

# Spatially Filtered Ridge Regression Modeling to Find Out the Rice Production Factors in East Java, Indonesia

Vita Dewi Islami<sup>1</sup>, Rahma Fitriani<sup>2\*</sup>, and Henny Pramoedyo<sup>3</sup>

<sup>1–3</sup>Master Program of Statistics, Faculty of Mathematics and Natural Sciences, Universitas Brawijaya

Jawa Timur 65145, Indonesia

Email: <sup>1</sup>vitadewi17@student.ub.ac.id, <sup>2</sup>rahmafutriani@ub.ac.id, <sup>3</sup>pramoedyohp@yahoo.com

**Abstract**—The research aims to model rice production in East Java using the Spatially Filtered Ridge Regression (SFRR) method and ensure that all violations of assumptions are resolved by knowing the direct and indirect effect of predictor variables. The data are secondary data sourced from the publication of Badan Pusat Statistik containing provincial food crop agriculture statistics in East Java and the 2018 publication of Dinas Pertanian Jawa Timur (literally translated as Agriculture Department of East Java). The data analysis process is done by RStudio and ArcMap 10.3 software. In the research, the observation unit is 38 regencies or cities in East Java. The analysis results show that SFRR with queen contiguity weighting can overcome spatial autocorrelation and multicollinearity in rice production data in East Java. As for the established model, the variables of rice field area, urea fertilizer, Phonska fertilizer, SP-36 fertilizer, and tractor have a significant effect on rice production. However, ZA fertilizer has no significant effect on rice production. Then, a large comparison of direct and indirect impacts for each predictor variable is also generated. Generally, direct impacts are greater than indirect impacts.

**Index Terms**—Spatially Filtered Ridge Regression, Rice Production, East Java

## I. INTRODUCTION

THE agricultural sector is essential in the development of the national economy. The role of the agricultural sector includes providing food and industrial raw materials, absorbing labor, and being the main source of rural household income. As a food provider, the agricultural sector should be constantly supervised because it is an important need for humans. Rice is the main food commodity of Indonesian society that almost the entire population in the country consumes rice every day [1].

The agricultural sector is closely related to geographical conditions. Geographical conditions include altitude, air temperature, soil type, soil pH, and water content, which are suitable for certain types of plants. Based on those conditions, the development of rice crops is related to spatial effects. Tobler proposed the first law on geography that the condition at one point or area related to other adjacent points or areas [2]. Like rice production, if an area has high rice production, the surrounding area will have high rice production potential. It is because the adjacent areas have almost the same geographical conditions. It indicates that there is spatial autocorrelation in rice production.

Rice production factors include the area of rice fields, the amount of labor, fertilizer, pesticides, and irrigation systems [3]. Among the factors that influence rice production, there is a correlation between predictor variables that the wider the rice fields are cultivated, the more the seeds and fertilizers are needed. If there are cases of two or more highly correlated predictor variables, it is difficult to separate the influence of each variable into a response variable called multicollinearity. If the issue is not addressed, it may lead to an estimator with a large variance or standard error. Also, multicollinearity issues cause high  $R^2$  values, but many predictor variables have no statistical effect. One method of analysis that can be used to solve multicollinearity is the ridge method [4].

Some assumptions must be fulfilled in the multiple linear regression model: the residuals should be normally distributed with a zero mean, and the variance should be consistent and not correlate with themselves or other variables [4]. If the non-autocorrelation assumptions are violated, the multiple linear regression model is not properly used. They will cause an estimator with large variance, so another model is required that can accommodate violations of non-

Received: Aug. 31, 2020; received in revised form: Nov. 11, 2020; accepted: Nov. 11, 2020; available online: Nov. 23, 2020.  
\*Corresponding Author

autocorrelation assumptions. Spatial Autoregressive (SAR) is a model to cope with observation value in a location depending on the observation value in other locations that are close to each other [5].

If SAR has a violation of non-multicollinearity assumptions, another method is needed to overcome the deviation of such assumptions. Spatially Filtered Ridge Regression (SFRR) is a combination of SAR and ridge method. It obtains stable parameter guessing value and resistant to multicollinearity problems and dependency between observation areas (spatial autocorrelation). The previous research studies the factors affecting the temperature in land cover issues on Urban Heat Island (UHI). Most of the methods used in these cases are Ordinary Least Squares (OLS), but there are cases of violated non-multicollinearity. It causes parameter estimations value to vary from one sample to another. Large sample variability results in an increased variance of estimation. The predictors should have a significant but insignificant effect. When the parameters to measure the degree of spatial spillover ( $\rho$ ) are specified, it is substituted into  $(I - \rho W)Y$ . The result is used as a response variable for the ridge regression method. The research shows that the SFRR model can produce the least Mean Squared Error values and Akaike’s Information Criterion (AIC) compared to spatial regression, Spatial Autoregressive Model and Ridge Models [6]. It is a guide to the selection of Spatially Filtered Ridge Regression models.

