimation result for all of subjects, and \bar{y} is a vector that include mean value for each subject. The best model that can be used for prediction met he goodness of criteria. The goodness of criteria is the smallest GCV value for an optimal oscillation parameter, the smallest MSE value, and the big of determination coefficient value [9].

3. Data, Variables, and Procedure

The in-sample data used is the number of foreign tourist arrivals to Indonesia through six entrances like in Kualanamu, Minangkabau, Soekarno Hatta, Adi Sucipto, Juanda, and Lombok Praya International Airport per month from January 2014 to February 2017 obtained from the website of Indonesian Central Bureau of Statistics and Ministry Tourism Indonesia. The data is used because the data is the latest data and 2018 data has not been released by the Indonesian Central Bureau of Statistics and Ministry Tourism Indonesia.

In this research, the response variable is the number of foreign tourist arrivals to Indonesia through six entrances and the predictor variable is the time in units of months from January 2014 to February 2017 which is estimated to affect the number of tourist arrivals to Indonesia. Researchers tested using nonparametric regression analysis method because the initial data plot did not form a specific pattern or spread, and used the approach of Fourier series estimator because the fluctuated data pattern. While the out-sample data is the number of foreign tourist arrivals per month period March 2017 to November 2017. The steps used for this data analysis are done with the help of statistical OSS-R software. The procedure of data analysis uses Fourier series estimator in nonparametric regression as follows [7]:

1. Enter the data pair for each entrance (y_i, t_i) , $i = 1, 2, \dots, n$.
2. Presenting scatter plots for the number of tourist and time in each entrance.
3. Determine the optimal oscillation parameters based on the smallest GCV using equation and determine value (λ) optimal.
4. Choose oscillation parameters based on the smallest GCV and consider the selected model.

5. Determine the goodness of the indicator for the selected oscillation.
6. Forecast the data on the number of foreign tourist arrivals to Indonesia per month period March 2017 to November 2017 based on the best model by entering the value of oscillation parameters (λ) that have been obtained and the number of forecasting desired for the coming months.

4. Result and Discussion

In this section is divided to three subsection. The first subsection about the descriptive statistics that presents the characteristics of tourist arrival in Indonesia based on collected data. The second subsection about the Fourier series method that be used to of the number of foreign tourist arrivals in Indonesia based on training data. The third subsection about the prediction of the number of foreign tourist arrivals in Indonesia with the selected model, that satisfies the goodness of model criteria based on testing data.

4.1. Characteristics of tourist arrival in Indonesia

TABLE 1: Statistics Descriptive of Tourist Arrival in Indonesia.

Entrance	Maximum Number of Foreign Tourists		Minimum Number of Foreign Tourists	
	Month	Average	Month	Average
Soekarno-Hatta,	August 2017	297.625	June 2015	18.752
Kualanamu	December 2014	29.243	January 2014	10.704
Juanda	August 2017	27.377	September 2015	1.886
Minangkabau	December 2014	6.563	May 2015	4
Lombok Praya	August 2017	12.656	June 2015	441
Adi Sucipto	August 2017	16.433	December 2015	1.001

The largest number of foreign tourists arriving is through the entrance of Soekarno-Hatta, followed by the entrance of Kualanamu, Juanda, Adi Sucipto, Lombok, and the lowest is through the entrance of Minangkabau. At the entrance of Soekarno-Hatta, the number of foreign tourists decreased in June 2015 was 18.752, while the increase in August 2017 as much as 297.625. At the entrance of Kualanamu the number of foreign tourists decreased in January 2014 as much as 10.704, while having an increase in December 2014 of 29.243. At the entrance of Juanda the number of foreign tourists decreased in September 2015 as much as 1.886, while have an increase in August 2017

of 27.377. At the entrance of Adi Sucipto the number of foreign tourists decreased in December 2015 as much as 1.001, while having an increase in August 2017 of 16.433. At the entrance of Lombok the number of foreign tourists decreased in June 2014 as much as 441, while having an increase in August 2017 of 12.656. While at the entrance of Minangkabau the number of foreign tourists decreased in May 2015 as much as 4, and having an increase in December 2014 as many as 6.563.

