Contribution Analysis of “Suroboyo Bus” in Waste Management Based on Two Form of Complete Fourier Series Estimator

M. Fariz Fadillah Mardianto*, Reynaldy Aries Ariyanto2, Raka Andriawan3, Devayanti Anugerahing Husada4

1,2,3,4Department of Mathematics, Universitas Airlangga, Surabaya, Indonesia

Article history:
Received Nov 11, 2020
Revised May 24, 2021
Accepted Jun 11, 2021

Kata Kunci:
Estimator deret Fourier, Regresi Nonparametrik, Manajemen Limbah, Suroboyo Bus

Keywords: Fourier series estimator, Nonparametric Regression, Waste Management, Suroboyo Bus

Abstract. Plastic waste is a problem that almost exists in all countries. This problem arises because of the lack of facilities that can handle the plastic waste. Suroboyo Bus is an innovation for this problem because Suroboyo Bus uses plastic bottles as payment. The purpose of this research is to predict the percentage contribution of Suroboyo Bus in handling plastic waste. The Fourier series estimator performs well for data modeling with seasonal trend patterns. This paper examines two approaches to the Fourier series. The difference between the approaches is the inclusion of the phi (π) function in the model. The result shows the goodness of fit criterion model with π function are for and 0,08% for MAPE whereas the fit criterion model without π function is 100% for and 0,07% for MAPE. In conclusion, the Fourier series model without the π function is better because the Fourier series model without the π function is more satisfy the goodness of fit criteria than the Fourier series model with the π.

How to cite:
1. Introduction

The issue of waste treatment is a crucial problem. In 2016, total garbage in Indonesia was 65,200,000 tons with population of 261,115,456 people [1]. Population in Indonesia is predicted to increase over time and predicted to increase the amount of waste. According to the Central Bureau of Statistics Indonesia, the production of waste happens massively in town. Surabaya’s city produces 9,896.78 waste every day, and Jakarta’s city produces 7,164.53 volume of waste every day. If the waste problem is not solved, it will negatively affect water pollution, land, and air. Some of the effects of the waste problem in Indonesia are the average quality pollution with heavy iron, emission of greenhouse effect from waste in 2014, and 30.26% to the emission of word greenhouse effect and it cause 1,805 floods in 2016-2017 [1]. Because of that, it is needed to reduce waste effectively. According to presidential decree number 97, released in 2017, Indonesia’s government has the target to reduce waste by 30% and handle waste 70% from Indonesia’s waste. It is still difficult because the amount of waste will increase in a row with an increase in Indonesia’s population.

Waste management is matter that mentioned in SDGs for responsible consumption and production category. In order to increase waste management performance, one of the innovations that are applied is Suroboyo Bus. Suroboyo Bus is public transportation with a payment system using garbage. People can use the Suroboyo bus for 2 hours for every ride by exchange three plastic bottles of 1.5 L or five plastic bottles of 500 ml or ten plastic glass of 240 ml. Department of Transportation says people who like the Suroboyo bus increase 15% every day and maximumly increase in the weekend. Surabaya’s government says that Suroboyo Bus can reduce plastic waste in Surabaya and make Surabaya’s people not littering. It is necessary to evaluate its function to solve the waste problem. Prior research has already reviewed the mechanism of Suroboyo Bus both as a means of public transportation and waste management effort [2]. Other previous research discusses the innovativeness and the implementation of the Suroboyo bus, which concludes that the Suroboyo bus is a good innovation as public transportation and its implementation is quite reasonable that positively accepted by the civilians [3]. However, until now, there is no further research of plastic waste management in Suroboyo Bus. As already mentioned, plastic waste is a means of payment used to enjoy this facility. In order to support the study of the Suroboyo bus, it should be research-related contributions to the Suroboyo bus in plastic waste management in Surabaya. The method to evaluate Suroboyo Bus is predicting the contribution of Suroboyo Bus. The contribution of Suroboyo Bus is measured as the total of waste loaded by Suroboyo Bus divided by estimated total waste production in Surabaya. Nonparametric regression with Fourier series estimation can be used to predict the contribution of the Suroboyo bus.

