International Journal of Innovation, Creativity and Change. www.ijicc.net
Volume 5, Issue 3, 2019 Special Edition: *Science, Applied Science, Teaching and Education*

# Prediction of National Strategic Commodities Production based on Multi - Response Nonparametric Regression with Fourier Series Estimator

M. Fariz Fadillah Mardianto*, Siti Maghfirotul Ulyah, Eko Tjahjono

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

*Email: m.fariz.fadillah.m@fst.unair.ac.id

## *Abstract*

The Ministry of Agriculture in Indonesia stated that there are 11 strategic commodities that have the largest contribution to food security and the formation of inflation rates in Indonesia. There are rice, corn, shallot, garlic, red chili, cayenne pepper, beef, chicken, broiler eggs, sugar, and cooking oil. The supply of strategic commodities that are suitable to the needs of the Indonesian people can maintain the stability of national food security. Indonesia's Government depends on provinces that become the main producer of most commodities, like East Java. However, a prediction can be made to determine the availability of these commodities in the coming period, based on data from the previous period. Because the data has an oscillation pattern, Fourier series estimators in multi-response case is used to forecast. Fourier series have the flexibility to approach the data pattern smoothly. The data from the East Java Province Government in Indonesia is taken for 11 commodities. The result is an optimal model based on the parsimony model with the small Mean Square Error (MSE), a Generalised Cross Validation (GCV) and the big determination coefficient value. The model that has been selected has a small goodness of fit criteria to forecast. So, Fourier series estimators with a multi-response case is suitable to predict national strategic commodity production in East Java, that give high contributions to Indonesia's achievement for food security.

***Keywords:*** *Fourier series estimator, multi response, nonparametric regression, food security, national strategic commodities.*

## Introduction

The Law of the Republic of Indonesia number 18 of 2002 states that food security is a condition of fulfilling food for the country up to the individuals, as reflected in the availability of sufficient food, both in quantity and quality, safe, diverse, nutritious, equitable and affordable and not contrary to religion, beliefs and community culture, to be able to live a healthy, active and productive life in a sustainable manner. Food security for a country is very important, especially for countries that have large populations such as Indonesia. Indonesia's population is estimated to reach 200 million by 2020 and projected to reach 270 million by 2025. Indonesia's historical development experience shows that the problem of food security is closely related to economic stability, especially inflation, aggregate economic production costs, and national political stability. Therefore, the government's decision related to food security is a national strategic policy (Hanafie, 2010).

Based on the Ministry of Agriculture of the Republic of Indonesia, there are 11 strategic commodities that have the largest contribution to food security and the formation of inflation figures in Indonesia. These commodities are rice, corn, shallot, garlic, red chili, cayenne pepper, beef, chicken, broiler eggs, sugar and cooking oil (The Ministry of Agriculture, 2015). The supply of strategic commodities that are suitable to the needs of the Indonesian people can maintain the stability of national food security. But in reality, at a certain time the supply of strategic commodities is not proportional to the needs of people in Indonesia. Some of the factors that influence these problems are the high needs of the community, for example when approaching the month of Ramadan, the feast means there are limited supplies due to failures and delays in the process of harvesting or production, distribution processes, and factors of weather conditions. However, statistically, a prediction can be made to determine the availability of these commodities in the coming period, based on data from the previous period. The prediction results are useful in evaluating government policies related to food security in Indonesia.

Several studies related to the prediction of food commodities simultaneously so far have never been done. Ahmad (2011) predicts the production of chicken eggs with the Neural Network (NN) approach. Koide, Robertson, Ines, Qian, Witt & Lucero (2011) predicts rice production in the Philippines based on climate forecasting using the seasonal time series method. Septiawan, Komarudin & Sulistya (2012) predict the price of chili based on climatic conditions with a linear regression approach. Penedo (2014) compared several forecasting methods to predict general sugar production in Brazil. These predictions only involve one commodity.

Prediction of strategic commodity production is carried out in East Java Province. East Java Province is the second most populated province in Indonesia. East Java Province has the largest area of agricultural land in Indonesia from year to year. In 2015, East Java Province was the largest producer of rice, corn, soybean, peanuts, beef, eggs and native chicken meat in Indonesia (Central Bureau of Statistics, 2017). These data conclude that East Java Province plays an important role in national food security, because this province has the largest population and can produce some of the largest superior commodities in Indonesia.

In predicting the production of strategic commodities, a regression approach is used. Regression approach is an approach used to model the causality relationship between response variables based on predictor variables. In regression analysis, the shape of the regression curve can be approximated by mathematical functions that fit the data pattern. If data patterns tend to form known patterns of mathematical functions such as linear, quadratic, or cubic functions, the regression approach used is the parametric regression approach. Besides that, in reality, not all data follows certain patterns. If the pattern of the relationship between the predictor and response variables is unknown, then the nonparametric regression approach is appropriate for modeling the relationship between these variables. In addition, there is a semiparametric regression approach that is formed from the sum of the parametric regression and nonparametric regression functions (Mardianto, Tjahjono & Rifada, 2019a). The approach used in this study is nonparametric regression.

Nonparametric regression approach develops because all of the data patterns are not suitable when estimated with the parametric regression approach (Takezawa, 2005). The advantage of nonparametric regression is it's flexibility properties. The flexibility means that the pattern of data that is presented in scatter plots can determine the shape of the regression curve based on estimators in the nonparametric regression study. This study uses Fourier series estimators in nonparametric regression. Nonparametric regression approach with Fourier series estimator has rapid research development. Fourier series estimators have flexibility in approaching repetitive or oscillating data patterns that are consistent with trigonometric function patterns (Mardianto, Tjahjono & Rifada, 2019b). Fourier series estimator is suitable to model seasonal, and trend – seasonal data pattern (Mardianto, Kartiko & Utami, 2019a). Some studies that use Fourier series estimators in nonparametric regression were initiated by Bilodeau (1992). Some other recent studies include Biedermann, Dette & Hoffmann (2009) examined the optimal design for obtaining refinement parameters from a nonparametric regression model using the Fourier series with cosine and sine component. Tjahjono, Mardianto & Chamidah (2018) determines Fourier series estimators in bi-response cases with application in predicting morning and night-time electricity consumption. Mardianto, Cahyono, Syarifah & Andriani (2019) applied Fourier series estimators for time series data in tourism data separately. Mardianto, Tjahjono, Rifada, Herawanto, Putra &

Utama (2019) used Fourier series estimator to predict rice production. None of these studies have been applied for responses in more than two or multi-response cases.

The nonparametric regression approach used in this study is a multi-response nonparametric regression approach based on the Fourier series estimator in time series data. In the context of time series analysis Fourier series functions are appropriate when used as periodic or seasonal data representations (Bloomfield, 2000). Bloomfield (2000) introduces the use of the Fourier series for alternative approaches to time series analysis. In the context of nonparametric multi-response Fourier series regression, it can be used to predict many response variables that are correlated with each other in a particular observation time. Fourier series estimators used in this study are Fourier series estimators based on sines and cosines based on Biedermann, Dette & Hoffmann (2009) and Dette, Melas & Shpilev (2016). This function has been used by Bloomfield (2000) in studying the use of the Fourier series in time series analysis. Theoretically, Fourier series estimators in nonparametric regression with longitudinal data based on Mardianto, Kartiko, & Utami (2019b) can be generalised for multi–response casse. This paper will give significant theoretical results about the development of the Fourier series estimator in nonparametric regression for multi–response cases, and in application to model more than two commodities simultaneously to make recommendations for food security.