In addition to information on spatial autocorrelation between observation units based on the position of each unit that is related to each other, it is interesting to know the direct and indirect impacts. Research on rice production with spatial influences is conducted by [7] with the Geographically Weighted Regression (GWR) model. The research also includes Phonska fertilizer and harvest ingestion, which influences the increase in rice production in East Java. The other previous research mentions that rice field area, harvest area, and rice productivity significantly affect rice production with Spatial Durbin Error Models (SDEM) approach [8]. Next, the amount of used fertilizer also has a significant effect on rice productivity. The increase in rice production can also be influenced by agricultural machinery that farmers use [9]. Similarly, previous research shows the effect of the tractor on agricultural income. It mentions that using tractors can reduce the labor to accelerate land preparation activities, implement timely planting in the growing season, and increase rice production [10].

The research aims to model rice production in East Java using SFRR method. It addresses the problem of multicollinearity and spatial autocorrelation on rice production. It is expected to assist the government in

TABLE I  
RESEARCH VARIABLES.

Notation	Variable	Unit
$Y$	Rice Production Results	Tons
$X_1$	Rice Field Area	ha
$X_2$	Urea Fertilizer	Tons
$X_3$	Phonska Fertilizer	Tons
$X_4$	SP-36 Fertilizer	Tons
$X_5$	ZA Fertilizer	Tons
$X_6$	Tractor	Unit

supervising areas with low rice production so that there is no shortage of rice, especially in highly populated areas with no potential as a rice producer.

## II. RESEARCH METHOD

### A. Data Sources and Research Variables

The used data are secondary data sourced from the Badan Pusat Statistik ‘Central Bureau of Statistics’ (BPS) containing East Java Provincial Food Crop Agriculture Statistics and the publication of Dinas Pertanian Jawa Timur 2018. The data analysis process is done using RStudio and ArcMap 10.3 software. In this research, the observation unit is 38 regencies or cities in East Java. The selection of location is because it is an area that has the potential to produce the largest rice in Indonesia. The variables used in the research are in Table I.

### B. Analysis Method

For the analysis, the researchers conduct several steps. First, explore the rice production distribution map ( $Y$ ) and descriptive statistical analysis of predictor variables ( $X_p$ ). Second, perform the multiple linear regression analysis so that the residual model is obtained. The multiple linear regression analysis model describes the relationship between response variables and predictor variables, as formulated by [4]. Third, define a spatial weighting matrix ( $W$ ) of standardized lines with the queen contiguity method [10]. Fourth, test spatial heterogeneity on the residuals of the multiple linear regression analysis model with the Breusch-Pagan test formulated by [11]. The test ensures that the residual elemental relationships of the multiple linear regression analysis model do not vary between locations [12]. Fifth, conduct spatial autocorrelation testing. It is performed on residuals of the multiple linear regression model and response variables of rice production. The presence of spatial autocorrelation on the residuals indicates a violation of the assumption of residual freedom at the time of hypothesis testing. It will cause the variance of estimation of correlated

residuals to no longer have a minimum variance [13]. The spatial autocorrelation test uses Moran's I test [5].

Sixth, perform multicollinearity testing using Variance Inflation Factor (VIF) value and correlation coefficients between predictor variables. Multicollinearity occurs if the VIF value is more than 10 [4]. The correlation coefficient ( $r$ ) describes the degree of closeness of the relationship between two or more variables [14]. Seventh, perform spatial filtering process with the following stages. It estimates the  $\hat{\rho}$  parameter to measure spatial spillover degrees on SAR models (Eq. (1)) [10].  $Y$  is a vector of dependent variable.  $X$  is a matrix of independent variables. Then,  $\beta$  is a vector of regression coefficient.  $I$  is a identity matrix.  $\nu$  is a vector of residual.

$$Y = (I - \rho W)^{-1}(X\beta + \nu). \quad (1)$$

When the  $\hat{\rho}$  parameter has been specified, it is substituted into Eq. (2).  $Y_{filtered}$  is used as a response variable for the ridge regression method [6]. It is a vector of observations on the spatially filtered dependent variable.

$$Y_{filtered} = (I - \rho W)Y. \quad (2)$$

Eighth, conduct SFRR model test (Eq. (3)) with several stages.  $e$  is a vector of residual.

$$Y_{filtered} = X\beta + e. \quad (3)$$

Firstly, transform it into the default form (standardization) of response variables and predictor variables like in Eq. (4) [15].  $X_{ip}^*$  is a default value of independent variables. Meanwhile,  $Y_i^*$  is a default value of dependent variable. Then,  $S_Y$  is a standard deviation for  $Y$ , and  $S_p$  is for each predictor variable.  $p$  is the total number of parameters, and  $n$  is the number of observations.

