

## COMPARISON OF RICE PRICE PREDICTION RESULTS IN EAST JAVA USING FOURIER SERIES ESTIMATOR AND GAUSSIAN KERNEL ESTIMATOR SIMULTANEOUSLY

**Adnan Syawal Adilaha Sadikin<sup>1</sup>, Aqil Azmi Reswara<sup>2</sup>,  
M. Fariz Fadillah Mardianto<sup>3\*</sup>, Ardi Kurniawan<sup>4</sup>**

<sup>1,2,3,4</sup>Statistics Study Program: Department of Mathematics, Faculty Sains & Technology, Universitas Airlangga  
Jl. Dr. Ir. H. Soekarno, Surabaya, 60115, Indonesia

Corresponding author's e-mail: \* [m.fariz.fadillah.m@fst.unair.ac.id](mailto:m.fariz.fadillah.m@fst.unair.ac.id)

### ABSTRACT

#### Article History:

Received: 14<sup>th</sup> March 2024

Revised: 2<sup>nd</sup> May 2024

Accepted: 16<sup>th</sup> July 2024

Published: 1<sup>st</sup> September 2024

#### Keywords:

Fourier Series;

Rice Price;

Gaussian Kernel;

MAPE;

Zero Hunger.

Extreme weather changes and the El Nino phenomenon in 2023 will cause drought, resulting in a decrease in rice production and an increase in rice prices. It has significantly impacted East Java Province as it is the most extensive rice supplier in Indonesia. This study aims to predict the price of rice with six different qualities using the Fourier series estimator and Gaussian kernel function simultaneously. The results show that the Gaussian kernel method, with a bandwidth value of 1, produces a better model with a MAPE value of 0.228259% than the sine function Fourier series method in predicting rice prices based on six different qualities. The prediction results using the Gaussian kernel function method are categorized as highly accurate because they are less than 10%. This research accelerates the realization of SDG 2 related to "Zero Hunger" through government policies to control the high price of rice in Indonesia. Recommendations that can be given through the research results include cooperation with the government, which can help access information and resources needed to manage price risks.



This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution-ShareAlike 4.0 International License.

#### How to cite this article:

A. S. A. Sadikin, A. A. Reswara, M. F. F. Mardianto and A. Kurniawan., "COMPARASION OF RICE PRICE PREDICTION RESULTS IN EAST JAVA USING FOURIER SERIES ESTIMATOR AND GAUSSIAN KERNEL ESTIMATOR SIMULTANEOUSLY," *BAREKENG: J. Math. & App.*, vol. 18, iss. 3, pp. 1963-1974, September, 2024.

Copyright © 2024 Author(s)

Journal homepage: <https://ojs3.unpatti.ac.id/index.php/barekeng/>

Journal e-mail: [barekeng.math@yahoo.com](mailto:barekeng.math@yahoo.com); [barekeng.journal@mail.unpatti.ac.id](mailto:barekeng.journal@mail.unpatti.ac.id)

**Research Article** · **Open Access**

## 1. INTRODUCTION

The fulfillment of food needs is the most essential thing in a country's human development [1]. Rice is a staple food in some Asian countries, especially Indonesia. Therefore, rice has become an important commodity in maintaining national food security in Indonesia. The availability of rice in Indonesia is a prerequisite for national food security [2]. Of the several countries in the world, Indonesia has a rice income of 80% [3]. Domestic rice production continues to increase from year to year, although there is a tendency for the growth rate to slope [4]. With the increasing population in Indonesia, the need for rice is also increasing. However, the price of rice on the market continues to rise, so many traders sell poor-quality rice [5].

Rice prices in the country continue to rise. This increase in rice prices far exceeded the Highest Retail Price (HET) set by the government and set a new record [6]. To date, the average price of rice has touched IDR 15,650/liter in all Indonesian provinces [7]. According to the government, the skyrocketing price of rice is caused by climate change, which has caused several regions to experience crop failure. Extreme climate change occurs due to the El Nino phenomenon, which affects global weather patterns, including Indonesia. According to the Meteorology, Climatology and Geophysics Agency (BMKG), El Nino in Indonesia started in June and peaked in August to September. This phenomenon is also predicted to last until early 2024. This phenomenon resulted in the retreat of the beginning of the rainy season, resulting in drought in Indonesia until the end of 2023 [8]. Drought due to El Nino can suppress the production of national rice prices and impact the increase in rice prices, which almost occurs in all countries.

The impact of rising rice prices is felt in all provinces in Indonesia, one of which is East Java. East Java has the most significant rice production in Indonesia, with a total rice production of 9.68 million tonnes and an agricultural land area of 1.7 million hectares in 2022 [9]. In September 2023, the price of rice in East Java reached IDR 16,000 per kg above the price ceiling. The rising price of rice in East Java is a concern for the community because East Java is an area that has abundant rice stocks, but rice prices are still soaring. The long chain of rice distribution also causes an increase in rice prices. It harms several parties, such as farmers and retailers [10]. As a result of the increase in the price of rice as a staple food, food stalls are faced with several problems regarding the quality of the products sold to survive and maintain consumer loyalty. The increase in rice prices will more or less affect the quality of the products sold in the food stall [11].

Therefore, it is important to analyze and predict the price of rice in East Java Province, using various qualities, to prepare market players, companies, and the government to design appropriate policies for increasing rice prices. Based on data from the Strategic Food Price Information Centre (PIHPS), the price of rice in each quality has tended to grow in the last two years. Therefore, a classical time series analysis was conducted using the kernel estimator method, one of the nonparametric methods [12]. We chose the kernel estimator method because it is flexible in modeling data patterns and does not depend on assumptions about the relationship between variables. In addition to the kernel estimator, in classical nonparametric time series analysis, the prediction method used is the Fourier series. It has flexibility in modeling data patterns whose oscillation shape is unknown [13]. A benefit of employing nonparametric regression with the Fourier series lies in its ability to address datasets exhibiting trigonometric distribution, such as sine and cosine patterns. Utilizing the Fourier series approach is particularly advantageous when dealing with repetitive data patterns, where the dependent variable values repeat for various independent variables [14].