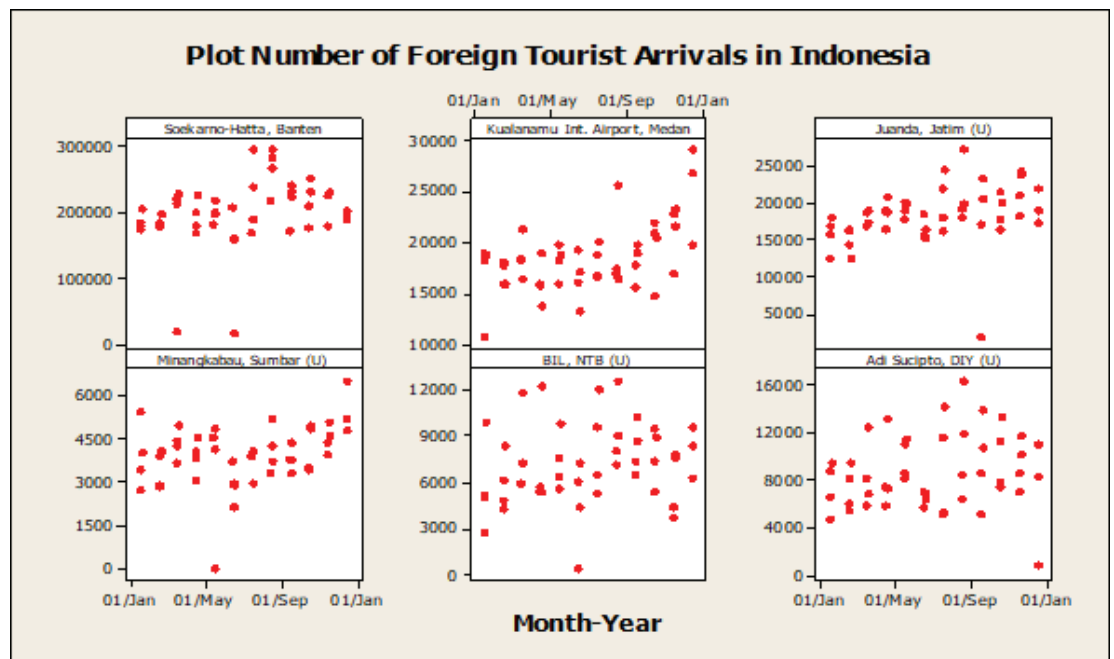


Figure 1: Plot Number of Foreign Tourist Arrivals in Indonesia.

Based on Figure 1, it can be seen that the data regarding the number of foreign tourist arrivals at each entrance has a pattern of spread data, so it is suitable to use nonparametric regression with a Fourier series estimator.

4.2. The estimation result

The nonparametric regression estimator of the Fourier series has an oscillation parameter (λ). Determining the value of the optimal λ is obtained by looking at the value of λ when the GCV value is minimum. Based on Figure 1 the λ optimum value for the entrance of Soekarno Hatta, Juanda, and Lombok Praya International Airport is 1, the entrance of Kualanamu and Minangkabau is 3, and the last entrance Adi Sucipto is 2. With the GCV value are presented in the following table:

Based on the results of calculations using R software, the parameter values in the model can be known for each entrance. In the data of the number of foreign tourist

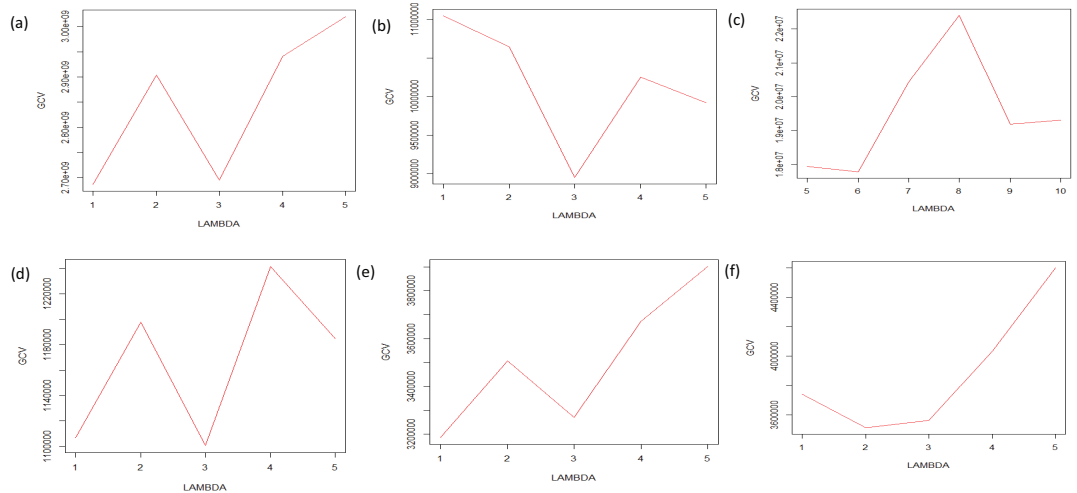


Figure 2: Plot Number GCV Plot and λ on Data on the Number of Foreign Tourist Arrivals (a) at the Soekarno-Hatta Entrance, (b) at the Kualanamu Entrance, (c) at the Juanda Entrance, (d) at the Minangkabau Entrance, (e) the Lombok entrance, and (f) the Adi Sucipto entrance.