$$X_{ip}^* = \frac{X_{ip} - \bar{X}_p}{\sqrt{n - 1}S_p} \text{ and } Y_i^* = \frac{Y_i - \bar{Y}}{\sqrt{n - 1}S_Y}. \quad (4)$$

Secondly, determine bias constant ( $\lambda$ ) based on VIF value in Eq. (5). The best bias constant ( $\lambda$ ) value results in a VIF value of less than 1.

$$VIF_i = \frac{(X^{*T}X^* + \lambda I)^{-1}(X^{*T}X^*)}{(X^{*T}X^* + \lambda I)^{-1}}. \quad (5)$$

Thirdly, estimate the parameters in the SFRR model.  $\hat{\beta}_{SFRR}$  is vector parameter of SFRR ( $p \times 1$ ), and  $\lambda$  is bias constant located on the hose  $0 \leq \lambda < 1$ .  $X^*$  is a matrix of predictor variables of standardized results ( $n \times p$ ).  $Y_{filtered}^*$  is a vector response variable of standardized results ( $n \times 1$ ). The equation is as follows:

$$\hat{\beta}_{SFRR} = (X^{*T}X^* + \lambda I)^{-1}(X^{*T}Y_{filtered}^*). \quad (6)$$

Fourthly, in forming the SFRR model, the SFRR regression guess is restored to the initial model form in Eqs. (7) and (8) [6].  $\hat{\beta}_k$  is a coefficient of SFRR, and  $\hat{\beta}_0$  is an intercept coefficient.

$$\hat{\beta}_k = \frac{S_Y}{S_k} \hat{\beta}_{Rk}, \quad (7)$$

$$\hat{\beta}_0 = \hat{\beta}_{R0} - \frac{\hat{\beta}_{R1}\bar{X}_1}{S_1} \dots - \frac{\hat{\beta}_{Rp}\bar{X}_p}{S_p}. \quad (8)$$

Ninth, test the residual assumptions of the SFRR model. It is to ensure that the assumption violation is resolved. The normality test uses the Jarque-Bera (JB). The test utilizes skewness and kurtosis measures [4]. Meanwhile, the spatial autocorrelation test uses Moran's I test [5]. Tenth, perform a test of the significance of the SFRR model parameters using Wald test [5]. Last, interpret the SFRR model and determine the direct and indirect effects.

### III. RESULTS AND DISCUSSION

#### A. Rice Production Spread Pattern in East Java

Kementerian Pertanian Republik Indonesia 'Ministry of Agriculture Republic Indonesia' published the rice production improvement program results in 2018. It noted that East Java was the highest rice producer in Indonesia in 2018 with 13,000,475 tons [16]. Moreover, the spread of rice production in each regency or city in East Java Province was mapped into three categories, namely the highest, moderate, and lowest rice production. The category division was based on ArcGIS with Natural Breaks classification method.

Figure 1 shows the results of mapping rice production data in tons throughout the observation site. There are 6 regencies or cities with the highest rice producers (532,815–924,212 tons). The majority are coastal areas near the coast. Lamongan Regency has the highest rice production of 924,212 tons, followed by regencies of Bojonegoro, Ngawi, Jember, Tuban, and Banyuwangi. About 19 regencies or cities fall into moderate rice production (179,915–532,814 tons). The remaining 13 regencies or cities are classified into low rice production (4,903–179,914 tons). Areas with low rice production are mostly in urban areas, such as Mojokerto City with 4,903 tons. This condition is due to differences in geographical conditions and economic activity in each region.

#### B. Description of Predictor Variables of Rice Production in East Java in 2018

Several factors can cause a high amount of rice production in each area. It includes the area of rice fields, fertilizer, and agricultural machinery. The geographical location of a region also affects rice production.