The regression model can be divided into three types: parametric, semiparametric, and nonparametric. The error term is assumed to follow specific distribution probabilities in parametric regression. This assumption is not required in nonparametric regression. Nonparametric regression uses a smoothing technique to obtain an estimate of the observation value. Because of that, nonparametric regression has high flexibility in approaching the pattern in data observation [4]. One of the methods in nonparametric regression is the Fourier series that uses the trigonometric function. A parameter, namely λ, determines the smoothing of the regression curve in the Fourier series. In general, nonparametric regression with Fourier series estimation is used in data with unknown patterns and has a seasonal trend.

This paper discusses two approaches to the Fourier series. The difference between those approaches is the inclusion of the phi (π) function in the model. The performance of the models is tested and compared on a case study. Compare both of model Fourier
series; it is used data contribution of Suroboyo Bus. The case study involves Suroboyo Bus, which accepts bottle plastic as payment. The goal of the study is to predict the percentage of Suroboyo Bus contribution to handling plastic waste. The best model of the Fourier series is determined with smaller MAPE and larger determination coefficient.

2. Nonparametric Regression based on Fourier series Estimator

2.1 Estimator with phi (\(\pi\)) Function

Fourier series is a trigonometric function which has show sine and cosine curve. Fourier series is best used on-trend seasonal data [5-8]. Based on Takezawa [5], a regression model for observation data \((t_r, y_r)\) can be written

\[
y_r = m(t_r) + \epsilon_r, \quad r = 1, 2, ..., n
\]

Distribution of is normal with mean 0 and variant 1. Assumed that \(m(t_r) \in L_2 [a,b]\) which is Hilbert's room, so that \(m(t_r)\) can be stated as linear combination with the element base from \(L_2 [a, b]\). If \(\{x_j\}_{j=1}^\infty\) is complete orthonomic system from \(L_2 [a, b]\) so \(m(t_r)\) can be stated

\[
m(t_r) = \sum_{j=1}^\infty \beta_j x_j(t_r)
\]

According to Eubank [4]. If \(\beta_j\) is a scalar, then equation (1) became.

\[
y_r = \sum_{j=1}^\infty \beta_j x_j (t_r)
\]

Observation data is time function and periodic, then to estimate function equation (2) is used linear model which has sine function and cosine function [4]. Complex exponential function can be stated in sine function and cosine function. Complex exponential function can be written:

\[
x_j (t_r) = e^{2\pi i j t_r}
\]

Substitution equation (4) to equation (2) and become as follows,

\[
\tilde{m}(t) = \sum_{j=-\infty}^{-1} \hat{\beta}_j e^{2\pi i j t_r}
\]

with \(e^{ix} = \cos(x) + i \sin(x)\) and \(e^{-ix} = \cos(x) - \sin(x)\) Fourier series estimator for \(\tilde{m}(t_r)\) is as follows

\[
\tilde{m}(t_r) = \sum_{j=-\infty}^{\infty} \hat{\beta}_j e^{2\pi i j t_r}
\]

\[
\tilde{m}(t_r) = \sum_{j=-\infty}^{-1} \hat{\beta}_j e^{2\pi i j t_r} + \hat{\beta}_0 + \sum_{j=1}^{\infty} \hat{\beta}_j e^{2\pi i j t_r}
\]

\[
\tilde{m}(t_r) = \hat{\beta}_0 + \sum_{j=1}^{\infty} [\hat{\beta}_j (\cos(2\pi j t_r) + i \sin(2\pi j t_r)) + \hat{\beta}_{-j} (\cos(2\pi j t_r) - i \sin(2\pi j t_r))]
\]

\[
\tilde{m}(t_r) = \hat{\beta}_0 + \sum_{j=1}^{\infty} [a_j \cos(2\pi j t_r) + b_j \sin(2\pi j t_r)]
\]

Where,

\[
a_j = \frac{1}{n} \sum_{r=1}^{n} y_r \cos(2\pi j t_r)
\]

\[
b_j = \frac{1}{n} \sum_{r=1}^{n} y_r \sin(2\pi j t_r)
\]

\[
t_r = \frac{j-1}{n}
\]
Substitute equation (5) into equation (1) and become.

\[ y_r = \hat{\beta}_0 + \sum_{j=1}^{J} [a_j \cos(2\pi t_r) + b_j \sin(2\pi t_r)] + \epsilon_r \]  

(7)

with estimator for regression curve is given as follows

\[ \hat{y}_r = \hat{\beta}_0 + \sum_{j=1}^{J} [\hat{a}_j \cos(2\pi j t_r) + \hat{b}_j \sin(2\pi j t_r)] \]  