In previous research, Widyadarma et al. [15] predicted the weekly price of shallots and common red chilies using the Radial Basis Function Neural Network (RBFNN), Elman Recurrent Neural Network (ERNN), and a combination of neural network and genetic algorithm. However, the accuracy obtained is still relatively low, around 75%. Similar research has also been carried out by Purwoko et al. [16] regarding Non-Oil and Gas Export Price Prediction in Indonesia with nonparametric regression Fourier series estimator. Prediction in Indonesia with nonparametric regression Fourier series estimator. The function used in the study is a Polynomial kernel with excellent prediction with a smoothing epsilon parameter value of 0.0266, a weighting parameter value of 0.0285, and obtained an absolute average percentage error value of 7.7513%, which gives prediction accuracy results with a Mean Absolute Percentage Error (MAPE) value of 15.26% on testing data.

Based on previous research, a study will be conducted to predict the price of rice with six different qualities by comparing the classical nonparametric Fourier series and Gaussian kernel methods. This research is important because the rice price is increasing, so it can determine the best way to overcome this. This research supports and accelerates the achievement of the 2nd SDG's point regarding the goal of zero hunger

as a form of recommendation to the Indonesian government to make policies to control rice prices, which are also influenced by natural factors.

## 2. RESEARCH METHODS

### 2.1 Data Sources

The data used in this study are secondary data, which are rice prices based on six different qualities, namely Lower Quality Rice I (BKB I), Lower Quality Rice II (BKB II), Medium Quality Rice I (BKM I), Medium Quality Rice II (BKM II), Super Quality Rice I (BKS I), and Super Quality Rice II (BKS II). The data is taken from the PIHPS website weekly from 24 February 2022 to 20 February 2024, and there are as many as 104 datasets. The data used is divided into training data and testing data. The training and testing data division in this study is 80% for training data or 84 data from 24 February 2022 to 26 September 2023. Meanwhile, the testing data used in this study is 20% or 21 from 3 October 2023 to 20 February 2024.

### 2.2 Research Variables

The variables used in this study are the time variable as the predictor variable and the rice price variable as the response variable. The multicollinearity test tests whether a regression model correlates with predictor variables (independent) [17]. Since there is only one predictor (time), multicollinearity testing of variables is unnecessary. Multicollinearity itself requires at least two predictors that correlate with each other. With only one predictor, no pair of variables can show a correlation. The description for each variable is presented in **Table 1** below.

**Table 1. Research Variables**

Variables	Description
Predictor	Time (X)
Response	BKB I (Y1), BKB II (Y2), BKM I (Y3), BKM II (Y4), BKS I (Y5), BKS II (Y6)

### 2.3 Research Method

The analysis steps in this study can be described in detail as follows:

- Collected weekly data on rice prices from the [bi.go.id/hargapangan](https://bi.go.id/hargapangan) website.
- Create a descriptive overview of weekly rice price data using time series plots and describe it.
- Perform Bartlett's Test.
- Determining training and testing data with an 80% and 20% ratio.
- Selects oscillation parameters (k) in the Fourier series and bandwidth (h) in the optimal Gaussian kernel based on minimum GCV values.
- Modeling the training data using the Fourier series approach and Gaussian kernel function.
- Predict the testing data using the Fourier series approach and Gaussian kernel function.
- Selecting the best model by comparing the prediction results of the Fourier series estimator and Gaussian kernel function based on the smallest MAPE value.

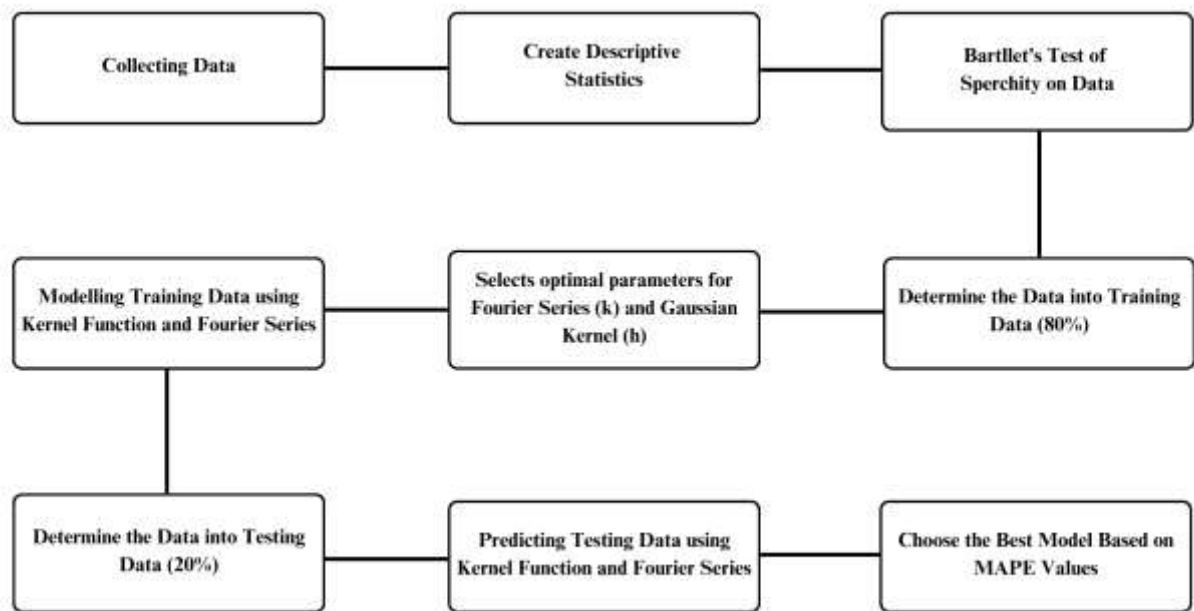


Figure 1. Flowchart research

## 2.4 Simultaneous Fourier Series Estimator

The Fourier series estimator is based on two parameters, namely the model parameter and the oscillation parameter as a representation of the bandwidth. One of the advantages of the Fourier series estimator approach is that it can handle data with patterns represented by trigonometric functions [18]. Therefore, Fourier series are characterized as trigonometric polynomial functions renowned for their great flexibility, since they embody curves representing sine and cosine functions [19]. To obtain the model parameters, the optimisation method in this study is the Weighted Least Square (WLS) method. This method is not tied to the shape of the distribution function of the error, and there is no penalty. Therefore, this method is easy to be used [20]. The complete Fourier series consists of cosine and sine functions. Given paired data  $(x_{ij}, y_{ij})$  with  $j = 1, 2, \dots, q$  then the simultaneous Fourier series equation is shown as follows..