## Materials and Method

### *Data and Variables*

The data of this study are secondary data related to the production of strategic commodities in East Java. Those commodities are rice, corn, shallot, garlic, red chili, cayenne pepper, beef, chicken, broiler eggs, sugar, and cooking oil. The data was taken annually from 2000 to 2018. The data from 2000 to 2013 are used as in sample data for estimation. The data from 2014 to 2018 are used as out sample data for prediction. The data was taken at the East Java Agriculture Office for data about rice, corn, shallot, garlic, red chili, cayenne pepper, sugar cane, and soybean production. The data about production of beef, chicken, broiler eggs were taken at the East Java Animal Husbandry Office.

This study consisted of the observation time ($t$) in months as a predictor. The observation unit in this study is annual data for 19 years. For more details, the data structure is presented in **Table 1**. There are 11 response variables in this study ($y_j$), $j = 1, \ldots, 11$. The response variables are presented as follows:

1. Annual sugar production in East Java, Indonesia (tonnes).
2. Annual shallot production in East Java, Indonesia (tonnes).
3. Annual garlic production in East Java, Indonesia (tonnes).
4. Annual red chili production in East Java, Indonesia (tonnes)
5. Annual cayenne pepper production in East Java, Indonesia (tonnes)
6. Annual rice production in East Java, Indonesia (tonnes)
7. Annual chicken production in East Java, Indonesia (tonnes).
8. Annual beef production in East Java, Indonesia (tonnes).
9. Annual broiler eggs production in East Java, Indonesia (tonnes).
10. Annual soybean production in East Java, Indonesia (tonnes).
11. Annual corn production in East Java, Indonesia (tonnes).

**Table 1.** The data structure for multi-response cases in this study

| Number | Year | Response Variable | | | |
|---|---|---|---|---|---|
| | | $y_{i1}$ | $y_{i2}$ | ... | $y_{i11}$ |
| 1 | 2000 ($t_1$) | $y_{11}$ | $y_{12}$ | ... | $y_{111}$ |
| 2 | 2001 ($t_2$) | $y_{21}$ | $y_{22}$ | ... | $y_{211}$ |

| | | | | | |
|---|---|---|---|---|---|
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 19 | 2018 ($t_{19}$) | $y_{19\,1}$ | $y_{19\,2}$ | … | $y_{19\,q}$ |

### *Fourier series Estimator in Nonparametric Regression in Multi-Response Cases*

Regression analysis is a statistical method used to model the relationship between response and predictor variables. The purpose of the regression model for making predictions is based on a function that estimates the data pattern. Based on the data pattern approached by mathematical function, there are three approaches in regression analysis (Budiantara, Ratnasari, Zain, Ratna & Mardianto, 2015). There are parametric, nonparametric, and semi parametric regression. Nonparametric regression has a high flexibility in modelling data patterns which is unknown or unrecognised, so the regression curve looks for the data pattern (Takezawa, 2005). One of estimators used in nonparametric regression is the Fourier series that be proposed by Bilodeau on 1992 at the first time (Bilodeau, 1992). In this case, the Fourier series estimator is applied for multi–response cases with a time series data structure.

Consider pairs of data with the form $(y_{i1}, y_{i2}, \dots, y_{iq}, t_i)$, where $t_i$ denotes time for $i^{th}$ observation. Response variable for $j^{th}$product, $j = 1,2, \dots, q$ in $i^{th}$ observation is denoted by $y_{ij}$. The pairs of data presented in **Table 1** follows nonparametric regression equation in multi–response cases.

$$\begin{cases} y_{i1} = m_1(t_i) + \varepsilon_{i1}; \varepsilon_{i1} \sim N(0, \sigma_1^2) \\ y_{i2} = m_2(t_i) + \varepsilon_{i2}; \varepsilon_{i1} \sim N(0, \sigma_2^2) \\ \vdots \\ y_{ij} = m_j(t_i) + \varepsilon_{ij};; \varepsilon_{ij} \sim N(0, \sigma_j^2) \end{cases} \quad (1)$$

$m_j(t_{ij})$ represents a regression curve. Random error for $i^{th}$ observation in $j^{th}$ response is denoted by $\varepsilon_{ij}$ that independent, identically normal distributed with mean 0, and variance $\sigma_j^2$. In this case, Equation (1) can be approached by Fourier series estimators. The Fourier series equation for paired data with the form $(y_{i1}, y_{i2}, \dots, y_{iq}, t_i)$, is given in equation (2) as follows:

$$m_j(t_i) = \frac{\alpha_{0j}}{2} + \gamma_j t_i + \sum_{k=1}^{K} (\alpha_{kj} \cos k\, t_i + \beta_{kj} \sin k\, t_i) \tag{2}$$

$k$ 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 Penalised Weighted Least Square (PWLS) result (Mardianto, Kartiko & Utami, 2019b), denoted by $\alpha_{0j}$, $\gamma_j, \alpha_{kj}$ and $\beta_{kj}$. The nonparametric regression equation based on the Fourier series estimator for multi–response cases can be formed with the substitution process from equation (2) to equation (1) with results as follows:

$$y_{ij} = \frac{\alpha_{0j}}{2} + \gamma_j t_i + \sum_{k=1}^{K} (\alpha_{kj} \cos k\, t_i + \beta_{kj} \sin k\, t_i) + \varepsilon_{ij}, \varepsilon_{ij} \sim N\left(0, \sigma_j^2\right) \tag{3}$$

Form equation (3) and based on PWLS result with weighted is estimated based on Wu & Zhang (2006), the estimation form for nonparametric regression curve can be obtained as follows:

$$\hat{y}_{ij} = \frac{\hat{\alpha}_{0j}}{2} + \hat{\gamma}_j t_i + \sum_{k=1}^{K} \left(\hat{\alpha}_{kj} \cos kt_i + \hat{\beta}_{kj} \sin kt_i\right) \tag{4}$$

In this case, three kinds of performance indicators which relate with goodness of fit that popular in nonparametric regression study are used to are selected as a good model for prediction. There are Generalised Cross Validation (GCV), Mean Square Error (MSE), and determination coefficients ($R^2$). The first performance indicator that be considered is GCV (Wahba, 1990). For determining an optimal quantity measure for smoothing it can be seen based on the smallest GCV value (Eubank, 1999). The smallest GCV implies the small MSE value. The last performance indicator considered is $R^2$. In this case, we consider to select a model with the bigger value of the determination coefficient based on parsimony model. Furthermore, we use three kinds of Fourier series to determine the best estimator result to predict the production of national strategic commodity. There are cosine, sine and Fourier series that include both of trigonometric components (Mardianto, Tjahjono, & Rifada, 2019a).

## Results and Discussion

### *Description about the Production of National Strategic Commodities in East Java*

In order to maintain the stability of national food security, and achieve the target of Indonesia in 2045 to become a world food storage country, East Java Province has a large role. Demographically, East Java Province based on the 2010 population census has the second largest population in Indonesia, amounting to 37, 565, 800 inhabitants. In other words, 15.75% of Indonesia's population is in East Java Province (Central Beureau of Statistics, 2017). In terms of agriculture, plantation, animal husbandry, fisheries, and trade, East Java Province has several achievements which are presented in **Table 2**. Based on **Table 2**, East Java has the largest agricultural land area in Indonesia. In addition, most of the national strategic commodities are produced in the East Java Province. This shows that East Java Province plays an important role in maintaining the stability of national food security.