(8)

2.2. Estimator without phi (\(\pi\)) Function

In some conditions, the form of regression’s curve can not be determined. It means some pattern can not be determined with some model of the parametric curve because it will produce high error and variance. If data assumed that the form of regression curve is unknown, it is suggested to use a nonparametric regression approach. Consider paired data, equation of nonparametric regression based on Mardianto et al., [7] given as follows:

\[ y_i = g(t_i) + \epsilon_i \]  

(9)

With \(y_i\) is respon variable, \(t_i\) is predictor variable for nonparametric regression and \(\epsilon_i\) is error of model with normal distribution. Form of \(g(t_i)\) function can not be determined and assumed with nonparametric regression function. Error of nonparametric regression assumed that identical, independent, and have normal distribution with 0 mean, and \(\sigma^2\) varians [8].

If \(g(t)\) is function which integrable and differentiable on interval \([a, a + 2L]\), so representation of Fourier series on that interval related with \(g(t)\) that have trigonometric component sine and cosine given as follows:

\[ g(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} [a_n \cos(\lambda^* t) + b_n \sin(\lambda^* t)] \]  

(10)

where

\[ \lambda^* \approx \frac{n\pi}{L} ; n = 1, 2, 3, ... \]

As for Fourier coefficient determine with formula:

\[ a_n = \frac{1}{L} \int_a^{a+2L} g(t) \cos(\lambda^* t) \, dt \]  

(11)

\[ b_n = \frac{1}{L} \int_a^{a+2L} g(t) \sin(\lambda^* t) \, dt \]  

(12)

If \(g(t)\) is even function, so Fourier coefficient \(b_n=0\). Therefore, the Fourier series is called cosine Fourier series. If \(g(t)\) can be integrable and differentiable on interval \([0, L]\), so that cosine Fourier series can be written

\[ g(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(\lambda^* t) \]  

(13)

where

\[ \lambda^* \approx \frac{n\pi}{L} ; n = 1, 2, 3, ... \]

According to Biederman et al., [9] for Fourier coefficient can be determined with formula:

\[ a_0 = \frac{2}{L} \int_0^L g(t) \, dt ; a_n = \frac{2}{L} \int_0^L g(t) \cos(\lambda^* t) \, dt \]  

(14)
If \( g(t) \) is odd function, so Fourier coefficient \( a_n = 0 \). Therefore the Fourier series called sine Fourier. If \( g(t) \) integrable and differentiable on interval \([0, L]\), so sine Fourier series given as follows:

\[
g(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} b_n \sin(\lambda^* t)
\]  

(15)

with

\[
\lambda^* \approx \frac{n\pi}{L} ; n = 1, 2, 3, ...
\]

As for Fourier coefficient determine with formula:

\[
a_0 = \frac{2}{L} \int_0^L g(t) \, dt \tag{16}
\]

\[
b_n = \frac{2}{L} \int_0^L g(t) \sin(\lambda^* t) \, dt \tag{17}
\]

Because it has periodic formula and can approach trend data, Fourier series can be used in nonparametric regression curve. Based on Bilodeau [10], by adjustment Fourier series formula constructed Fourier series in equation (10) with add trend function

\[
g(t_i) = \frac{a_0}{2} + \omega t_i + \sum_{k=1}^{\lambda} a_k \cos(k t_i) + b_k \sin(k t_i)
\]  

(18)

Therefore, equation of nonparametric regression approach with Fourier series estimator for paired data \((t_i, y_i)\) on equation (9) can be written as

\[
y_i = \frac{a_0}{2} + \omega t_i + \sum_{k=1}^{\lambda} (a_k \cos(k t_i) + b_k \sin(k t_i)) + \varepsilon_i
\]  

(19)

with \(a_0, \omega, a_k\) are regression parameters that, \(k = 1, 2, ..., \lambda\) is oscillation parameter. Value of regression parameter estimator in vector form can be determined according to optimization method with Least Square (LS) approach. The estimation form for nonparametric regression can be obtained as follows:

\[
\hat{y}_i = \frac{\hat{a}_0}{2} + \hat{\omega} t_i + \sum_{k=1}^{\lambda} (\hat{a}_k \cos(k t_i) + \hat{b}_k \sin(k t_i))
\]  

