Modeling on Electricity Consumption's Average of Households Group in Surabaya with Nonparametric Approach based on Fourier Estimator

Eko Tjahjono, Sediono, M. Fariz Fadillah Mardianto and Ajeng Novy Lestari
Department of Mathematics, University of Airlamgga, Surabaya, Indonesia

Keywords: Electricity, Temperature, Nonparametric Regression, Fourier Series.

Abstract: Demand of electrical power increase in line with the number of populations, technological advances, and information. Household groups are the biggest users of electricity every year. The case study in this research that be used about electricity usage in Surabaya city. Surabaya is the capital of the province with very high electrification that equal to 99,79%. This proves that Surabaya can be used as an indicator of electricity. One factor that influence the high usage of electricity is temperature. The higher the air temperature caused the use of electrical equipment is increasing so that the used of electricity increases. Prediction of electrical used is needed to anticipate electricity supply as well as anticipation action against power outage which can disturb human activities. In this study, electricity consumption of household group in Surabaya was observed monthly, then predicted using nonparametric regression method based on Fourier series estimator. One of the advantages of the Fourier series is that it can overcome data that has repeating pattern. This study results a minimum GCV value when the oscillation parameter reaches 11. The chosen model based on minimum GCV has met the model goodness criteria. Then, the selected model used to predict the household electricity usage in Surabaya for some time ahead.

SCIENCE AND TECHNOLOGY PUBLICATIONS

1 INTRODUCTION

Electricity is one of the most important and vital energy for human life, so it cannot be separated every day. Nowadays the need for electric energy is increasing along with the increase of population and technological progress and information. Currently electrical energy has been classified as the basic needs of an area used by five groups of power users. Household groups are the largest group of electric energy users in each year. The average population growth of Surabaya City in 2007 was 0,1172 with a population of 2.720.156 people and the number of households was 755.914. The condition of electrification ratio in 2007 is very high, amounting to 99,79%. Thus, this is the basic research why use the city of Surabaya.

Determination of the future of electrical energy is influenced by various factors. Weather changes cause changes to consumer comfort and affect the use of equipment including electrical equipment. The higher temperature of the air will cause the use of electrical equipment is greater so that electricity consumption is increasing.

In this research, electricity consumption's average of households group in Surabaya is modeled using nonparametric regression approach based on Fourier series estimator. One of the advantages of nonparametric regression approach using Fourier series can overcome data having trigonometric spreading, in this case sine and cosine. The data patterns corresponding to the Fourier series approach are repetitive data patterns. Research on the Fourier series includes Bilodeau (1992), Tripena and Budiantara (2006).Tjahjono (2009) studied Fourier series estimators on nonparametric regression, and Tjahjono et al., (2018) developed an estimated nonparametric regression model of the Fourier bi-response series.

2 FOURIER SERIES ESTIMATOR

Fourier series is a trigonometric polynomial function that has a high flexibility. It is because the pattern of

Tjahjono, E., Sediono, ., Mardianto, M. and Lestari, A.
Modeling on Electricity Consumption's Average of Households Group in Surabaya with Nonparametric Approach based on Fourier Estimator.
DOI: 10.5220/0008518501470150
In *Proceedings of the International Conference on Mathematics and Islam (ICMIs 2018)*, pages 147-150
ISBN: 978-989-758-407-7
Copyright © 2020 by SCITEPRESS – Science and Technology Publications, Lda. All rights reserved

data can determine the shape of regression curve based on estimators in the nonparametric regression, and it is engaging trigonometric function (Prahutama, 2013). Fourier series estimators are generally used when the data used are investigated unknown patterns and there is a tendency of seasonal patterns (Tripenaand Budiantara, 2006 and Bilodeau, 1992). Suppose given the observation data $(x(t_r), y_r)$ following the regression model

$$y_r = m(x(t_r)) + \varepsilon_r \; ; \qquad r = 1,2, \ldots, n \qquad (1)$$

and functions $m(x(t_r))$as follows:

$$m(x(t_r)) = \sum_{j=-\lambda}^{\lambda} \beta_j x(t_r) \qquad (2)$$

withλ is an integer, then equation (2) become :

$$y_r = \sum_{j=-\lambda}^{\lambda} \beta_j x(t_r) + \varepsilon_r; \qquad r = 1,2, \ldots, n \qquad (3)$$