**Table 2.** East Java Province achievements in 2015

| Number | Achievement | Achievement Value | Ranking in Indonesia | National Contribution |
|---|---|---|---|---|
| 1. | Agricultural land area | 1,091,752 hectares | 1 | 13.5% |
| 2. | Rice production | 13,154,967 tonnes | 1 | 17.45% |
| 3. | Corn production | 6,131,163 tonnes | 1 | 31.26% |
| 4. | Soybean production | 344,998 tonnes | 1 | 35.82% |
| 5. | Shallot production | 30,783 tonnes | 2 | 25.21% |
| 6. | Chili production | 68,218 tonnes | 1 | 26.68% |

| | | | | |
|---|---|---|---|---|
| 7. | Sugar production | 1,207,300 tonnes | 1 | 48.33% |
| 8. | Beef production | 4,267.3 tonnes | 1 | 27.68% |
| 9. | Chicken production | 194,604.9 tonnes | 2 | 12.7% |
| 10. | Broiler eggs production | 390,055 tonnes | 1 | 28.41% |

**Table 3.** Descriptive Statistics of National commodities produced in East Java

| Commodities | Average | Maximum Result (Year) | Minimum Result (Year) |
|---|---|---|---|
| Rice | 10,655,550 | 13,154,967 (2015) | 8,672,791 (2001) |
| Corn | 4,681,224 | 6,647,219 (2017) | 2,309,899 (2016) |
| Sugar | 1,186,785 | 1,338,182 (2001) | 1,013,600 (2010) |
| Soybean | 325,536 | 385,212 (2000) | 252,027 (2007) |
| Broiler eggs | 263,420 | 465,838 (2018) | 93,044 (2000) |
| Shallot | 242,344 | 344,642 (2001) | 181,489 (2009) |
| Cayenne pepper | 191,375 | 339,022 (2017) | 94,645 (2000) |
| Chicken | 160,221 | 277,654 (2018) | 72,664 (2000) |
| Beef | 91,876 | 112,447 (2011) | 71,650 (2002) |
| Red chili | 84,136 | 111,022 (2014) | 60,747 (2005) |
| Garlic | 1,829 | 9,507 (2002) | 528 (2015) |

**Table 3** summarises the commodities with the highest average production in East Java, with maximum and minimum yields along with the year of achievement. Based on **Table 3**, rice is the main commodity and highest commodity produced in East Java. The resulting value cannot be matched by the achievements of other commodities. Corn is the second largest commodity produced in East Java. The resulting value cannot be matched by the achievements of other commodities except rice. Sugar is the third largest commodity produced in East Java. The resulting value cannot be matched by the achievements of other commodities except rice and corn. Meanwhile, fourth to tenth ranked is occupied by soybeans, broiler eggs, shallot, cayenne pepper, chicken, beef, and red chili. Every year, the highest rank of production for these commodities are

always changing because the ranges are close to each other. East Java's strategic commodity with the lowest production is garlic. Garlic production achievements cannot approximate other commodity production and always become the lowest production value than all of strategic commodities. The efforts from the related stakeholders in East Java for the last three years to increase the production of several national strategic commodities such as rice, corn, broiler eggs and chicken production results the highest production achievement for those commodities.

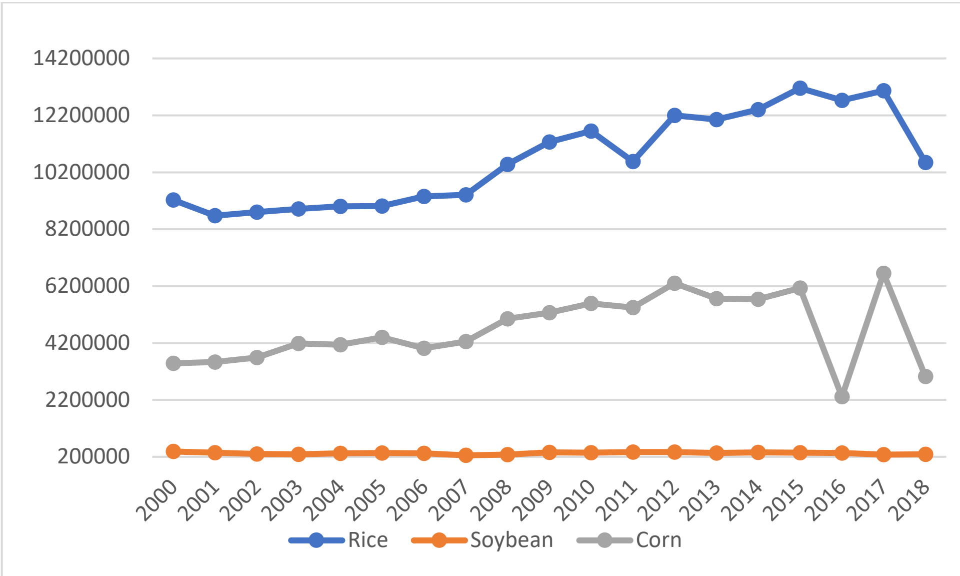


**Figure 1.** Development of strategic commodity production for agriculture sector

In modeling using time series data, what is considered is the data pattern. **Figure 1** shows the development of strategic commodity production in the East Java agricultural sector, consisting of rice, soybeans, and corn. Based on **Figure 1** the trend of increasing national strategic commodity production in East Java in the agriculture sector can be seen. A clear upward trend was occurring in rice commodities. In the case of corn, the trend had not been maintained since 2016, an upward trend that tended to be sloping for soybeasn. However, there was no perfect upward trend in production because of data fluctuations or oscillation patterns occurring in all agricultural commodities.

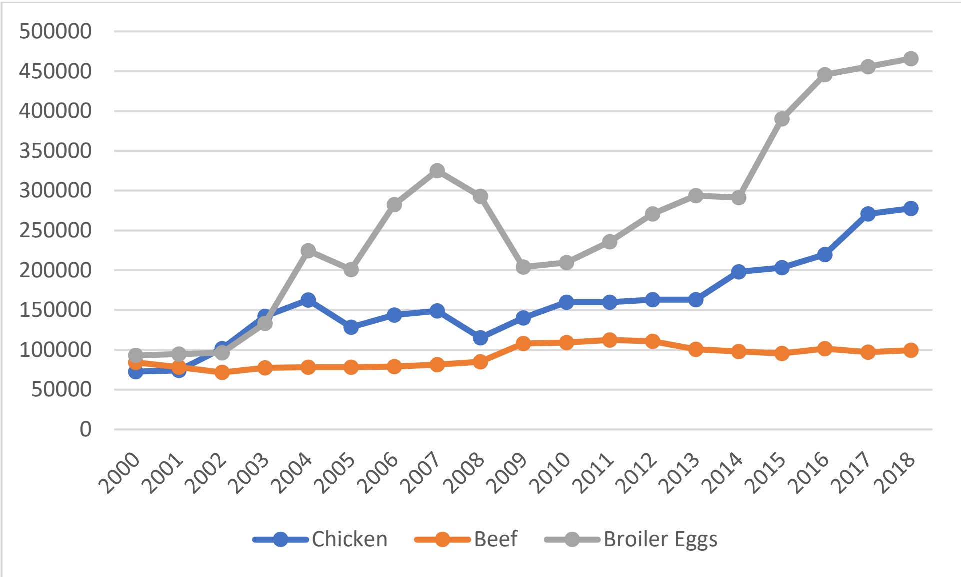


**Figure 2.** Development of strategic commodity production for the livestock sector

**Figure 2** shows the development of strategic commodity production in the East Java livestock sector which consists of chicken, beef and broiler eggs. Based on **Figure 2**, it can be seen that the trend pattern of increasing national strategic commodity production in East Java livestock sector. The increasing trend with a rapid and significant production rate occured in broiler eggs, followed by chicken. For beef, the rate of increase was not significant and since 2002, beef production was the lowest compared to broiler eggs and chicken. However, there was no perfect upward trend in production because data fluctuations or oscillation patterns occur in all livestock commodities.