(20)

### 2.3. Selection of Oscillation Parameters

The selection of \( \lambda \) values must be carried out optimally. Determination of optimal \( \lambda \) can use Generalized Cross Validation (GCV) method. Based on [15, 16], formula of GCV can be written as follows:

\[
GCV(\lambda) = \frac{MSE(\lambda)}{(N - \text{trace}(I - A(\lambda)))^2}
\]  

(21)

where

\[
MSE(\lambda) = N^{-1} y^T (I - A(\lambda)) W (I - A(\lambda)) y
\]  

(22)

The GCV value depends on Mean Square Error (MSE) value because the numerator of GCV formula is the MSE formula. Measurement of goodness model is determined by the value of the determination coefficient \((R^2)\) which shows the percentage contribution of the predictor variable to the response variable.

The best model that can be used for prediction needs to pass the goodness of criteria. The goodness of criteria is the smallest GCV value for an optimal oscillation.
parameter, the smallest MSE value, and the enormous determination coefficient value [15, 16].

2.4. Measure of Goodness of Model

The goodness of model can be measured by Mean Absolute Percentage Error (MAPE). MAPE is calculated using the absolute error in each period divided by the observed value that is evident for that period [13]. The formula of MAPE can be written:

\[
MAPE = \frac{\sum_{i=1}^{n} |Y_i - \bar{Y}_i|}{\bar{Y}_i} \times 100\% 
\]

(23)

The goodness of model also can be measured by the determination coefficient \( R^2 \). The determination coefficients determine how much influence between independent variable and dependent variable [19]. Formula of \( R^2 \) can be written

\[
R^2 = 1 - \frac{\sum_{i=1}^{n} (Y_i - \bar{Y}_i)^2}{\sum_{i=1}^{n} (Y_i - \bar{Y})^2}
\]

(24)

3. Research Method

3.1. Data Source

Data collected from Surabaya government office that handling sanitary and green open space in Surabaya city. The obtained data is a time-series data of contribution of Suroboyo Bus in waste management measured as the total of waste loaded by Suroboyo Bus divided by estimated total waste production in Surabaya. The data start from May 1, 2018 to January 31, 2019. It uses data from 2018 as in sample and data from 2019 as out sample.

3.2. Data Analysis Procedure

The procedure in data analysis that is related to estimation contribution of Suroboyo bus with nonparametric regression used complete Fourier series estimator for both models from equation (7), and equation (19) was given as follows:

a. Study literature related to the contribution of Suroboyo bus and its relationship to the related predictor variables.

b. Perform descriptive statistics for each variable based on minimum, maximum, and average values.

c. Determine the GCV value for each model of nonparametric regression with Fourier series estimation that is used based on data.

d. Choose the smallest GCV value for every Fourier series used to determine oscillation parameter optimally, and calculate MAPE using equation (23) and calculate using equation (24).

e. Comparing two Fourier series estimators to determine the best model based on the smallest MAPE value and the largest \( R^2 \).

4. Result and Discussion

Suroboyo Bus was public transportation with a payment system using garbage. Every day, from Suroboyo Bus payment system, waste is collected and distributed in a garbage bank. The contribution of Suroboyo Bus from May 1, 2018, until January 31, 2019, can be explained in Figure 1.
Figure 1. Graph of Contribution of Suroboyo Bus

From Figure 1, in May 2018 until December 2018, it increased slowly. The highest contribution of Suroboyo Bus in October 2018 with 1.1%, and the lowest contribution in June 2018 with 0% (no operation). The data had an increased trend pattern and had a repetitive pattern. So, the data had a seasonal pattern, and the Fourier series estimator was suitable to estimate the contribution percentage of Suroboyo bus with nonparametric regression using the Fourier series estimator.

The results of the optimal GCV value that be calculated from R software used training data are presented in Table 1.