$$\begin{aligned}
 y_{i1} &= \frac{a_{01}}{2} + \omega_1 x_{i1} + \sum_{k=1}^k (a_{k1} \cos kx_{i1} + b_{k1} \sin kx_{i1}) + \varepsilon_{i1}; \varepsilon_{i1} \sim N(0, \sigma_1^2) \\
 &\quad \vdots \\
 y_{iq} &= \frac{a_{0q}}{2} + \omega_q x_{iq} + \sum_{k=1}^k (a_{kq} \cos kx_{iq} + b_{kq} \sin kx_{iq}) + \varepsilon_{iq}; \varepsilon_{iq} \sim N(0, \sigma_q^2)
 \end{aligned} \tag{1}$$

Where  $k$  is the representation of the oscillation parameter,  $K$  is the number of the oscillation parameter. Parameters where their values can be determined is based on the Penalized Weighted Least Square (PWLS) result  $a_{0q}, a_{kq}, b_{kq}$ , and  $\omega_q$ . In general, Equation (1) can be written as follows.

$$y_{ij} = \frac{a_{0j}}{2} + \omega_j x_{ij} + \sum_{k=1}^k (a_{kj} \cos kx_{ij} + b_{kj} \sin kx_{ij}) + \varepsilon_{ij}; \varepsilon_{ij} \sim N(0, \sigma_j^2) \tag{2}$$

From Equation (2), it can be expressed in a matrix equation as follows:

$$\mathbf{y} = \mathbf{X}[k]\boldsymbol{\beta} + \boldsymbol{\varepsilon}; \boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma_j^2 \mathbf{I}) \tag{3}$$

with

$$\mathbf{y} = (y_1', y_2', \dots, y_q)$$

$$\mathbf{y}_j = (y_{1j}, y_{2j}, \dots, y_{nj})'$$

and

$$\mathbf{X}[k] = \mathbf{X}_j[k]\mathbf{I}$$

where

$$\mathbf{X}_j[k] = \begin{bmatrix} 1 & x_1 & \cos x_1 & \cdots & \cos kx_1 & \sin x_1 & \cdots & \sin kx_1 \\ 1 & x_2 & \cos x_2 & \cdots & \cos kx_2 & \sin x_2 & \cdots & \sin kx_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & \cos x_n & \cdots & \cos kx_n & \sin x_n & \cdots & \sin kx_n \end{bmatrix}$$

Then, given the parameter vector  $\boldsymbol{\beta}$

$$\boldsymbol{\beta} = (\boldsymbol{\beta}'_1, \boldsymbol{\beta}'_2, \dots, \boldsymbol{\beta}'_j)'$$

with

$$\boldsymbol{\beta}'_j = \left( \frac{a_{0j}}{2}, \omega_j, a_{1j}, \dots, a_{kj}, b_{1j}, \dots, b_{kj} \right)$$

and

$$\boldsymbol{\varepsilon} = (\boldsymbol{\varepsilon}'_1, \boldsymbol{\varepsilon}'_2, \dots, \boldsymbol{\varepsilon}'_q)'$$

with

$$\boldsymbol{\varepsilon}_j = (\varepsilon_{1j}, \varepsilon_{2j}, \dots, \varepsilon_{nj})'$$

The estimator for the regression curve of the multiresponse nonparametric model using the Fourier series approach with sine basis is

$$\begin{aligned} \hat{y}_{i1} &= \frac{\hat{a}_{i1}}{2} + \hat{\omega}_1 x_{i1} + \hat{\beta}_1 x_{i1} + \sum_{p=1}^K (\hat{a}_{k1} \sin kx_{i1}) \\ &\quad \vdots \\ \hat{y}_{iq} &= \frac{\hat{a}_{iq}}{2} + \hat{\omega}_q x_{iq} + \hat{\beta}_q x_{iq} + \sum_{p=1}^K (\hat{a}_{kq} \sin kx_{iq}) \end{aligned} \quad (4)$$

The estimator for the regression curve of the multiresponse nonparametric model using the Fourier series approach with a cosine basis is

$$\begin{aligned} \hat{y}_{i1} &= \frac{\hat{a}_{i1}}{2} + \hat{\omega}_1 x_{i1} + \hat{\beta}_1 x_{i1} + \sum_{p=1}^K (\hat{a}_{k1} \cos kx_{i1}) \\ &\quad \vdots \\ \hat{y}_{iq} &= \frac{\hat{a}_{iq}}{2} + \hat{\omega}_q x_{iq} + \hat{\beta}_q x_{iq} + \sum_{p=1}^K (\hat{a}_{kq} \cos kx_{iq}) \end{aligned} \quad (5)$$

The estimator for the regression curve of the multiresponse nonparametric model using the Fourier series approach with sine cosine basis is

$$\begin{aligned} \hat{y}_{i1} &= \frac{\hat{a}_{i1}}{2} + \hat{\omega}_1 x_{i1} + \hat{\beta}_1 x_{i1} + \sum_{p=1}^K (\hat{a}_{k1} \cos kx_{i1} + \hat{b}_{k1} \sin kx_{i1}) \\ &\quad \vdots \\ \hat{y}_{iq} &= \frac{\hat{a}_{iq}}{2} + \hat{\omega}_q x_{iq} + \hat{\beta}_q x_{iq} + \sum_{p=1}^K (\hat{a}_{kq} \cos kx_{iq} + \hat{b}_{kq} \sin kx_{iq}) \end{aligned} \quad (6)$$

The estimator for the parameters of the simultaneous Fourier series model based on Weighted Least Squared (WLS) optimisation is  $\hat{\boldsymbol{\beta}} = (\mathbf{X}[\mathbf{k}]' \mathbf{W} \mathbf{X}[\mathbf{k}])^{-1} \mathbf{X}[\mathbf{k}]' \mathbf{W} \mathbf{y}$ . The  $\mathbf{W}$  matrix plays a crucial role in optimizing Weighted Least Squares (WLS) to derive the parameter estimator,  $\hat{\boldsymbol{\beta}}$ , in nonparametric multi-response multi-predictor regression using the Fourier series approach. WLS optimization aims to minimize the discrepancy between the observed data and the model's predictions. It is achieved by incorporating weights derived from the inverse of the covariance-variance matrix of the errors, denoted as  $\mathbf{W} = \mathbf{V}^{-1}$ .