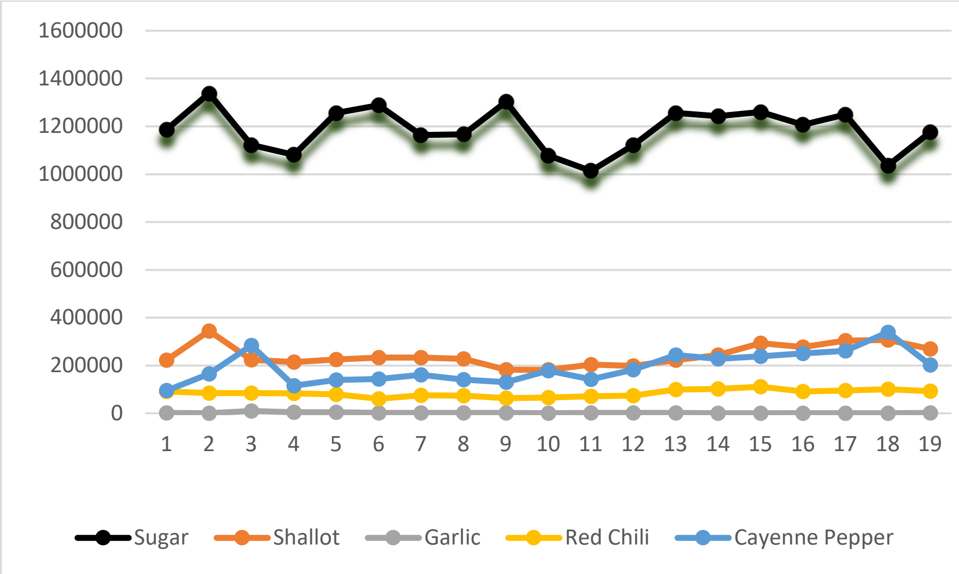


**Figure 3.** Development of strategic commodity production for plantation sector

**Figure 3** shows the development of strategic commodity production in the East Java horticultural plantation sector which consists of sugar, shallot, garlic, cayenne pepper, and red chili. Based on **Figure 3**, it can be seen that a clear fluctuation pattern exists for each commodity. Sugar was a strategic plantation commodity with the highest production, if compared with other commodities. A clearer upward trend occurred in shallot, cayenne pepper and red chilli, but those were not very significant. Sloping and invisible trends occured in garlic commodities because garic had the lowest production in East Java, if compared with other commodities.

Furthermore, because the production of eleven strategic commodities in East Java is modeled simultaneously, it must be confirmed statistically the eleven of commodities have correlation each other in production result. In other words, there is a dependent relationship between the production results from one commodity with another commodity. The test is carried out using the Bartlett Sphericity test based on Johnson & Wichern (2007), by involving the correlation matrix (**R**) which is presented in **Table 4**. **Table 4** shows that not all relationships between variables have a high correlation, so it is necessary to confirm the relationship with the Bartlett Sphericity test with the following hypothesis:

Hypothesis

$H_0$: Between variables have correlation with each other

$H_1$: Between variables do not have correlation with each other

Test Statistics:

$$G^2 = -\left\{n - 1 - \frac{2q+5}{6}\right\} ln|\mathbf{R}|$$
$$= -\left\{19 - 1 - \left(\frac{16}{6}\right)\right\} ln|0.02278|$$
$$= 163.8777$$

compared with $\chi^2_{0.05,55}$= 73.3115, because $G^2 > \chi^2$, so $H_0$ does not rejected. It means that between variables have correlation with each other. Thus, simultaneous modeling can be carried out for all commodities.

**Table 4.** The structure of correlation matrix

| | Y1 | Y2 | Y3 | Y4 | Y5 | Y6 | Y7 | Y8 | Y9 | Y10 | Y11 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Y1 | 1 | 0.3257 | -0.233 | 0.1044 | -0.116 | -0.127 | -0.247 | -0.261 | -0.018 | 0.1802 | -0.236 |
| Y2 | 0.3257 | 1 | -0.241 | 0.615 | 0.5025 | 0.2705 | 0.3672 | -0.068 | 0.3499 | 0.0327 | -0.16 |
| Y3 | -0.233 | -0.241 | 1 | -0.114 | 0.0745 | -0.511 | -0.259 | -0.518 | -0.445 | -0.28 | -0.311 |
| Y4 | 0.1044 | 0.615 | -0.114 | 1 | 0.6249 | 0.5335 | 0.4644 | 0.2565 | 0.3479 | 0.1719 | 0.1541 |
| Y5 | -0.116 | 0.5025 | 0.0745 | 0.6249 | 1 | 0.6602 | 0.593 | 0.3279 | 0.5249 | -0.056 | 0.3289 |
| Y6 | -0.127 | 0.2705 | -0.511 | 0.5335 | 0.6602 | 1 | 0.669 | 0.7666 | 0.6847 | 0.1861 | 0.5776 |
| Y7 | -0.247 | 0.3672 | -0.259 | 0.4644 | 0.593 | 0.669 | 1 | 0.4856 | 0.8916 | -0.264 | 0.2124 |
| Y8 | -0.261 | -0.068 | -0.518 | 0.2565 | 0.3279 | 0.7666 | 0.4856 | 1 | 0.4283 | 0.4146 | 0.4864 |
| Y9 | -0.018 | 0.3499 | -0.445 | 0.3479 | 0.5249 | 0.6847 | 0.8916 | 0.4283 | 1 | -0.334 | 0.1577 |
| Y10 | 0.1802 | 0.0327 | -0.28 | 0.1719 | -0.056 | 0.1861 | -0.264 | 0.4146 | -0.334 | 1 | 0.1174 |
| Y11 | -0.236 | -0.16 | -0.311 | 0.1541 | 0.3289 | 0.5776 | 0.2124 | 0.4864 | 0.1577 | 0.1174 | 1 |

***The Selected Model***

The best Fourier series estimator based on three forms of Fourier series components, cosine Fourier series, sine Fourier series, and complete Fourier series, i.e. Fourier series with cosine and sine components, are used to predict national commodities production in East Java simultaneously. Before predicting the step, the model is determined by using a sample of the data. The 14 periods were done to get the sample of the data. The periods were conducted from 2000 until 2013. **Table 5** presents the GCV value for oscillation parameter inputted for cosine Fourier series, with visually supported by **Figure 4**. According to **Table 5**, the minimum GCV value is 0.0001 with k equal to 12. **Figure 4** shows the change in GCV for k = 1 to k = 13. Based on **Figure 4**, it can be seen that there is a decrease in GCV value until the optimal oscillation parameters, then after reaching the optimal oscillation parameters, the GCV value is increasing. For k equal to 12, the smallest MSE value is 0.25167 and the greatest coefficient of determination is 100%.