<table>
<thead>
<tr>
<th>Model</th>
<th>λ</th>
<th>GCV</th>
</tr>
</thead>
<tbody>
<tr>
<td>With π function</td>
<td>38</td>
<td>0.01836</td>
</tr>
<tr>
<td>Without π function</td>
<td>122</td>
<td>5.1994 x 10^{-13}</td>
</tr>
</tbody>
</table>

Based on the result in Table 1, the minimum GCV value for nonparametric regression using Fourier series with π function was 0.01836 with λ equal to 38 was chosen. For nonparametric regression using the Fourier series without π function, the minimum GCV value was 5.1994 x 10^{-13} with equal to 122 was chosen. For nonparametric regression using Fourier estimation with π model’s, obtained model Fourier series in nonparametric regression as follows

\[
\hat{y}_i = \hat{\beta}_0 + \hat{a}_1 \cos(2\pi t_r) + \hat{b}_1 \sin(2\pi t_r) + \cdots + \hat{a}_{38} \cos(76\pi t_r) + \hat{b}_{38} \sin(76\pi t_r)
\]  

(25)

where

\[
t_r = \frac{t_{i-1}}{\pi}
\]

Based on the results of, the parameter values in the nonparametric regression using Fourier series estimator with π model’s can be written as follows

\[
\hat{y}_i = 0.33079 - 0.05716 \cos(2\pi t_r) - 0.076012 \sin(2\pi t_r) - \cdots - 0.02365 \sin(76\pi t_r)
\]

(26)

For nonparametric regression using Fourier series without π, obtained model as follows:
\begin{equation}
\hat{y}_i = \frac{\hat{a}_0}{2} + \hat{\omega}_i t_i + \hat{a}_1 \cos(t_i) + \cdots + \hat{a}_{122} \cos(122t_i) + \hat{b}_1 \sin(t_i) + \cdots + \hat{b}_{122} \sin(122t_i)
\end{equation}

Based on the results, the parameter values in the nonparametric regression using Fourier series estimator without \( \pi \) function model’s can be written as follows

\begin{equation}
\hat{y}_i = 0.40293 - 0.00115 t_i + 0.001581 \cos(t_i) + \cdots + 0.02863 \sin(t_i) + \cdots + 0.0033 \sin(122t_i)
\end{equation}

Figure 3 showed comparison between data estimation using Fourier series estimation without \( \pi \) function, and data estimation using Fourier series estimation with \( \pi \) function.

![Figure 3. Graph of comparison between actual data and result](chart.png)

Based on Figure 3, estimation used Fourier series estimator without \( \pi \) function closed to actual data than used Fourier series estimator with \( \pi \) function. It means Fourier series estimator without \( \pi \) function was better than Fourier series estimator with \( \pi \) function for data because Fourier series estimator without \( \pi \) function was smoother than Fourier series estimator with \( \pi \) function. It was seen from graph for a model that closed to actual data. \( R^2 \), and MAPE from Fourier series estimator with \( \pi \) function in a row was 99.96\% for \( R^2 \), 0.08\% for MAPE. For Fourier series estimator without \( \pi \) function, the result of \( R^2 \), MAPE in a row was 100\% for \( R^2 \), and 0.07\% for MAPE. Based on the result, Fourier series estimator without \( \pi \) function was better than Fourier series estimator without \( \pi \) function to estimate contribution percentage of Suroboyo bus because Fourier series estimator without \( \pi \) functions was more flexible than Fourier series estimator with \( \pi \) functions.

5. Conclusions

Suroboyo Bus had a chance to solve the plastic waste problem in a big city because the Suroboyo Bus can increase waste management percentage. From both models, Fourier series estimator without \( \pi \) function was more suitable than Fourier series estimator with \( \pi \)
function because MAPE’s Fourier series estimator without π function was smaller than Fourier series estimator with π function. Furthermore, determination coefficient’s Fourier series estimator without π function was higher than the Fourier series estimator with π function. However, if we used the Fourier series estimator without π function, it had an extended model because optimum λ in the Fourier series estimator without π function was equal to 122. This value was enormous, so it would difficult to interpret the model. Fourier series estimator with π function was more parsimony than Fourier series estimator without π function because optimum λ in Fourier series estimator with π function was smaller than Fourier series estimator without π function.

There is still low research of Suroboyo Bus as a research object, especially about its ability to handle plastic bottle waste and waste management flow after the waste has been collected. Therefore, it is necessary to have further research to become input and feedback for related services, particularly the department of transportation and also Surabaya government office that handling sanitary and green open space in Surabaya city.

References


M. Fariz Fadillah Mardianto, R. Aries Ariyanto, Raka Andriawan, and Devayanti A. Husada,
Contribution Analysis of “Suroboyo Bus” in Waste Management Based on Two Form of Complete Fourier Series Estimator