$$W = \begin{bmatrix} w_1 & 0 & \cdots & 0 \\ 0 & w_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & w_n \end{bmatrix}$$

while

$$V = \begin{bmatrix} \Sigma_1 & \Sigma_{12} & \cdots & \Sigma_{1j} \\ \Sigma_{21} & \Sigma_2 & \cdots & \Sigma_{2j} \\ \vdots & \vdots & \ddots & \vdots \\ \Sigma_{j1} & \Sigma_{j1} & \cdots & \Sigma_j \end{bmatrix} \quad (7)$$

With  $\Sigma_1$  is a matrix containing:

$$\Sigma_j = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_j^2 \end{bmatrix}$$

and

$$\Sigma_{ij} = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}; i \neq j$$

$\sigma_j^2$  is the  $j$ -th response error variance and  $\sigma_{ij}$  is the  $j$ -th response error covariance with other responses.

Because of the independent assumption between  $i$  and  $j$ .

## 2.5 Simultaneous Kernel Estimator

This research uses the Nadaraya-Watson kernel estimator, often called local polynomial regression of order 0. In general, local polynomial regression expands the Taylor series around  $\mathbf{x}$ . The Taylor series is the main tool for deriving numerical methods and is helpful in approximating (functions into polynomial forms) [21]. The following is the general equation of local polynomial regression for the multi-response case:

$$y(\mathbf{x}_{it}) = f(\mathbf{x}) + (\mathbf{x}_{it} - \mathbf{x})f'(\mathbf{x}) + \frac{(\mathbf{x}_{it} - \mathbf{x})^2}{2!}f''(\mathbf{x}) + \cdots + \frac{(\mathbf{x}_{it} - \mathbf{x})^p}{p!}f^{(p)}(\mathbf{x}) \quad (8)$$

with,

$y(\mathbf{x}_{it})$  is the observed response value at the point  $\mathbf{x}_{it}$ .

$f(\mathbf{x})$  is the value of the function at point

$(\mathbf{x}_{it} - \mathbf{x})f'(\mathbf{x})$  is the first-order term, representing the linear approximation (slope) at  $\mathbf{x}$ .

$\frac{(\mathbf{x}_{it} - \mathbf{x})^2}{2!}f''(\mathbf{x})$  is the second-order term, representing the curvature at  $\mathbf{x}$ .

$\frac{(\mathbf{x}_{it} - \mathbf{x})^p}{p!}f^{(p)}(\mathbf{x})$  is the higher-order terms account for more complex shapes of the function near  $\mathbf{x}$ .

$i$  is the response index =  $0, 1, 2, \dots, m$

$t$  is the time index =  $0, 1, 2, \dots, n$ ,

Let  $\beta_k = \frac{f^{(k)}(\mathbf{x})}{k!}$ ,  $k$  is a local polynomial order, then Equation (8) can also be written as:

$$y(\mathbf{x}_{it}) = \beta_0(\mathbf{x}) + (\mathbf{x}_{it} - \mathbf{x})\beta_1(\mathbf{x}) + (\mathbf{x}_{it} - \mathbf{x})^2\beta_2(\mathbf{x}) + \cdots + (\mathbf{x}_{it} - \mathbf{x})^p\beta_p(\mathbf{x}) + \varepsilon \quad (9)$$

Equation (9) can be written in the following matrix form:

$$\mathbf{y} = \mathbf{X}\beta + \varepsilon \quad (10)$$

with elements  $\mathbf{y}$  and  $\boldsymbol{\varepsilon}$  re the same as Fourier series modeling where  $\mathbf{X}\boldsymbol{\beta}$  contains elements of the Gaussian kernel. In this research, the WLS method was used with the result  $\hat{\mathbf{y}} = \mathbf{H}\mathbf{y}$ , where:

$$\mathbf{H} = (\mathbf{X}'\mathbf{W}(\mathbf{X}_{it})\mathbf{X})^{-1}\mathbf{X}'\mathbf{W}(\mathbf{X}_{it}) \quad (11)$$

So, we obtain a Gaussian kernel estimator simultaneously in Equation (12) as follows:

$$\begin{aligned} y(x_{i1}) = m_1(x_{i1}) &= \frac{\sum_{i=1}^n K_h(x_1 - X_{i1})w_1y_{i1}}{\sum_{i=1}^n K_h(x_1 - X_{i1})} \\ &\vdots \\ y(x_{iq}) = m_q(x_{iq}) &= \frac{\sum_{i=1}^n K_h(x_q - X_{iq})w_qy_{iq}}{\sum_{i=1}^n K_h(x_q - X_{iq})} \end{aligned} \quad (12)$$

Where  $\mathbf{X}_{iq}$  is the input value,  $\mathbf{h}$  is the bandwidth,  $\mathbf{w}_q$  is the weighted variance error, and  $\mathbf{K}_h$  shows the univariate Gaussian kernel function, presented in Equation (13).

$$K_h(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right); -\infty < x < \infty \quad (13)$$

$\mathbf{K}_h(\mathbf{x})$  be the  $n\mathbf{r} \times n\mathbf{r}$  diagonal matrix of weights.

## 2.6 Selection of the Best Model

The bandwidth selection method commonly uses Generalized Cross Validation (GCV) criteria. The general formula for GCV is as follows:

$$GCV(\mathbf{h} \vee \mathbf{k}) = \frac{MSE}{\left(\frac{1}{n} \text{trace}(\mathbf{I} - \mathbf{H}(\mathbf{h} \vee \mathbf{k}))\right)^2} \quad (14)$$

where  $\mathbf{h}$  is the bandwidth in the kernel approach and  $\mathbf{k}$  is the oscillation parameter in the Fourier series approach.

$MSE = \frac{1}{n} \mathbf{y}'(\mathbf{I} - \mathbf{H}(\mathbf{h} \vee \mathbf{k}))' \mathbf{W}(\mathbf{I} - \mathbf{H}(\mathbf{h} \vee \mathbf{k})) \mathbf{y}$  and  $\mathbf{I}$  are identity matrices.