TABLE 2: GCV Value of Each Entrance.

Entrance	GCV Minimum
Soekarno-Hatta	2687164618
Kualanamu	8956488
Juanda	17788724
Minangkabau	1100834
Lombok Praya International Airport	3187072
Adi Sucipto	3512589

arrivals to Indonesia from January 2014 to February 2017, obtained data modelling for Soekarno–Hatta entrance as follows:

$$\hat{y} = 58168,41 + 135726,29 t_i - 318,892\cos t_i - 23638,5\sin t_i$$

The goodness of the model is supported by the relative MSE value that is equal to 5722,87 and the value of R^2 is 99,99%. For Kualanamu entrance as follows:

$$\hat{y} = 5431,191 + 12672,779 t_i + 1436,852\cos t_i + 1356,847\sin t_i - 411,2386\cos 2t_i - 1726,192\sin 2t_i - 30,18479\cos 3t_i - 1726,192\sin 3t_i$$

The goodness of the model is supported by the relative MSE that is equal to 491697,3 and the value of R^2 is 92,3%. For Juanda entrance as follows:

$$\hat{y} = 5305.134 + 12378,646t_i + 1794,174\cos t_i - 888,1193\sin t_i - 986,5665\cos 2t_i \\ -744,634\sin 2t_i - 537,2086 \cos 3t_i + 116,0539\sin 3t_i - 1056,284\cos 4t_i \\ +32,41143\sin 4t_i -127,9144\cos 5t_i +1012,665\sin 5t_i -1511,016\cos 6t_i +680\sin 6t_i$$

The goodness of the model is supported by the relative MSE that is equal to 196297,1 and the value of R^2 is 95,2%. For Minangkabau entrance as follows:

$$\hat{y} = 1164,0711 + 2716,16t_i +518,0488\cos t_i -289,6571 \sin t_i -109,8265 \cos 2t_i \\ + 191,0503\sin 2 t_i + 169,0783\cos 3t_i -568,2408 \sin 3t_i$$

The goodness of the model is supported by the relative MSE that is equal to 9195,932 and the value of R^2 is 97,8%. For Lombok entrance as follows:

$$\hat{y} = 1940,48 + 4527,78 t_i +980,8844 \cos t_i -1309,329 \sin t_i$$

The goodness of the model is supported by the relative MSE that is equal to 18467,84 and the value of R^2 is 98,9%. For Adi Sucipto entrance as follows:

$$\hat{y} = 2346,16 + 5474,37 t_i +1454,643 \cos t_i -1354,45 \sin t_i -215,6766 \cos 2t_i \\ -980,1504 \sin 2 t_i + 169,0783\cos 3t_i -568,2408 \sin 3t_i$$

The goodness of the model is supported by the relative MSE that is equal to 287,5336 and the value of R^2 is 99,9%.

The estimator based on Fourier series in nonparametric regression is the best estimator because when compared with Fourier sines series and Fourier series that have basis sines and cosines the estimator is the most parsimony or has simple estimator form. The estimator based on Fourier series in nonparametric regression has a big determination coefficient value for each entrance. So, we use Fourier series in nonparametric regression to make prediction of the number of foreign tourist arrivals in Indonesia.

4.3. The prediction result

The number of foreign tourist arrivals in Indonesia from March 2017 to November 2017 is predicted based on the best model with the Fourier series estimator using testing data that be observed at January 2014 to February 2017, the prediction results are presented in Table 2.

TABLE 3: Prediction the Number of Foreign Tourist Arrivals in Indonesia as Testing Data.

Date	Soekarno-Hatta Entrance		Kualanamu Entrance		Juanda Entrance	
	Original Data	Prediction	Original Data	Prediction	Original Data	Prediction
March 17	229485	214862.3	21335	21011.96	17477	13789.41
April 17	227240	208817	19068	19095.15	20911	21971.56
May 17	219371	215684.9	18795	19067	19871	21803.21
June 17	160705	232252.6	17195	16656.65	16539	13810.84
July 17	297194	250767.8	20134	20873.77	24655	25648.47
August 17	297625	262567.1	25800	24948.04	27377	28475.88
Sep 17	233314	262129.4	19028	19889.39	23382	22287.08
Oct 17	231616	249659.5	20523	19756.07	20154	18911.31
Nov 17	231183	230992.3	23356	23935.96	24360	28028.24
Date	Minangkabau Entrance		Lombok Entrance		Adi Sucipto Entrance	
	Original Data	Prediction	Original Data	Prediction	Original Data	Prediction
March 17	5026	4757.875	11860	9999.1	12506	13412.98
April 17	4598	4828.786	12319	10227	13239	13108.07
May 17	4155	3989.389	9908	10519.1	11564	10039.73
June 17	2945	3025.46	7303	10738.7	6588	9031.855
July 17	4136	4150.394	12130	10783.1	14197	12543.45
August 17	5246	5138.487	12656	10631.5	16433	16233.44
Sep 17	4395	4582.664	10373	10354.82	13974	15450.83
Oct 17	4870	4624.82	9025	10082.53	13405	12244.15
Nov 17	4652	4925.124	7704	9942.048	11781	11622.5