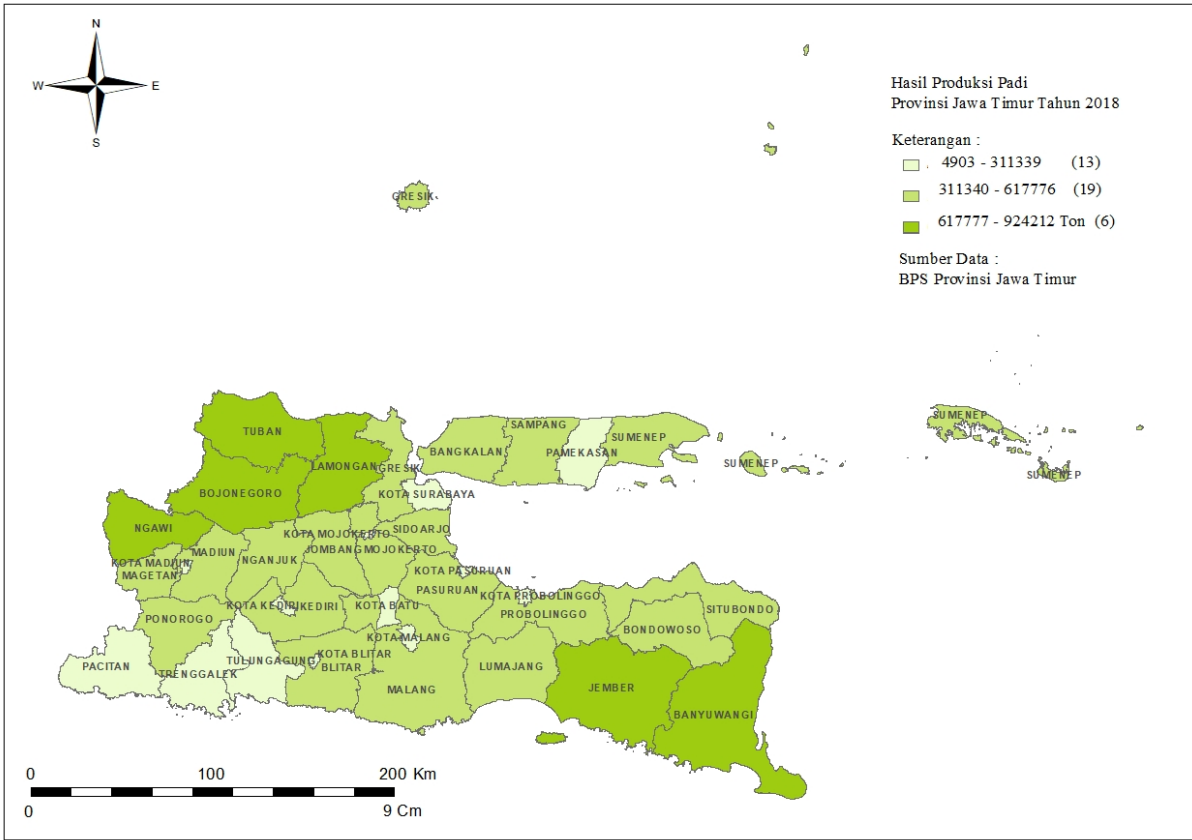


Fig. 1. East Java Province Rice Production Distribution Map in 2018. In the figure, *Hasil Produksi Padi Provinsi Jawa Timur Tahun 2018* (the rice production in East Java in 2018), *Keterangan* (description), and *Sumber Data: BPS Provinsi Jawa Timur* (Source: Central Bureau of Statistics of East Java).

TABLE II  
DESCRIPTIVE STATISTICS OF PREDICTOR VARIABLES.

Variable	Median	Min.	Max.	Standard Deviation
$X_1$ (ha)	1,606	630.89	3,664.05	696.80
$X_2$ (tons)	27,813.5	311.50	94,919.9	22,556.53
$X_3$ (tons)	12,287.4	171	43,527.8	14,636.84
$X_4$ (tons)	3,707.5	0	15,761.1	4,012.75
$X_5$ (tons)	9,545.5	52	46,351.8	12,461.51
$X_6$ (unit)	91	45	159.0	32

Table II provides information on descriptive statistics of predictor variables.

The pattern of functional relationships between predictor and response variables can be estimated by creating scatterplots that contain information about both variables. Based on Fig. 2, it provides information that the rice field area (Fig. 2a), urea fertilizer (Fig. 2b), Phonska fertilizer (Fig. 2c), SP-36 fertilizer (Fig. 2d), and ZA fertilizer (Fig. 2e) have a positive and linear pattern of a relationship with rice production. It is characterized by the distribution of data that follows the regression line. However, the relationship between

tractor (Fig. 2f) and rice production does not form linear patterns.

### C. Multiple Linear Regression Analysis

Multiple linear regression analysis ensures that there is no spatial autocorrelation and spatial heterogeneity in the residuals. The residuals in the multiple linear regression model are derived from a subtraction in the value of rice production ( $Y_i$ ) and the estimated value of the multiple linear regression analysis model ( $\hat{Y}_i$ ) as shown in Eq. (9).

$$\hat{Y}_i = -1.158 \times 10^5 + 45.59\hat{X}_1 + 4.73\hat{X}_2 + 9.54\hat{X}_3 + 9.64\hat{X}_4 - 6.87\hat{X}_5 + 816.4\hat{X}_6. \quad (9)$$

The model of multiple linear regression analysis is not interpreted since the results cannot be trusted. Based on the theory, rice production requires a relationship with spatial interaction. The multiple linear regression model is only used to obtain residual values for spatial effect testing.