**Table 5.** The result for cosine Fourier series

| k | GCV | MSE | $R^2$ |
|---|---|---|---|
| 1 | 7,465,699 | 18.2113 | 0.996351 |
| 2 | 5,252,616 | 15.5036 | 0.996894 |
| 3 | 4,131,492 | 15.0549 | 0.996983 |
| 4 | 3,216,984 | 14.8363 | 0.997027 |
| 5 | 2,226,587 | 13.4122 | 0.997313 |
| 6 | 1,152,493 | 9.4419 | 0.998107 |
| 7 | 776,642 | 9.1693 | 0.998163 |
| 8 | 469,538 | 8.6618 | 0.998264 |
| 9 | 196,136 | 6.4324 | 0.998711 |
| 10 | 34,624 | 2.5549 | 0.999488 |
| 11 | 5,363 | 1.583 | 0.999683 |

| | | | |
|---|---|---|---|
| 12 | 0.0001 | 0.25167 | 1 |
| 13 | 1,819,833 | 10.8198 | 0.997704 |

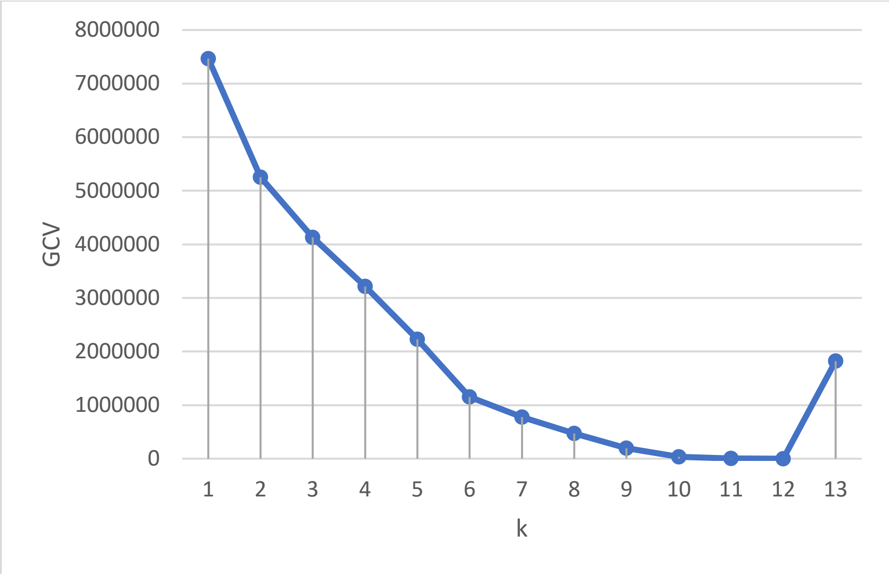


**Figure 4.** Plot oscillation parameter versus GCV for cosine Fourier series estimator

**Table 6** presents the GCV value for oscillation parameter inputted for sine Fourier series, with visually supported by **Figure 5**. According to **Table 6**, the minimum GCV value is 0.0004 with k equal to 12. **Figure 5** shows the change in GCV for k = 1 to k = 13. Based on **Figure 5**, it can be seen that there is a decrease in GCV value until the optimal oscillation parameters, then after reaching the optimal oscillation parameters, the GCV value is increasing. For k equal to 12, the smallest MSE value is 0.6325 and the greatest coefficient of determination is 100%. **Table 7** presents the GCV value for oscillation parameter inputted for Fourier series with cosine and sine components, with visually supported by **Figure 6**. According to **Table 7**, the minimum GCV value is 0.00001 with k equal to 6. **Figure 6** shows the change in GCV for k = 6 to k = 7. Based on **Figure 6**, it can be seen that there is a decrease in GCV value until the optimal oscillation parameters, then after reaching the optimal oscillation parameters, the GCV value is increasing.

For k equal to 12, the smallest MSE value is 0.1621 and the greatest coefficient of determination is 100%.

**Table 6.** The result for the sine Fourier series

| k | GCV | MSE | $R^2$ |
|---|---|---|---|
| 1 | 7,572,054 | 18.4708 | 0.9963 |
| 2 | 5,895,690 | 17.4017 | 0.996513 |
| 3 | 4,242,819 | 15.4606 | 0.996902 |
| 4 | 3,301,676 | 15.2269 | 0.996949 |
| 5 | 2,340,855 | 14.1005 | 0.997175 |
| 6 | 575,973 | 4.7223 | 0.999054 |
| 7 | 366,889 | 4.3316 | 0.999132 |
| 8 | 218,208 | 4.0235 | 0.999193 |
| 9 | 192,940 | 3.4732 | 0.999304 |
| 10 | 26,434 | 1.9506 | 0.999609 |
| 11 | 1,929 | 0.9543 | 0.999886 |
| 12 | 0.0004 | 0.6325 | 1 |
| 13 | 850,898 | 8.509 | 0.998252 |

**Table 7.** The result for the Fourier series with a cosine and sine basis

| k | GCV | MSE | $R^2$ |
|---|---|---|---|
| 1 | 5,991,716 | 17.65811 | 0.99646 |
| 2 | 3,013,462 | 13.89768 | 0.997215 |
| 3 | 1,446,508 | 11.85974 | 0.997624 |
| 4 | 629,324 | 11.60943 | 0.997674 |
| 5 | 114,825 | 8.47288 | 0.998302 |
| 6 | 0.00001 | 0.1620956 | 1 |
| 7 | 52,507 | 1.36524 | 0.99985 |