Calculations need to be carried out to determine the correspondence between actual data and predicted data in measuring the accuracy of predictions. Several calculations are generally used to measure the accuracy of prediction results, namely MAPE [22]. MAPE defines the percentage error of prediction results from actual data over time by providing information regarding the percentage error that is high or low. MAPE calculation by dividing the absolute error of each period according to the number of responses. The resulting MAPE value has the following interpretation [23]:

**Table 2.** MAPE value interpretation interval

MAPE (%)	Interpretation
$MAPE \leq 10\%$	High accurate predictions
$10\% < MAPE \leq 20\%$	Good prediction
$20\% < MAPE \leq 50\%$	Predictions are still within reason
$MAPE > 50\%$	Inaccurate predictions

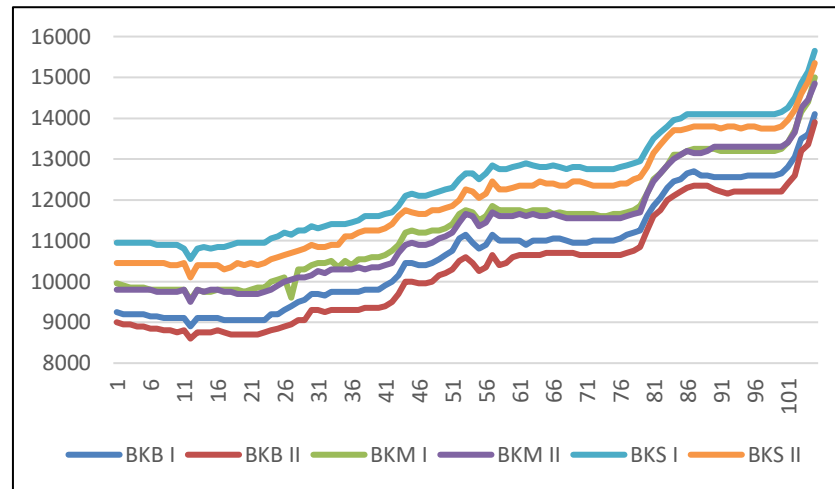
## 3. RESULTS AND DISCUSSION

### 3.1 Data Exploration

The data used in this research is weekly secondary data obtained from PIHPS via the website [bi.go.id/hargapangan](http://bi.go.id/hargapangan) from February 2022 to February 2024. Empirical studies prove that the best prediction results are obtained when the data set is divided into 80-90% data training and 20-30% test data [24]. The



data has been divided into 80% training data for or as many as 84 data from the period 24 February 2022 to 26 September 2023 and testing data, which has been divided into 20% or as many as 21 data from the period 3 October 2023 to 20 February 2024, presented in a time series plot Rice prices of the six different qualities are shown in **Figure 2**.



**Figure 2. Rice price plot data**

Generally, rice prices increased for all qualities from February 2022 to February 2024. In 105 weekly data on rice prices, an average of IDR 11,352.22 was obtained with a standard deviation of IDR 1,497.37. The characteristics of rice price data for all qualities are presented in **Table 3** as follows.

**Table 3. Characteristics of rice prices for the period 24 February 2022 – 20 February 2024**

Descriptive	Rice Quality					
	Min	Max	Average	St.Dev	Periode Min	Periode Max
<b>BKB I</b>	8900	14100	10706.19048	1348.373223	10/5/2022	20/2/2024
<b>BKB II</b>	13900	8600	10328.1	1356.092	10/5/2022	20/2/2024
<b>BKM I</b>	15000	9550	11392.38	1325.014	22/2/2022	20/2/2024
<b>BKM II</b>	14850	9500	11304.29	1364.349	10/5/2022	20/2/2024
<b>BKS I</b>	15650	10550	12396.19	1254.754	22/2/2022	20/2/2024
<b>BKS II</b>	15350	10100	11986.19	1312.12	10/5/2022	20/2/2024

Multiresponse nonparametric regression aims to identify and describe the relationship between several response variables and predictor variables, where the relationship pattern cannot be explained or known previously [25]. In nonparametric regression, the response variables must be correlated with each other.

**Table 4. Bartlett's Test of Sphericity**

<b>KMO Measure of Sampling Adequacy</b>		.919
Bartlett's Test of Sphericity	Approx. Chi-Square	2880.061
	df	15
	Sig.	.000

**Table 4** states that the Bartlett Test produces a P-value ( $0.00 < 0.05$ ). So, there is a relationship between variables. Therefore, modeling can be carried out simultaneously for all rice price values based on quality by comparing the Gaussian kernel estimator and the Fourier series estimator.