When the observation data is a function of time and shows periodic phenomena then in estimate m function used linear model which includes the sine and cosine functions. This study uses temperature(x)as the predictor variable in the time function(t_r), so$x(t_r)$is the temperatureat thet_r time.

$$x(t_r) = e^{2\pi i t_r} \qquad (4)$$

By taking *n* paired samples$(x(t_r), y_r)$which fulfill equation (4) then obtained equation as follows:

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta}_{\boldsymbol{\lambda}} + \boldsymbol{\varepsilon} \qquad (5)$$

with

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \boldsymbol{X} = \begin{bmatrix} e^{2\pi i t_1} & e^{2\pi i t_1} \cdots & e^{2\pi i t_1} \\ e^{2\pi i t_2} & e^{2\pi i t_2} \cdots & e^{2\pi i t_2} \\ \vdots & \vdots \quad \vdots & \vdots \\ e^{2\pi i t_n} & e^{2\pi i t_n} \cdots & e^{2\pi i t_n} \end{bmatrix},$$

$$\boldsymbol{\beta}_{\boldsymbol{\lambda}} = \begin{bmatrix} \beta_{-\lambda} \\ \beta_{-\lambda+1} \\ \vdots \\ \beta_{\lambda} \end{bmatrix}, \text{ and} \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

3 RESEARCH METHODOLOGY

The data used in this research is secondary data of household usage for 2011-2016 in Surabaya, there are 63 observations that be obtained from PT. PLN, branch of North Surabaya, and temperature data of Surabaya City obtained from BMKG, station of Perak Surabaya.

The following steps data analysis used to answer the purpose of this study as follows:

1. Describe the data of electricity consumption in Surabaya
   a) Entering variable response in the form of data electricity consumption and predictor variable of temperature
   b) Creating plots for data on electricity consumption and temperature
   c) Calculate and create tables for the mean, median, mode, and variance values of data on electricity consumption and temperature
2. Model data of electricity consumption in Surabaya based on Fourier series estimator using R software as follows:
   a) Defines response variables (Y) and predictor variables (X). Specifies the function m (x (t_r)) to use Calculate the value of MSE
   b) Calculate the smoothing parameter value (λ) with GCV
   c) Entering the initial smoothing parameter value (λ) and the final finishing parameter value (λ) Based on step e selected minimum GCV value based on looping result.
   d) The value of the fining parameter corresponding to the minimum GCV value is the optimum parameter value (λ) which will be used in the next step Input the refined parameter value (λ) optimal that has been obtained Calculate the value of a_j and b_j
   e) Calculate the value of $\widehat{m}(x(t_r))$
3. Interpret data model of electricity consumption in Surabaya based on Fourier series estimator
   a) Making the results of analysis based on software output R then make a conclusion from the analysis results.
   b) Creating the table of actual electricity usage data and estimation data from the model that has been obtained based on the Fourier series criteria.

4 RESULT AND DISCUSSION

4.1 Characteristics of Electricity Usage of Household Group in Surabaya City

The data analyzed in this study is data of electricity consumption of household groups in Surabaya area from April 2011 to June 2016 taken every month in Kilowatt Hour and monthly temperature data of Surabaya City during the same period. Overall data amounted to 63 data, including data from April 2011 to December 2015 as many as 57 data used as in-sample data while in January 2016 until June 2016 as much as 6 data used as out-sample data. In-sample data will be analyzed by Fourier series method to find out the best model for electricity consumption and out-sample data used for electricity consumption

forecasting in Surabaya affected by temperature and time for several periods ahead.