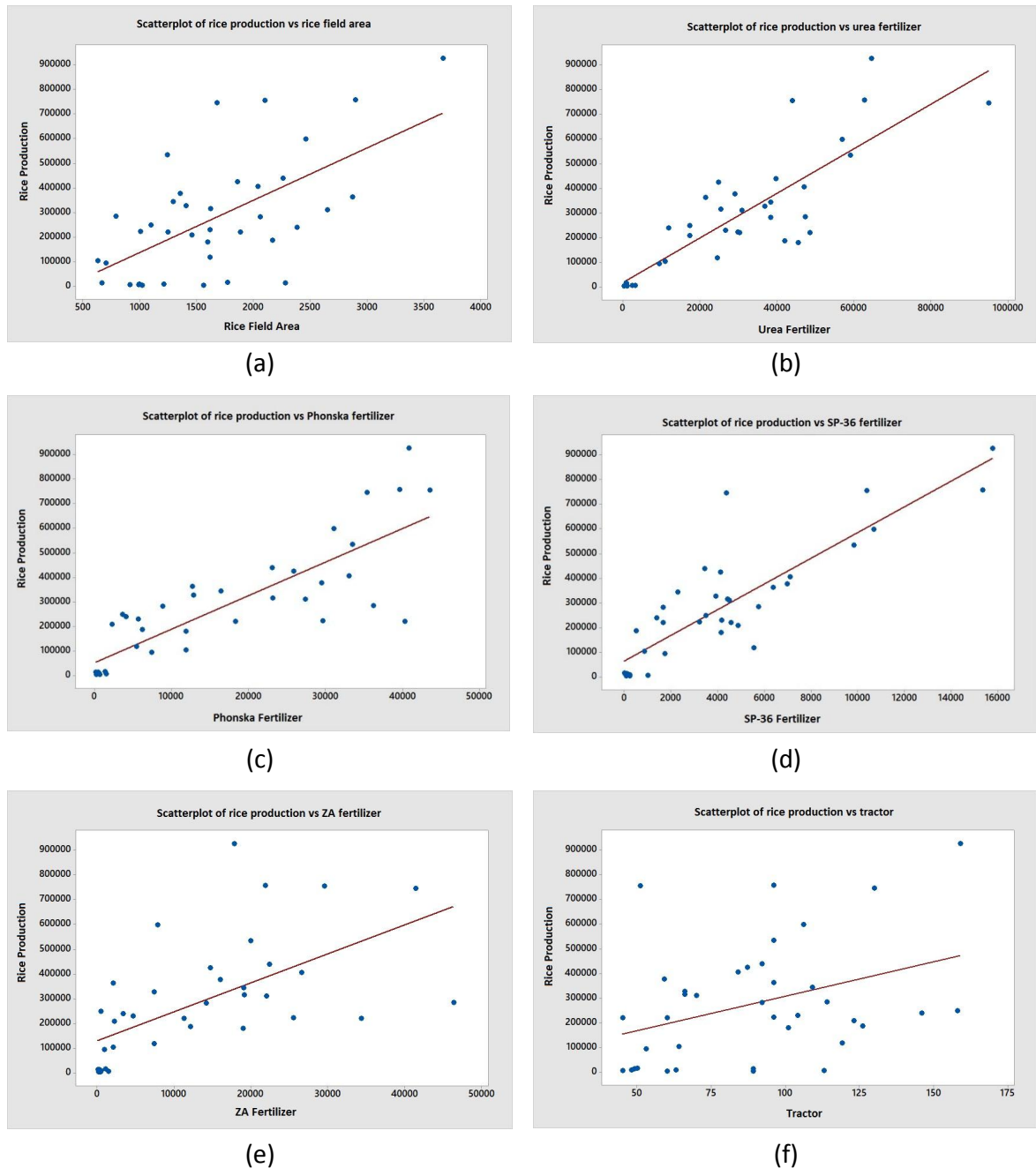


Fig. 2. Scatterplot of six predictor and response variables.

#### D. Spatial Heterogeneity Test

This test is carried out to ensure that the residual elemental relationship of the multiple linear regression model does not vary between locations. It obtains a Breusch Pagan (BP) value of  $12.4 < X^2_{(0.05;6)} = 12.59$ . It means  $H_0$  is accepted that the variance of estimation for all observations is the same. Hence, it can be concluded that there is no spatial heterogeneity

in rice production data in East Java of 2018.

#### E. Spatial Autocorrelation Test

In Fig. 1, the adjacent locations have the same color categories. Thus, there may be a spatial dependency.

In Table III, Moran's I test results on residuals of multiple linear regression model are p-value of  $< \alpha = 10\%$  so that  $H_0$  is rejected. Similarly, spatial

TABLE III  
THE RESULTS OF MORAN'S I TEST.

Data	The Moran's I Index	Expected Index	P-Value
Residuals of multiple linear regression model	0.13	-0.04	0.09000
Rice Production	0.39	-0.03	0.00051

TABLE IV  
VARIANCE INFLATION FACTOR (VIF) VALUE FOR EACH PREDICTOR VARIABLE.