### 3.2 Fourier Series Estimation Results

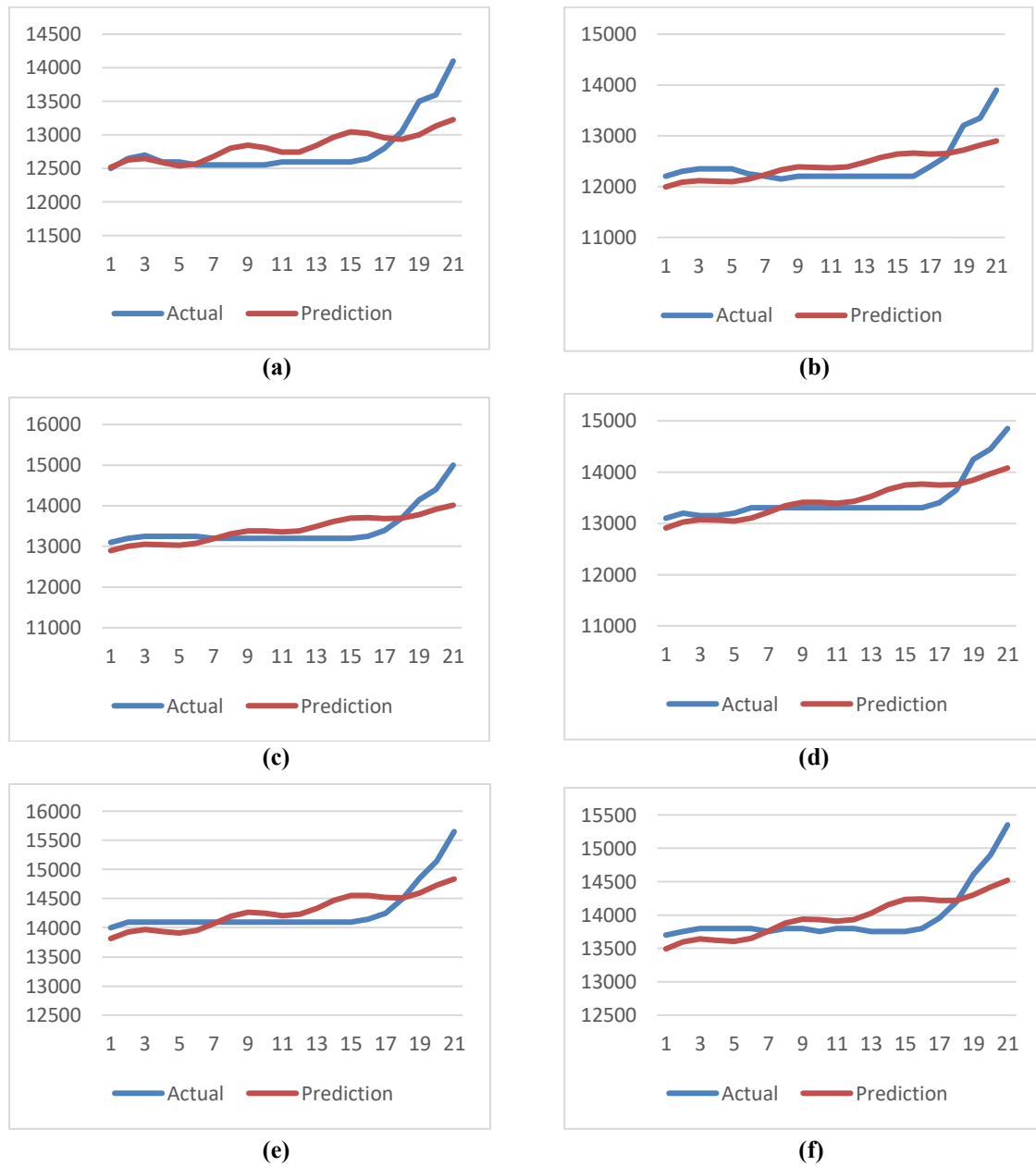
In nonparametric regression utilizing Fourier series estimation, an oscillation parameter ( $k$ ) exists. The selection of the optimal value for ( $k$ ) is crucial in constructing the model. It is determined by minimizing the GCV value. Next, the most optimal function will be selected by comparing the minimum GCV of the three functions. A compares the Fourier series estimation results can be seen in **Table 5**.



**Table 5. Comparison of Fourier series estimation results**

Function	k	GCV	MSE
Sine	1	0.002337377	1.095395
Cosine	1	0.002372061	1.111649
Cosine – Sine	1	0.00239241	1.093673

It was found that the best Fourier series model is the model with the sine function because it has the minimum GCV value. The sine function has an optimum oscillation parameter value (k) of 1 with a GCV value of 0.002337377, MSE of 1.095395, and  $R^2$  of 93.86499. Thus, we get the results of predicting the price of rice with six different qualities on test data or 21 periods into the future. A comparison plot of predicted values with actual values of rice prices is presented in **Figure 3**.

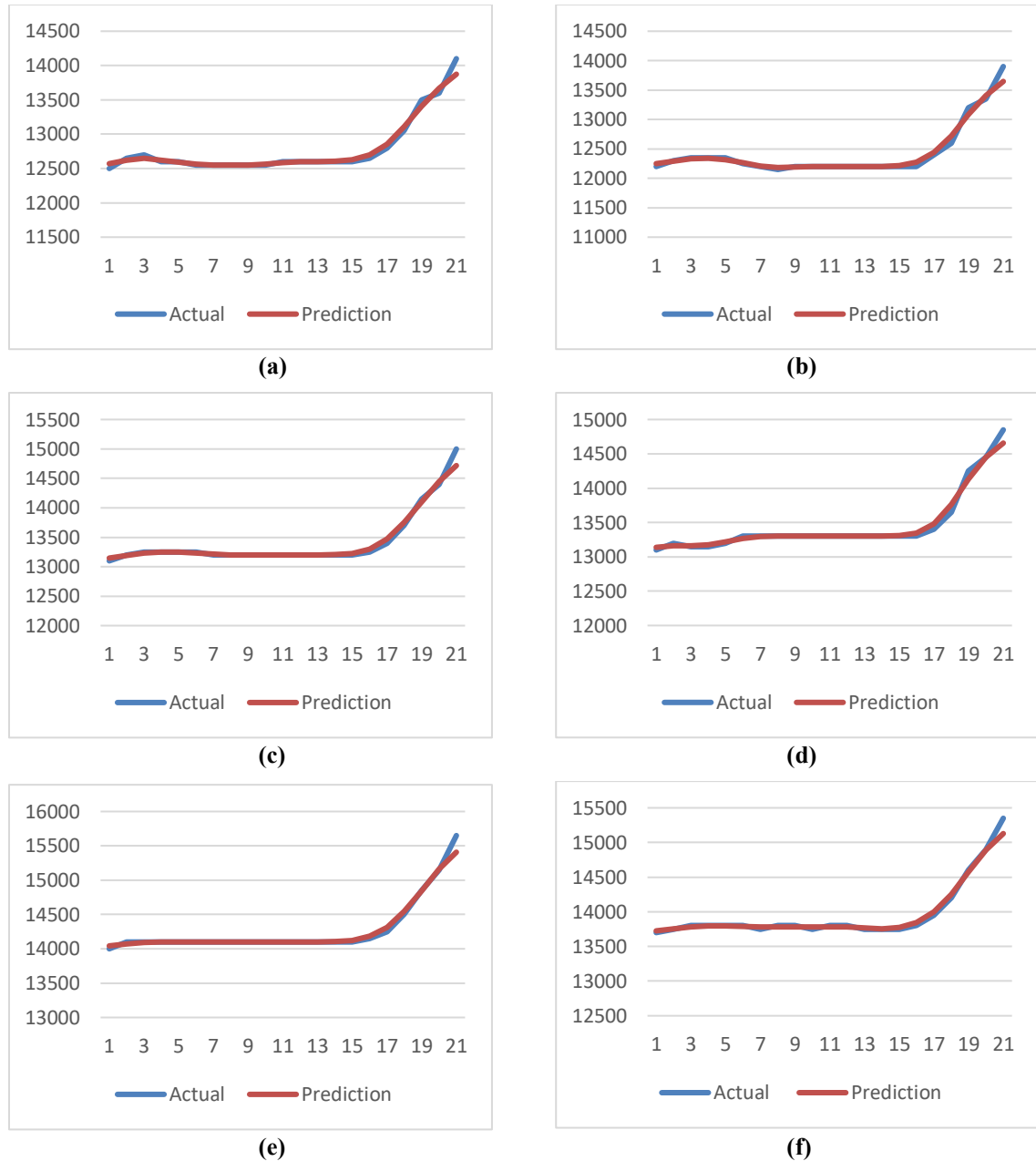


**Figure 3. Plot of actual data and predicted data for variables (a) BKB I, (b) BKB II, (c) BKM I, (d) BKM II, (e) BKS I, and (f) BKS II using the Fourier series estimator**