4.2 Relationship between Household Electricity Usage with Temperature in Surabaya City

The variable used in this study to identify household usage in Surabaya is the monthly temperature of that city. In addition to time or period (months), temperatures can also affect the size of electricity consumption, especially in household groups. This case the temperature effect on the conductor of electricity so that it can affect the electrical current. From the results it is known that the relationship or the influence of heat on the resistance of the wire is directly proportional. This means that if the temperature is greater than the resulting resistance the greater also. This is in accordance with opinion (Suroso, 2003) that the conductor of electricity is one important component in the distribution of electric power.

4.3 Model Data of Electricity Usage of Household Group in Surabaya City

Data used in this research is secondary data of electricity usage in Surabaya city on April 2011 up to June 2015 57 data as observation data (in sample) and last 6 data as prediction data (out sample). Response variable in this research is data of electricity usage in Surabaya City and predictor variable is Surabaya Kota temperature data and time (month). Here is a plot between response variables and predictor variables.

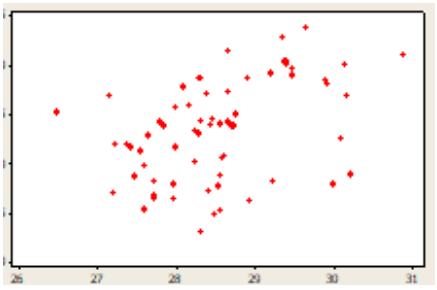

Figure 1: Scatterplot Temperature and Electricity Consumption in kWh.

Based on Figure 1 which is the average scatterplot of electricity consumption in Surabaya City against the monthly temperature of Surabaya City, it can be assumed that the data plot shows the absence of certain relationship pattern, the data pattern is periodic and shows the trigonometric distribution of sinus and cosine. So, the analysis of the regression model uses a nonparametric approach based on a Fourier sequence estimator.

The model of the relationship between the response variable and the best predictor variable is obtained by determining the optimal smoothing parameter (λ) with minimum GCV criterion.

Table 1: GCV minimum.

Lambdas	GCV	MSE	R^2
1	15,45084	13,86724	0,3778355
2	15,59415	12,97832	0,4177173
3	16,20542	12,46954	0,4405442
4	16,12573	11,43542	0,4869407
5	15,9596	10,39413	0,5336593
6	17,18157	10,23808	0,5406606
7	17,60686	9,559401	0,5711099
8	19,28306	9,49612	0,5739491
9	17,91869	7,963864	0,642695
10	10,31607	4,114997	0,8153774
11	**9,397186**	**3,343536**	**0,8499896**
12	10,59919	3,340589	0,8501218
13	10,66285	2,953697	0,86748
14	10,54307	2,544096	0,8858571

Based on Table 1 above which is a table of GCV values along with MSE and R^2 values on data of electricity consumption in Surabaya city can be seen that minimum GCV value is 9,397186 with MSE value equal to 3,343536 and R2 value equal to 0,8499896 or 85%. The next step is to make a plot between lambdas with GCV value on the data of electricity consumption in Surabaya. After obtaining the minimum GCV value and optimum lambda then do the estimation model using software R. The program is made to obtain the value of a_j as a cosine function and the value of b_j as a form of sine function derived from the Fourier derivative model estimator. From result of output program in software R obtained value of β_0 for data of electricity usage that is equal to 62,43035 and obtained value of λ, a_j and b_j as in Table 2 below:

Table 2: Value of coefficient a_j and b_j.