Variable	VIF
Rice Field Area ( $X_1$ )	1.619
Urea Fertilizer ( $X_2$ )	5.277
Phonska Fertilizer ( $X_3$ )	13.754
SP-36 Fertilizer ( $X_4$ )	6.128
ZA Fertilizer ( $X_5$ )	8.145
Tractor ( $X_6$ )	1.726

autocorrelation test results in response variables (rice production) with p-value of  $< \alpha = 5\%$ , so  $H_0$  is also rejected. It means that the residuals have a correlation or relationship between locations of regency or city. Moran's I value is positive, so the spatial pattern formed is a grouping pattern with positive spatial autocorrelation properties.

There is spatial autocorrelation in residuals of multiple linear regression model and response variables from the results of spatial autocorrelation tests. As a result of linear regression analysis is less precise, the analysis is performed using spatial regression. The model that can be used is SAR. It is used when there is an influence of the on-site ( $j$ ) response variable on the observation location ( $i$ ).

#### F. Multicollinearity Test

Multicollinearity testing on the research is conducted using Variance Inflation Factor criteria and correlation coefficients. Based on the VIF value in Table IV, it can be seen that the Phonska fertilizer ( $X_3$ ) has a value of  $VIF > 10$ . It indicates that there is a multicollinearity issue between predictor variables. To determine the predictor variables with a relationship with Phonska fertilizer ( $X_3$ ), the researchers calculate the correlation coefficient between predictor variables.

Table V shows that a very strong correlation between the Phonska fertilizer with urea and ZA fertilizers. The correlation coefficient ( $r$ ) is close to +1 and p-value is  $0.00 < \alpha = 0.05$ . It indicates the stronger presumption of multicollinearity in the case of rice production. Phonska is a compound fertilizer containing three main nutrients namely Nitrogen, Phosphorus, and Potassium

that are indispensable for plants. If the Phonska predictor variable is issued, it will cause a specification error in the factor of increasing rice production.

#### G. Estimation of Spatial Spillover ( $\rho$ ) Degree Parameters in Spatial Autoregressive (SAR) Model

The first step in the spatial filtering process is to estimate the spatial spillover ( $\rho$ ) degree parameter of the SAR model (Eq. (1)). The estimation of SAR parameters is done by the Maximum Likelihood Estimator (MLE) method. The equation is:

$$\hat{Y}_i = -1.77 \times 10^5 + 0.301 \sum_{j=1, i \neq j}^{38} W_{ij} Y_j + 31.3X_{1i} + 4.79X_{2i} + 8.55X_{3i} + 6.54X_{4i} - 5.91X_{5i} + 7.9 \times 10^2 X_{6i}. \quad (10)$$

Based on the estimation of SAR model parameters, it obtains  $\rho$ -value of 0.301. SAR model in Eq. (10) is not interpreted. The results cannot be trusted due to the violation of non-multicollinearity assumptions.

#### H. Spatial Filtering Process in Response Variables

The spatial filtering process substitutes the  $\rho$ -value of 0.301 and the response variable (rice production) to the function  $(I - \rho W)Y$  as Eq. (2). The result is used as a new response variable for the ridge regression method.

#### I. Parameter Estimation of Spatially Filtered Ridge Regression (SFRR)

The SFRR parameter estimation utilizes the bias constant ( $\lambda$ ) and spatial filtering of the response variable ( $Y_{(filtered_i)}$ ). Forming an SFRR model is to return the SFRR parameter estimation to the original variable into Eqs. (7) and (8). The SFRR model formed using rice production data in East Java of 2018 is in Eq. (11). This SFRR model has not been interpreted before it confirms that all assumptions are met. The equation is:

$$(I - \rho W)Y = -1.558 + 41.64X_1 + 2.97X_2 + 3.81X_3 + 14.72X_4 + 0.1X_5 + 638.1X_6. \quad (11)$$

#### J. Residual Assumption Test of Spatially Filtered Ridge Regression (SFRR)

In the assumption test of SFRR, there are two stages of testing: normality and spatial autocorrelation. These tests ensure that the assumption violation is resolved. First, for the normality test, the result shows Jarque Bera values of  $1.39 < \chi^2$  of 5.99 and p-value of  $0.49 \geq \alpha$  of 0.05. It accepts  $H_0$ , meaning the residuals

TABLE V  
CORRELATION COEFFICIENT BETWEEN PREDICTOR VARIABLES.

Variable	Rice Field Area	Urea Fertilizer	Phonska Fertilizer	SP-36 Fertilizer	ZA Fertilizer	Tractor
Rice Field Area	0.000					
Urea Fertilizer	0.443					
Phonska Fertilizer	0.403	0.812	0.000			
SP-36 Fertilizer	0.591	0.702	0.774	0.000		
ZA Fertilizer	0.199	0.794	0.848	0.466	0.000	
Tractor	0.270	0.390	0.608	0.309	0.142	0.000

TABLE VI  
PARTIAL TEST RESULTS ON SPATIALLY FILTERED RIDGE REGRESSION PARAMETERS.