Based on the plot data in **Figure 3**, it is shown that the predicted data using the Fourier series estimator is similar to the actual data in several periods for each type of rice. The Mean Average Percentage Error (MAPE) value obtained from prediction results using testing data based on the Fourier series estimator of the sine function with an oscillation parameter (k) equal to 1 is 1.872354%, which means that the forecasting results for rice price data with six different qualities are high accurate predictions.

### 3.3 Kernel Estimation Result

The best kernel estimator was selected using a Gaussian kernel, and an optimal bandwidth of 1 was obtained. From the results of the optimal bandwidth, the GCV value is 5703.129167, and the R-squared is 99.75448333. A comparison plot of the predicted value with the actual value of rice prices is presented in **Figure 4**.



**Figure 4.** Plot of actual data and predicted data for variables (a) BKB I, (b) BKB II, (c) BKM I, (d) BKM II, (e) BKS I, and (f) BKS II using Gaussian Function Kernel estimator

Based on the plot data in **Figure 4**, it is shown that the predicted data using the Gaussian kernel estimator function has a minimal difference from the actual data for the entire period for each type of rice. The Mean Average Percentage Error (MAPE) value obtained from prediction results using testing data based on a Gaussian kernel function estimator with a bandwidth parameter ( $h$ ) equal to 1 is 0.228259%, which means that the emission results for rice price data with six different qualities are high accurate prediction.

### 3.4 Comparison of prediction results

In the previous subchapter, predictions were made based on the Fourier series estimator and Gaussian kernel to select the best method based on the smallest MAPE value. A comparison of the MAPE value results for the two methods is presented in **Table 6** as follows.

**Table 6. Performance comparison of Fourier series methods and Kernel functions**

Method	MAPE
Gaussian Function Kernel	0.228259%
Fourier Sine Function	1.872354%

Based on the prediction results, both have resulted in the very accurate prediction category with MAPE values below 10 percent. When the two methods are compared, the best prediction result is using the Gaussian kernel method because it has the smallest MAPE value.

## 4. CONCLUSIONS

Based on the analysis results that have been conducted, several conclusions can be drawn as follows:

1. Prediction of rice prices in East Java can be done simultaneously using the Fourier series estimator approach and the Gaussian kernel function.
2. The prediction performance with both models is highly accurate prediction because it produces a MAPE of less than 10%. Both approaches also include parsimony because they produce a minimum GCV on the oscillation parameters for the Fourier series, and the bandwidth for the Gaussian kernel is assessed as 1. Then, predictions using the Fourier series method with the sine function produce a greater MAPE value of 1.872354% compared to the Gaussian function kernel with a MAPE value of 0.228259%.
3. The kernel method with the Gaussian function is better for predicting rice prices for each quality because it has a smaller MAPE value and is categorized as a highly accurate prediction.

From the results of this prediction, the government can implement policies to control rice prices to overcome the increase in rice prices that occurred in Indonesia. Especially in controlling the quality and quantity of rice production, import and export policies, rice subsidies, and rice trade regulations, which influence the production, distribution, and price of rice in East Java, the largest rice-producing province in Indonesia.

## ACKNOWLEDGMENT

The researcher would like to thank the Ministry of Education and Culture of Higher Education for the MBKM policy, which allows students to participate in the Research Internship course in the Statistics Study Program, Faculty of Science and Technology, Airlangga University. We also appreciate Bank Indonesia (PIHPS) for providing research data and all parties who have supported this research and publication process.

## REFERENCES

- [1] E. Nupuku, "Analisis Forecasting Produksi dan Konsumsi Beras di Propinsi Sumatera Utara," *Jurnal Darma Agung*, vol. 29, no. 3, pp. 370–377, Dec. 2021.
- [2] M. Ohlyver and H. Pudjihastuti, "Arima Model for Forecasting the Price of Medium Quality Rice to Anticipate Price Fluctuations," in *Procedia Computer Science*, Elsevier B.V., 2018, pp. 707–711. doi: 10.1016/j.procs.2018.08.215.
- [3] K. Sukiyono and R. Rosdiana, "Pendugaan Model Peramalan Harga Beras Pada Tingkat Grosir," *Jurnal AGRISEP*, vol. 17, no. 1, pp. 23–30, Mar. 2018, doi: 10.31186/jagrisep.17.1.23-30.
- [4] D. Setiawati, E. Daris, and D. M. Najamuddin, "Analisis Faktor-Faktor yang Mempengaruhi Pembentukan Harga Beras di Indonesia," *Jurnal Agribisnis*, vol. 12, no. 1, 2018.
- [5] A. A. Fardhani, D. I. N. Simanjuntak, and A. Wanto, "Prediksi Harga Eceran Beras Di Pasar Tradisional Di 33 Kota Di Indonesia Menggunakan Algoritma Backpropagation," *Jurnal Infomedia*, vol. 3, no. 1, Jun. 2018.