Lambda	a_j	b_j
1	-0,9814987	-3,984919
2	-0,2589094	-1,307973
3	0,1466272	-0,9980298
4	0,517541	-1,341785
5	-0,6823104	-1,271631
6	0,2800438	-0,4834005
7	-0,03172626	-1,164622
8	0,2975692	-0,1949733
9	1,164072	-1,307459
10	-2,721578	-0,5392087
11	-0.8037324	-0.9470675

Based on Table 2 which is the table of values of λ, a_j, and $b_j$of household electricity consumption data in Surabaya in 2011 to 2016, it can be found that household electricity consumption data model based on nonparametric regression method based on Fourier series estimator is as follows:

$$\begin{aligned}\hat{y}_r = 62{,}43035 &- 0{,}9814987\cos(2\pi x(t_r)) \\ &- 3{,}984919\sin(2\pi x(t_r)) \\ &- 0{,}2589094\cos(4\pi x(t_r)) \\ &- 1{,}307973\sin(4\pi x(t_r)) \\ &+ 0{,}1466272\cos(6\pi x(t_r)) \\ &- 0{,}9980298\sin(6\pi(t_r)) \\ &+ 0{,}517541\cos(8\pi x(t_r)) \\ &- 1{,}341785\sin(8\pi x(t_r)) \\ &- 0{,}6823104\cos(10\pi x(t_r)) \\ &- 1{,}271631\sin(10\pi x(t_r)) \\ &+ 0{,}2800438\cos(12\pi x(t_r)) \\ &- 0{,}4834005\sin(12\pi x(t_r)) \\ &- 0{,}03172626\cos(14\pi x(t_r)) \\ &- 1{,}164622\sin(14\pi x(t_r)) \\ &+ 0{,}2975692\cos(16\pi x(t_r)) \\ &- 0{,}1949733\sin(16\pi x(t_r)) \\ &+ 1{,}164072\cos(18\pi x(t_r)) \\ &- 1{,}307459\sin(18\pi x(t_r)) \\ &- 2{,}721578\cos(20\pi x(t_r)) \\ &- 0{,}5392087\sin(20\pi x(t_r)) \\ &- 0{,}8037324\cos(22\pi x(t_r)) \\ &- 0{,}9470675\sin(22\pi x(t_r))\end{aligned}$$

4.4 Interpret Model of Electricity Usage Data Usage in Surabaya City based on Fourier Series Estimator

Table 3: Value of coefficient a_j and b_j.

observations	Y	$\hat{Y}$
1	53,02	59,3564
2	55,246	54,7319
3	56,558	54,8453
4	58,317	56,8829
⋮	⋮	⋮
56	70,997	71,4272
57	72,891	66,7502

Based on data of household electrical usage that have been analyzed using software R obtained optimal GCV value at 11th lambda that is equal to 9,397186 with value of MSE equal to3,343536 and value of R^2 equal to 0,8499896 or equal to 85% which mean model have good because can represent data of 85%.

5 CONCLUSIONS

Based on the results of analysis and discussion in this study it can be concluded that any rise in temperature can cause changes in the amount of electricity consumption each month. The model of Fourier estimator series has been obtained is good with R^2value of 85% and MSE value of 3,343536 for the data in sample, while for the data out sample obtained R^2 value of 99% and MSE of0,08893838. This means the model has been good and in accordance with the data used in this study.

REFERENCES

Bilodeau, M., (1992), Fourier Smoother and Additive Models, *The Canadian Journal of* statistics, 3: 257 – 269

Tripena, A., and Budiantara, I. N., 2006, Fourier Estimator in Nonparametric Regression, *Proceeding International Conference on Natural Sciences and Applied Natural Scienes Ahmad Dahlan University*, Yogyakarta.

Tjahjono, E.,Mardianto, M. F. F, Chamidah, N., 2018, Prediction of Electricity Consumption Using Fourier Series Estimator in Bi-Response Nonparametric Regression Model, *Far East Journal of Mathematical Sciences (FJMS)*, vol. 103, Number 8, 2018, pp 1251 - 1263.

Tjahjono, E, 2009, weighted Fourier Series Estimator for Nonparametric Regression, *Thesis*, ITS, Surabaya.