Based on Table 3, it can be seen that the prediction value is not different significantly from the original data at each entrance. This indicates that the model is made using the Fourier series estimator in nonparametric regression have good performance to predict the number of foreign tourist arrivals in Indonesia theoretically. Statistically for the Soekarno-Hatta prediction has the best criteria with MSE equals 0.0004 and R^2 approaches 99,9%, for the Kualanamu prediction has the best criteria with MSE equals 0.00037 and R^2 approaches 99,9%, for the Juanda prediction has the best criteria with MSE equals 0.1496 and R^2 approaches 97,1%, for the Minangkabau prediction has the best criteria with MSE equals 0.0008 and R^2 approaches 99,5%, for the Lombok prediction has the best criteria with MSE equals 0.0556 and R^2 approaches 89,3%, and for the Adi Sucipto prediction has the best criteria with MSE equals 0.0003 and R^2 approaches 99,9%. So, the model has met the criteria for the goodness of the model.

The results of this prediction are relevant because the difference between the original data with predictive results is not much different and can be seen from the relatively small value of MSE and MAPE.

5. Conclusion

The largest number of foreign tourists arriving is through the entrance of Soekarno-Hatta, followed by the entrance of Kualanamu, Juanda, Adi Sucipto, Lombok, and the lowest is through the entrance of Minangkabau. For make a plan and target, the number of foreign tourists in Indonesia is modeled based on the best nonparametric regression based on Fourier series estimator. The estimator satisfies parsimony model, the smallest GCV and MSE value, and the biggest determination coefficient value. Using testing data, obtained the results of prediction of the number of foreign tourists in Indonesia which is not much different from training data. The government needs to make improvements to the facilities at the tourist attractions or access to tourist attractions. With forecast the number of foreign tourists, the government is expected to predict the amount of foreign exchange that will be received by the state so that the Indonesian economy remains stable and the government can anticipate it before the occurrence of the number of foreign tourist arrivals. In addition, in the months in which there is an increase in the number of foreign tourists arriving at the three entrances above, the government can monitor and explore what factors influence the attractiveness of foreign tourists to come to Indonesia, so it can be used as a reference for improvement in sectors affecting the attractiveness.

Acknowledgements

The authors give high appreciation for Indonesia Central Bureau of Statistics who have provided data, Tourism Ministry which have received the results of the study with good response, and Airlangga University has funded this publication.

References

- [1] Song, H., & Li, G. 2010. *Tourism Demand Modelling and Forecasting how should demand be measured?*. United Kingdom: University of Surrey.
- [2] Crotti, R., & Misrahi, T. 2017. *The Travel and Tourism Competitiveness Report 2007: Paying the Way for a more Sustainable and Inclusive Future*. Geneva, Switzerland:

World Economic Forum.

- [3] Gossen, C. 1993. *Eco Culture and Survival: Sustainability vs. Habit & Behavior Sustainable Structures*. Portland, Oregon: Unpublished document.
- [4] Bilodeau, M. 1992. *Fourier Smoother and Additive Models*. The Canadian of Statistic, 3, 257–259.
- [5] Eubank, R.L. 1999. *Spline Smoothing and Nonparametric Regression 2nd Edition*. New York.
- [6] Budiantara, I.N. 2006. Spline Model with Optimal Knot. *Journal of Basic Science FMIPA Jember University*, 7 (1), 77–85.
- [7] Suslov, S.K. 2003. *An Introduction to Basic Fourier series*. Springer-Science, Arizona.
- [8] Wahba, G. 1990. *Spline Model for Observational Data*. Philadelphia: SIAM.
- [9] Mardianto, M.F.F., and Budiantara I.N. 2014. Estimation of Semiparametric Regression Models Birespon with Fourier series Approach. *Proceedings of the National Seminar on Mathematics Unud, Denpasar*.
- [10] Wang, K., & Wu, C. 2003. *A Study of Cempetitiveness of International Tourism in The Southeast Asian Region*. University of Chicago Press.
- [11] Prasetyo, W.B. 2018. *The Indonesian Tourism Sector Can Produce 2.4 Million Workers*. Article posted February 3, 2018. Article Accessed September 29, 2018. Website beritasatu.com.
- [12] Djakfar, M. 2017. *Multidimensional Perspective Halal Tourism, Road Map Towards Academic & Halal Industry Development in Indonesia*. Malang. UIN Maliki Press.