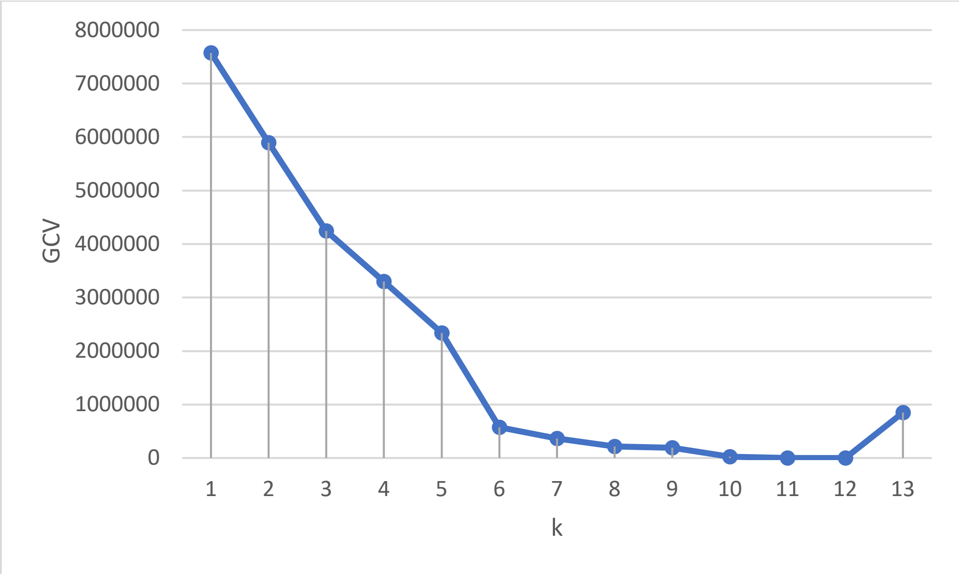


**Figure 5.** Plot oscillation parameter versus GCV for the sine Fourier series estimator

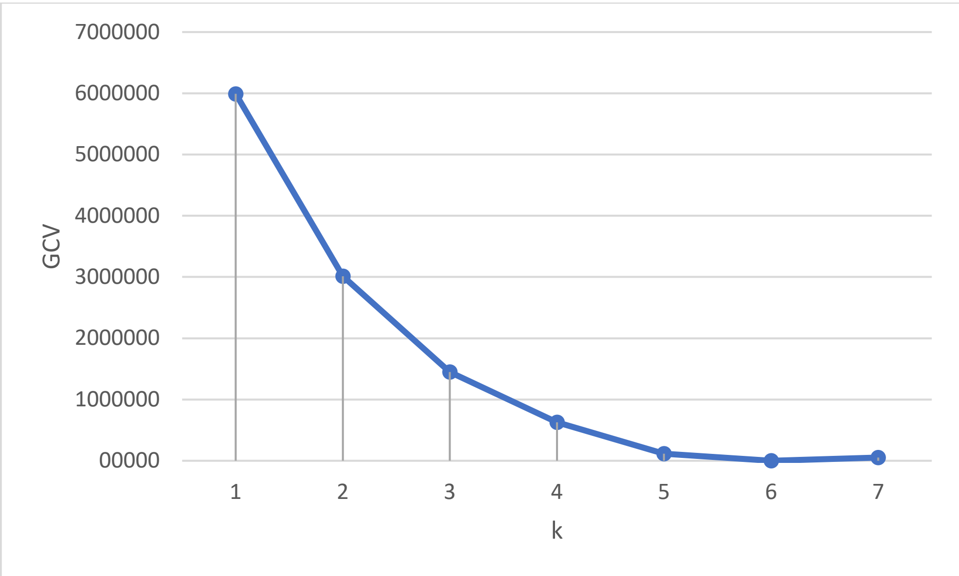


**Figure 6.** Plot oscillation parameter versus GCV for the Fourier series estimator

Using the principle of the parsimony model, the selected estimator for the prediction is the Fourier series estimator that includes cosine and sine with smaller oscillation parameters, the smallest MSE and GCV value when compared to the Fourier series estimators which only contain the cosine and sine bases. Based on the results of calculations using OSS-R, all of estimated parameter values in the model are presented in **Table 8.** So, the reason to select an estimator based on the parsimony model is reasonable. By selecting the parsimony model it makes it easier in computing complexity and application. The estimator form for selected model can be stated into equation (5) with general form for each response variables are accorded in equation (4).

$$\begin{cases} \hat{y}_{i1} = \dfrac{\hat{\alpha}_{01}}{2} + \hat{\gamma}_1 t_{i1} + \sum_{k=1}^{6} (\hat{\alpha}_{k1} \cos k\, t_{i1} + \hat{\beta}_{k1} \sin k\, t_{i1}) \\ \hat{y}_{i2} = \dfrac{\hat{\alpha}_{02}}{2} + \hat{\gamma}_2 t_{i1} + \sum_{k=1}^{6} (\hat{\alpha}_{k2} \cos k\, t_{i2} + \hat{\beta}_{k2} \sin k\, t_{i2}) \\ \vdots \\ \hat{y}_{i11} = \dfrac{\hat{\alpha}_{011}}{2} + \hat{\gamma}_{11} t_{i11} + \sum_{k=1}^{6} (\hat{\alpha}_{k11} \cos k\, t_{i11} + \hat{\beta}_{k11} \sin k\, t_{i11}) \end{cases} \quad (5)$$

The estimator in equation (5) whose parameter estimator values are presented in **Table 8** is used to predict the production of national strategic commodities in East Java.

**Table 8.** The estimated parameter value for selected model

| Parameter | Y1 | Y2 | Y3 | Y4 | Y5 | Y6 |
|---|---|---|---|---|---|---|
| alpha0/2 | -57,872,130 | -50,777,300 | -1,087,570 | -4,583,558 | -65,038,150 | -51,260,900 |
| gamma | 29,442.91 | 25,425.49 | 543.2325 | 2,328.057 | 32,508.97 | 30,764.4 |
| alpha1 | -62,285.76 | -29,968.95 | 1,018.104 | 1,709.568 | -18,067.66 | -125,293.4 |
| alpha2 | 85,311.7 | 18,501.88 | -783.704 | -2,972.58 | -20,579.08 | -329,665.5 |
| alpha3 | -52,062.46 | -21,305.78 | 1,809.455 | 1,563.98 | 30,955.48 | -234,380.2 |
| alpha4 | 46,068.1 | 15,946.55 | -782.769 | 771.0075 | -2,490.481 | -117,674.1 |
| alpha5 | -4,318.924 | -30,893.01 | 808.5473 | 5,793.541 | -14,313.2 | -129,156.2 |
| alpha6 | 94,272.95 | 84,064.29 | 2,233.06 | 23,853.38 | 102,487.8 | 235,450 |
| beta1 | 101,219.3 | 2,331.753 | -1,954.86 | 2,919.572 | -26,075.24 | -125,409.9 |
| beta2 | 71,129.18 | -1,962.385 | 534.363 | -3,449.68 | 40,990.63 | 54,443.27 |
| beta3 | -23,834.99 | -508.9179 | -194.759 | -1,635.38 | 615.3941 | -190,252.1 |
| beta4 | -5,061.696 | -5,646.006 | -741.348 | -2,548.73 | -37,693.43 | -184,100.8 |
| beta5 | -94,829.1 | 8,168.842 | 1,678.955 | 1,735.739 | 30,063.92 | -302,755.2 |
| beta6 | -146,402.1 | -117,253.4 | -3,880.54 | -4,522.53 | -111,335 | 1,576,820 |

| Y7 | Y8 | Y9 | Y10 | Y11 | Parameter |
|---|---|---|---|---|---|
| -20,231,860 | 7,565,459 | -18,762,050 | 2,472,795 | -68,919,150 | alpha0/2 |
| 10,145.18 | -3,722.72 | 9,428.616 | -1,055.18 | 36,793.62 | gamma |
| 24,055.2 | 5,155.236 | -9,878.933 | 7,640.868 | 216,439.9 | alpha1 |
| -2,041.773 | -1,483.24 | 20,111.88 | -4,047.88 | -85,968.95 | alpha2 |
| 7,527.194 | 143.0394 | 11,425.52 | 5,249.304 | -308,690.6 | alpha3 |
| 873.6749 | 1,922.172 | 7,376.735 | 15,888.79 | -69,071.81 | alpha4 |
| 13,101.07 | -1,012.04 | 31,552.41 | -29,954.2 | 139,015 | alpha5 |
| -5,117.928 | -2,974.55 | -103,151.6 | 45,869.31 | 872,26.35 | alpha6 |
| 77.97743 | -2,720.53 | 52,631.55 | 3,365.008 | -9,172.392 | beta1 |
| -12,271.6 | 785.4149 | -16,729.42 | 16,775.95 | 157,174.7 | beta2 |
| 1,662.821 | 1,551.248 | -7,531.175 | 1,067.178 | -70,661.84 | beta3 |
| -3,867.478 | -1,881.73 | 3,398.408 | -5,323.23 | -74,889.22 | beta4 |
| 5,131.534 | 263.6159 | -14,451.07 | -8,132.53 | -197,939.6 | beta5 |
| -36,277.91 | 36,939.71 | -28,635.5 | 37,696.14 | 909,129.7 | beta6 |

## *Predicting Results*

Prediction of national strategic commodity production is carried out simultaneously based on equations (5) using out sample data. There are two steps to the process of prediction, which first compares the out sample data with the predicted results in the year corresponding to the out sample data. The year used as a prediction is 2014 to 2018. The prediction results are very good because the predicted value is the same as the value of the out sample data with Mean Absolute Percentage Error (MAPE) 0.005637 %, it means that the accuracy of the prediction is very good. This is evidenced by the plots those be presented separately for the leading commodities of each sector, **Figure 7** for the agricultural sector, **Figure 8** for the livestock sector and **Figure 9** for the horticultural plantation sector which has a very large value so that some of the plot data is not visible. Thus, equation (5) can be used to predict the production of strategic commodities in East Java from the current period to the future. The prediction is carried out until 2024. Based on the prediction, some commodities need to be maintained so that production does not fall, and import activities can be reduced, the future demands of the people can be satisfied**.**