Parameters	Estimation	Standard Error	Wald
$\beta_0$	$4.62 \times 10^{-19}$	0.0079	$3.4 \times 10^{-33}$
$\beta_1$	0.135	0.0450	9.19
$\beta_2$	0.312	0.0480	42.30
$\beta_3$	0.259	0.0390	42.33
$\beta_4$	0.275	0.0450	36.86
$\beta_5$	$5.82 \times 10^{-3}$	0.0430	0.02
$\beta_6$	$9.43 \times 10^{-2}$	0.0420	5.04

on the SFRR model are normally distributed. Then, for spatial autocorrelation using Moran's I test. The residuals of SFRR show p-value of  $0.315 > \alpha$  of 5%. Hence,  $H_0$  is accepted implying that the residuals in SFRR have no correlation or relationship between locations of regency or city.

#### K. Parameter Test of Spatially Filtered Ridge Regression (SFRR)

Once it is confirmed that the assumption violation has been resolved. The results of the estimated parameters of the SFRR model are partially tested. It is to indicate that the parameter has a significant or insignificant effect. Partial testing uses Wald test statistics. If Wald test statistical value is more than  $X_{\alpha,1}^2$  of 3.841, then it is decided that the  $H_0$  or model parameters have a significant effect. Table VI shows that the predictor variables of rice field area, urea fertilizer, Phonska fertilizer, SP-36 fertilizer, and tractor on the SFRR model have a significant effect on rice production. ZA fertilizer has no significant effect on rice production.

#### L. Direct and Indirect Effects

Dependency between site units affects rice production in East Java. It indicates that changes in conditions of one predictor variable at one location will affect rice production at the same location (direct effect). In turn, it will affect rice production in another location (indirect effect) [12]. Direct effect values of rice field area, urea fertilizer, Phonska fertilizer, SP-36 fertilizer,

TABLE VII  
THE DIRECT EFFECT VALUE OF RICE PRODUCTION CHANGES DUE TO PREDICTOR VARIABLE CHANGES IN LAMONGAN REGENCY.

Predictor Variables	Change value
Rice Field Area (ha)	42.6960
Urea Fertilizer (tons)	3.0460
Phonska Fertilizer (tons)	3.9070
SP-36 Fertilizer (tons)	15.0920
ZA Fertilizer (tons)	0.1029
Tractor (unit)	654.3145

TABLE VIII  
THE INDIRECT EFFECT VALUE OF RICE PRODUCTION DUE TO SIX PREDICTOR VARIABLES OF LAMONGAN REGENCY TO GRESIK REGENCY.

Predictor Variables	Change value
Rice Field Area (ha)	2.748
Urea Fertilizer (tons)	0.196
Phonska Fertilizer (tons)	0.251
SP-36 Fertilizer (tons)	0.971
ZA Fertilizer (tons)	0.007
Tractor (unit)	42.100

ZA fertilizer and tractor are obtained from the diagonal matrix element of  $(I - \rho W)\beta_k$ . The values of  $\rho$  and  $\beta_k$  are from the SFRR model. For example, the application of direct effects and interpretation to Lamongan Regency is presented in Table VII.

The direct effect value indicates that an increase in the area of rice fields in Lamongan Regency by 1 ha will directly increase rice production by 42.696 tons. Then, if there is an increase in using urea fertilizer by 1 ton, it will increase rice production by 3.046 tons. The rise in Phonska fertilizer by 1 ton will have a direct effect on the increase in rice production by 3.907 tons. Similarly, the changes in SP-36 fertilizer by 1 ton will directly increase rice production by 15.092 tons. If the use of ZA fertilizer rises by 1 ton, it will have a direct effect on the increase in rice production by 0.1029 tons. Last, a change tractor by 1 unit will affect rice production about 654.3145 tons.

Moreover, the value of indirect effect changes in rice field area, urea fertilizer, Phonska fertilizer, SP-36 fertilizer, ZA fertilizer, and tractor for each regency or city is obtained from the average element instead

of diagonal matrix of  $(I - \rho W)\beta_k$ . An example of the indirect effect applied to Lamongan Regency against Gresik Regency is in Table VIII.

Every increase in rice fields in Lamongan Regency by 1 ha will indirectly increase rice production in Gresik Regency by 2.748 tons. Then, the increase in using urea fertilizer by 1 ton will have an indirect effect on the rise in rice production by 0.196 ton in Gresik Regency. Next, if there is an increase in Phonska fertilizer by 1 ton in Lamongan Regency, it will increase rice production in Gresik Regency by 0.251 tons. Similarly, the changes in SP-36 fertilizer in Lamongan Regency by 1 ton will affect rice production by 0.971 tons in Gresik Regency indirectly. If the ZA fertilizer in Lamongan Regency increases by 1 ton, it will indirectly affect the increase in rice production by 0.007 tons in Gresik Regency. Last, the increase in a tractor by 1 unit in Lamongan Regency will indirectly increase rice production in Gresik Regency by 42.1 tons.