- [6] Parmentaria, "Pemerintah Harus Bongkar Dugaan Kartel Penyebab Harga Beras Naik," <https://www.dpr.go.id/berita/detail/id/48585/t/Pemerintah+Harus+Bongkar+Dugaan+Kartel+Penyebab+Harga+Beras+Nai+k>. Accessed: Feb. 26, 2024. [Online]. Available: <https://www.dpr.go.id/berita/detail/id/48585/t/Pemerintah+Harus+Bongkar+Dugaan+Kartel+Penyebab+Harga+Beras+Nai+k>
- [7] PIHPS, "Informasi Harga Pangan Antar Daerah," Bank Indonesia. Accessed: Feb. 26, 2024. [Online]. Available: <https://www.bi.go.id/hargapangan>
- [8] BMKG, "Mengenal Fenomena El Nino yang Mengancam Indonesia. BMKG Stasiun Klimatologi Jawa Timur," BMKG. Accessed: Mar. 02, 2024. [Online]. Available: <https://staklim-jatim.bmkg.go.id/>
- [9] BPS, "Luas Panen, Produktivitas, dan Produksi Padi menurut Kabupaten/Kota di Provinsi Jawa Timur, 2021 dan 2022," BPS Jawa Timur. Accessed: Mar. 03, 2024. [Online]. Available: <https://jatim.bps.go.id/statictable/2023/03/16/2521/luas-panen-produktivitas-dan-produksi-padi-menurut-kabupaten-kota-di-provinsi-jawa-timur-2021-dan-2022.html>
- [10] M. Sihombing, "Inilah Jalur Distribusi Beras Hingga Ke Konsumen," *Bisnis.com*. Accessed: Feb. 22, 2024. [Online]. Available: <https://ekonomi.bisnis.com/read/20180117/99/727664/inilah-jalur-distribusi-beras-hingga-kekonsumen>
- [11] M. Farid and N. Ishari, "Pengaruh Kenaikan Harga Beras terhadap Kualitas Produk Warung Makan Milik Orang Muslim di Desa Wonorejo Kecamatan Kedungjajang Kabupaten Lumajang," 2021.
- [12] M. F. F. Mardianto, Sediono, I. Syahzaqi, S. A. D. Safitri, and N. Afifah, "Prediction of Indonesia Strategic Commodity Prices during the COVID-19 Pandemic based on a Simultaneous Comparison of Kernel and Fourier Series Estimator," *Journal of Southwest Jiaotong University*, vol. 55, no. 6, 2020, doi: 10.35741/issn.0258-2724.55.6.43.
- [13] M. Fariz, F. Mardianto, H. Kartiko, and H. Utami, "Prediction the Number of Students in Indonesia who Study in Tutoring Agency and Their Motivations based on Fourier Series Estimator and Structural Equation Modelling."
- [14] M. D. Pasarella, S. Sifriyani, and F. D. T. Amijaya, "Nonparametric Regression Model Estimation with The Fourier Series Approach and Its Application to The Accumulative Covid-19 Data in Indonesia," *BAREKENG: Jurnal Ilmu Matematika dan Terapan*, vol. 16, no. 4, pp. 1167–1174, Dec. 2022, doi: 10.30598/barekengvol16iss4pp1167-1174.
- [15] D. Made, R. Widyadarma, D. Saepuddin, and F. Nhita, "Agricultural Commodities Price Prediction Using Hybrid Elman Neural Network with Genetic Algorithm."
- [16] C. F. F. Purwoko, S. Sediono, T. Saifudin, and M. F. F. Mardianto, "Prediksi Harga Ekspor Non Migas di Indonesia Berdasarkan Metode Estimator Deret Fourier dan Support Vector Regression," *Inferensi*, vol. 6, no. 1, p. 45, Mar. 2023, doi: 10.12962/j27213862.v6i1.15558.
- [17] Effilyadi, Pasaribu. J.P.K, Suratno. E, Kadar. M, Gunardi, Naibaho. R, Hati. S.K, Aryati. V, "Penerapan Uji Multikolinearitas dalam Penelitian Manajemen Sumber Daya Manusia," *Jurnal Ilmiah Manajemen dan Kewirausahaan*, vol. 1, no. 2, 2022, [Online]. Available: <https://ejournal.unama.ac.id/index.php/jumanage>
- [18] P. Bloomfield, *Fourier Analysis of Time Series: An Introduction*, Second., vol. 74, no. 368. North Caroline: John Wiley & Sons, Inc, 1979.
- [19] F. Faisol, P. Ukhrowi, M. F. F. Mardianto, I. Yudistira, and K. Kuzairi, "Comparison of Salinity and Seawater Temperature Predictions Using VAR and Biresponse Fourier Series Estimator," *BAREKENG: Jurnal Ilmu Matematika dan Terapan*, vol. 16, no. 4, pp. 1465–1476, Dec. 2022, doi: 10.30598/barekengvol16iss4pp1465-1476.
- [20] M. F. F. Mardianto *et al.*, "Comparison of Simultaneously Nonparametric Regression Based on Spline and Fourier Series Estimator Related Social Aid Distribution in Indonesia," *Communications in Mathematical Biology and Neuroscience*, vol. 2022, 2022, doi: 10.28919/cmbn/7545.
- [21] M. Rizaldi, "Estimasi Kurva Regresi Nonparametrik Fungsi Kernel Pada Data Longitudinal (Studi Kasus: Laju Pertumbuhan Ekonomi Provinsi Nusa Tenggara Barat 2016-2020)," Universitas Mataram, Mataram, 2023.
- [22] X. Guo, Y. Gao, Y. Li, D. Zheng, and D. Shan, "Short-term household load forecasting based on Long- and Short-term Time-series network," *Energy Reports*, vol. 7, pp. 58–64, Apr. 2021, doi: 10.1016/j.egy.2021.02.023.
- [23] J. J. Montaña Moreno, A. Palmer Pol, A. Sesé Abad, and B. Cajal Blasco, "El índice R-MAPE como medida resistente del ajuste en la previsión," *Psicothema*, vol. 25, no. 4, pp. 500–506, 2013, doi: 10.7334/psicothema2013.23.
- [24] E. Karyadiputra and E. Hijriana, "Analisis Penerapan Algoritma Naive Bayes untuk Klasifikasi Prioritas Pengembangan Jalan di Provinsi Kalimantan Selatan," *Jurnal Ilmiah Teknologi*, vol. 9, no. 2, pp. 500–506, Apr. 2018.
- [25] D. Paramita, "Estimator Campuran Spline Truncated dan Kernel dalam Regresi Nonparametrik Birespon untuk Pemodelan Faktor-Faktor yang Mempengaruhi Keberhasilan Pembangunan Di Provinsi Jawa Timur," Institut Teknologi Sepuluh Nopember, Surabaya, 2020.