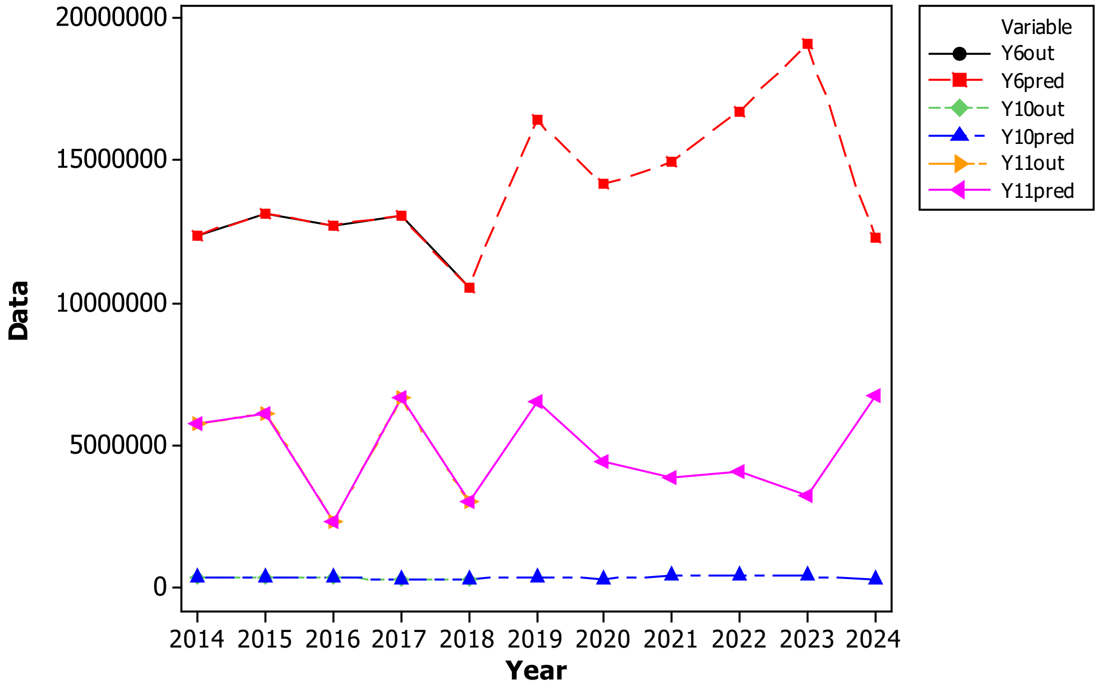


**Figure 7.** Prediction of strategic commodity production for agriculture sector

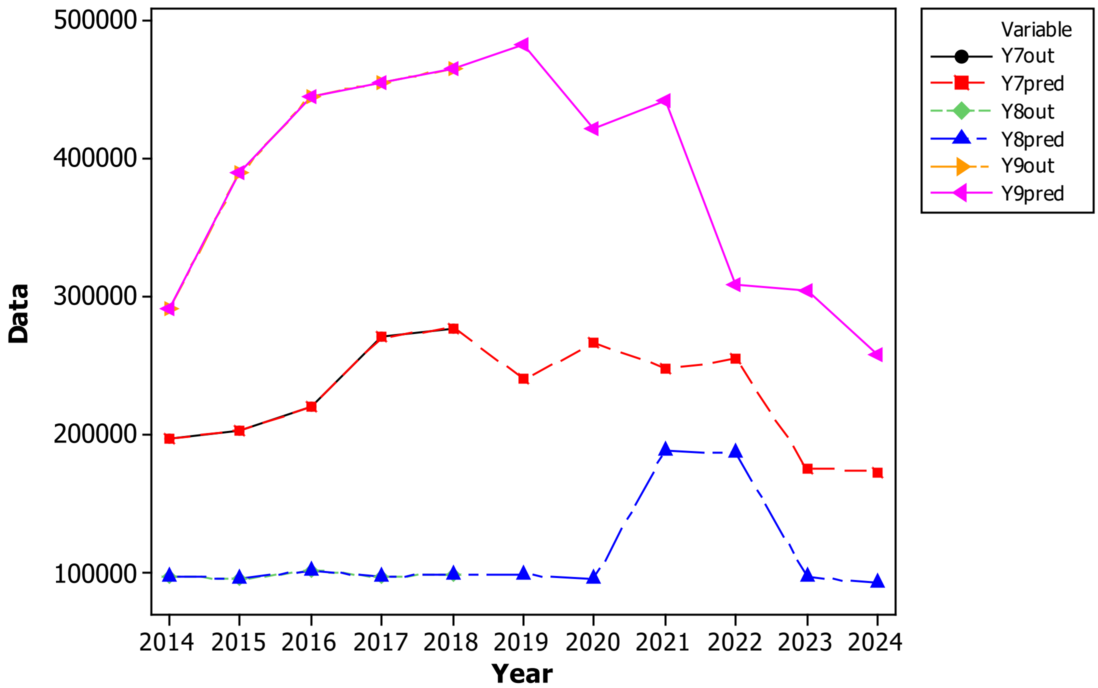

**Figure 8.** Prediction of strategic commodity production for livestock sector

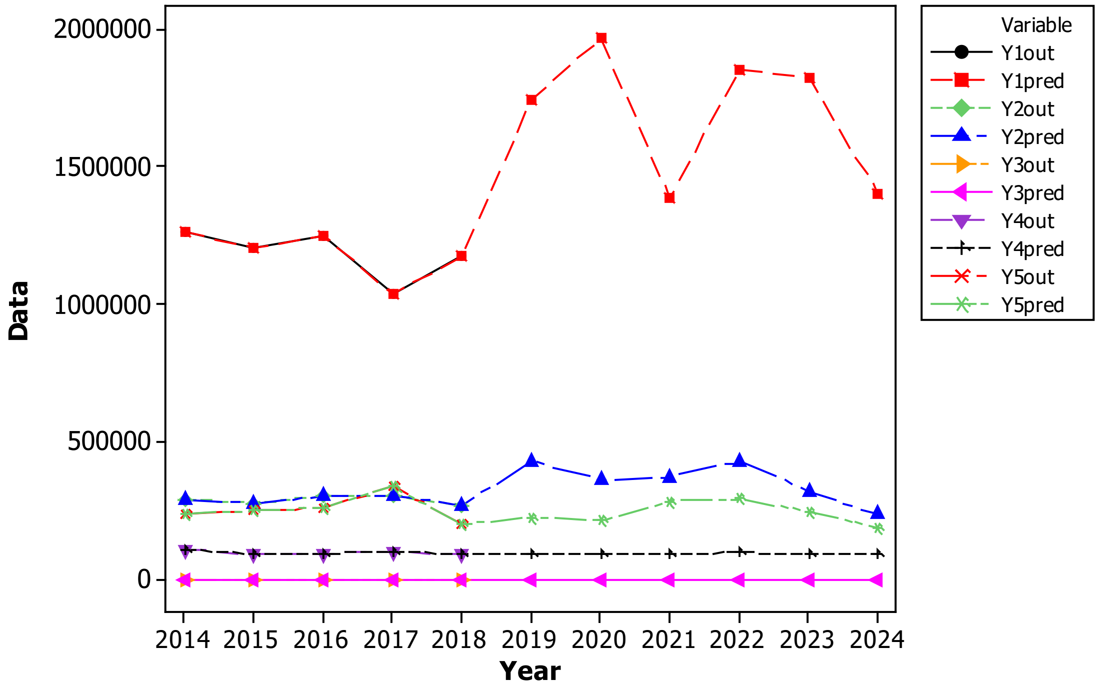


**Figure 9.** Prediction of strategic commodity production for plantation sector

## Conclusion

The production of national strategic commodities in East Java has fluctuated each year, but most of them have an increasing pattern of production results, especially in the agriculture, livestock and horticultural plantation sectors. Agricultural commodities, especially rice followed by corn, and sugar is the highest production in East Java. In the livestock sector, broiler eggs and chicken is the highest production in East Java. However, in the horticultural plantation sector in East Java, garlic production is very low. The production of national strategic commodities in East Java have correlation with each other, so that simultaneous modeling is carried out. Nonparametric regression based on the Fourier series estimator for multi-response cases is used for the prediction of 11 national strategic commodity production in East Java. The best model is chosen based on the minimum GCV value and the parsimony model concept. The selected model is achieved when k equals to 6, with the Fourier series estimator that contain cosine and sine bases. The selected model has goodness measure values like GCV equals to 0.00001, MSE equals to 0.1621 and $R^2$ equals to 100%. The selected model is used for predicting national strategic commodities in East Java with using out sample data. The model has a goodness of prediction measure like the small value of MAPE equals to 0.005637%. Based on the prediction result, the Government needs to pay attention to a number of commodities so that production can be mantained, import activities can be reduced, and future Indonesian demands can be fulfilled. So, food security in Indonesia will have better achievement.

**Acknowledgements**

The authors would like to give high appreciation for Airlangga University that funded this publication, East Java Agriculture Office and East Java Animal Husbandry Office for the valuable opportunities during the data collection.

**References**

Ahmad, H.A. (2011). Egg Production Forecasting: Determining Efficient Modelling Approaches. *Journal of Applied Poultry Research*, *4* (20), 463–473. https://doi.org/10.3382/japr.2010-00266.

Biedermann, S., Dette, H., & Hoffmann, P. (2009). Constrained Optimal Discrimination Designs for Fourier Regression Models, *Ann Inst Stat Math Journal*, *61* (1), 143–157. https://doi.org/10.1007/s10463-007-0133-5.

Bilodeau, M. (1992). Fourier Smoother and Additive Models. *The Canada Journal of Statistics*, *20* (3). 257–269. https://doi.org/10.2307/3315313.

Bloomfield, P. (2000). *An Introduction Fourier Analysis for Time Series.* New York: John Wiley & Sons.

Budiantara, I. N., Ratnasari, V., Zain, I., Ratna, M., & Mardianto, M. F. F. (2015). Modeling of HDI and PQLI in East Java (Indonesia) using Biresponse Semiparametric Regression with Fourier Series Approach. *Asian Transactions on Basic and Applied Sciences Journal*, *5*(4), 21–28.

Central Beureau of Statistics. (2017). *Indonesia Statistics 2017.* Jakarta: Central Beureau of Statistics in Indonesia.

Dette, H., Melas, V.B., & Shpilev, P. (2016). T -optimal Discriminating Designs for Fourier Regression Models. *Journal of Computational Statistics and Data Analysis*, *113* (26), 196-206. https://doi.org/10.1016/j.csda.2016.06.010.