#### IV. CONCLUSION

Spatially Filtered Ridge Regression models are formed with queen contiguity weighting that can solve spatial autocorrelation and multicollinearity problems. The analysis has been carried out to determine the direct and indirect effects. The variables of the rice field area, urea fertilizer, Phonska fertilizer, SP-36 fertilizer, and tractor have a significant impact on rice production. Those variables have a significant effect directly and indirectly. However, ZA fertilizer has no significant effect on rice production. Then, a large comparison of direct and indirect effects for each predictor variable is generated, generally direct effects are greater than indirect effects.

The research limitation is that the used predictor variables are only six variables. Hence, for future research, it is suggested that the selection of variables are more various from many aspects, such as rice seeds, harvest area, and soil nutrients. Moreover, for weighting matrix based on the concept of the area, the future research can use another weighting matrix approach with the concept of distance.

#### REFERENCES

- [1] A. Rohman and A. D. Maharani, "Proyeksi kebutuhan konsumsi pangan beras di Daerah Istimewa Yogyakarta," *Caraka Tani: Journal of Sustainable Agriculture*, vol. 32, no. 1, pp. 29–34, 2017.
- [2] W. R. Tobler, "A computer movie simulating urban growth in the Detroit region," *Economic Geography*, vol. 46, no. Sup1, pp. 234–240, 1970.
- [3] Mahananto, S. Sutrisno, and C. F. Ananda, "Faktor-faktor yang mempengaruhi produksi padi studi kasus di Kecamatan Nogosari, Boyolali, Jawa Tengah," *WACANA, Jurnal Sosial dan Humaniora*, vol. 12, no. 1, pp. 179–191, 2009.
- [4] D. N. Gujarati and P. D. C., *Basic econometrics*. McGraw-Hill Irwin, 2009.
- [5] L. Anselin, *Spatial econometrics: Methods and models*. Springer Science & Business Media, 2013, vol. 4.
- [6] C. Fan, S. J. Rey, and S. W. Myint, "Spatially Filtered Ridge Regression (SFRR): A regression framework to understanding impacts of land cover patterns on urban climate," *Transactions in GIS*, vol. 21, no. 5, pp. 862–879, 2017.
- [7] A. D. P. Sari and W. S. Winahju, "Pemodelan faktor-faktor yang memengaruhi produksi padi di Jawa Timur," *Jurnal Sains dan Seni ITS*, vol. 5, no. 2, pp. 414–419, 2016.
- [8] F. Arnanda and A. Karim, "Pemodelan produksi padi di Provinsi Jawa Tengah dengan pendekatan spatial econometrics," *Jurnal Statistika Universitas Muhammadiyah Semarang*, vol. 4, no. 2, pp. 20–27, 2016.
- [9] R. Q. Muslim, B. Barus, and K. Munibah, "Analisis spasial indeks pertanaman dan produktivitas padi sawah di Dusun 1 Desa Purwasari, Kabupaten Bogor," Skripsi, Departemen Ilmu Tanah dan Sumber daya Lahan, Fakultas Pertanian, Bogor Agricultural University (IPB), 2017.
- [10] J. P. LeSage and R. K. Pace, *An introduction to spatial econometrics*. Boca Raton: Chapman & Hall/CRC, 2009.
- [11] G. Arbia, *Spatial econometrics: Statistical foundations and applications to regional convergence*. Germany: Springer Science & Business Media, 2006.
- [12] R. Fitriani and A. Efendi, *Ekonometrika spasial terapan dengan R*. Malang: UB Press, 2019.
- [13] N. U. Putri, Maiyastri, and H. Yozza, "Permasalahan autokorelasi pada analisis regresi linier sederhana," *Jurnal Matematika UNAND*, vol. 2, no. 2, pp. 26–34, 2013.
- [14] H. Pramoedyo, *Statistika inferensia terapan*. Malang: Danar Wijaya, 2013.
- [15] M. H. Kutner, C. J. Nachtsheim, J. Neter, and W. Li, *Applied linear statistical models*. Irwin, 2004.
- [16] Kementerian Pertanian Republik Indonesia. (2018) Tabel 2.1.3. Produksi padi menurut provinsi, 2014-2018. [Online]. Available: [https://www.pertanian.go.id/Data5tahun/TPATAP-2017\(pdf\)/20-ProdPadi.pdf](https://www.pertanian.go.id/Data5tahun/TPATAP-2017(pdf)/20-ProdPadi.pdf)