Eubank, R.L. (1999). *Spline Smoothing and Nonparametric Regression 2nd Edition*. New York: Marcel Dekker.

Hanafie, R. (2010). *Pengantar Ekonomi Pertanian (text in Indonesia)*. Yogyakarta: Andi Offset.

Koide, N., Robertson, A.W., Ines, A.V.M., Qian, J.H., Witt, D.G.D., & Lucero, A. (2011). Prediction of Rice Production in Philippines using Seasonal Climate Forecasts, *Journal of Applied Meteorology and Climatology*, *11*, 1 – 46. https://doi.org/10.1175/JAMC-D-11-0254.1.

Mardianto, M. F. F., Cahyono, E. F., Syarifah, L., & Andriani, P. (2019). Prediction of the Number of Foreign Tourist Arrival in Indonesia Halal Tourism Entrance using Simultaneously Fourier series Estimator. *Proceedings of the 2nd International Conference on Islamic Economics, Business, and Philanthropy*, *KnE Social Sciences,* 1093–1104.

Mardianto, M. F. F., Kartiko, S. H., & Utami, H. (2019a). Forecasting Trend-Seasonal Data Using Nonparametric Regression with Kernel and Fourier Series Approach. *Proceedings of the Third International Conference on Computing, Mathematics and Statistics (ICMS2017), Langkawi Malaysia. Springer*, 343–349. https://doi.org/10.1007/978-981-13-7279-7_42.

Mardianto, M. F. F., Kartiko, S. H., & Utami, H. (2019b). Prediction the Number of Students in Indonesia who Study in Tutoring Agency and Their Motivations based on Fourier Series Estimator and Structural Equation Modelling. *International Journal of Innovation, Creativity and Change, 5* (3), 343–349.

Mardianto, M. F. F., Tjahjono, E., & Rifada, M. (2019a). Semiparametric regression based on three forms of trigonometric function in Fourier series estimator. *Journal of Physics: Conference Series. IoP*, *1227* (1), 1–10.

Mardianto, M. F. F., Tjahjono, E., & Rifada, M. (2019b). Statistical modelling for prediction of rice production in Indonesia using semiparametric regression based on three forms of Fourier series estimator. *ARPN Journal of Engineering and Applied Sciences*, *14* (15), 2763 – 2770.

Mardianto, M. F. F., Tjahjono, E., Rifada, M., Herawanto, A., Putra, A. L., & Utama, K. A. (2019). The Prediction of Rice Production in Indonesia Provinces for Developing Sustainable Agriculture. *Proceedings of the 1st International Conference on Food and Agriculture, Denpasar*, 325–333.

Penedo, A.B.T. (2014). Forecast of Production of Cane Sugar in Brazil. *Asian Journal of Business and Management Sciences*, *3* (9), 8 – 14.

Septiawan, R., Komarudin, A., & Sulistya, B. (2012). Prediction Model for Chili Productivity Based on Climate Data. *Proceedings IEEE of 6th Computer Modeling and Simulation (CMS) Symposium, Malta, IEEE*. https://doi.org/10.1109/EMS.2012.67.

Takezawa, K. (2005). *Introduction to Nonparametric Regression*. New Jersey: John Wiley & Sons.

The Ministry of Agriculture. (2005). *2015 Policy Analysis Report: National Strategic Food Policy Outlook for 2015 - 2019*. Jakarta: The Ministry of Agriculture in Indonesia.

Tjahjono, E., Mardianto, M. F. F., & Chamidah, N. (2018). Prediction of Electricity Consuption using Fourier Series Estimator in Bi-Response Nonparametric Regression Model. *Far East Journal of Mathematical Sciences (FJMS)*, *103* (8), 1251–1263. http://dx.doi.org/10.17654/MS103 08125 1.

Wahba, G. (1990). *Spline Model for Observational Data*. Philadelphia: SIAM XII.

Wu, H., & Zhang, J. T. (2006). *Nonparametric Regression Methods for Longitudinal Data Analysis*. New Jersey: John Wiley & Sons.