Parameter: Journal of Statistics
Volume 1 Nomor 1 Tahun 2021

THE INFLUENCE OF CLIMATE FACTORS ON COCOA PRODUCTIVITY IN SULAWESI, 2019

Marta Sundari[1], Pardomuan Robinson Sihombing[2*]
[1]Indraprasta University, Nangka Raya Street Number 58, South Jakarta
[2]BPS-Statistics Indonesia, dr. Sutomo Street No 6-8, Central Jakarta

***e-mail**: *robinson@bps.go.id*

ABSTRACT

Cocoa is one of the plantation commodities that has an important role in Indonesia's economic activity and is one of Indonesia's export commodities which is quite important as a source of foreign exchange and oil and gas. Sulawesi Island is one of the cocoa-producing islands in Indonesia. This study aims to determine a spatial regression model between the average cocoa productivity per month with the average drinking temperature per month, the average monthly rainfall and the average length of sunshine per month and the climatic factors that affect cocoa productivity in Sulawesi. The best model estimation uses the AIC value; the best model has the smallest AIC value. In this study, the SARMA spatial regression model is the best model with the specified criteria. The AIC value of the SARMA model is -61.145, the smallest of the other models.

Keywords: *Cocoa, SARMA, Spatial, Sulawesi*

INTRODUCTION

The agricultural sector has an important role in Indonesia's economic activity as seen from the contribution of the agricultural industry to Gross Domestic Product (GDP) of 12.72% in 2019 or is in third place after the Manufacturing and Wholesale and Retail Trade (19.70%), Car and Motorcycle Repair sector (13.01%). When an economic crisis occurs, the agricultural industry is a sector that is strong enough to face economic shocks and can be relied on in the recovery of the national economy (Badan Pusat Statistik, 2020). Part from that, being an agricultural country, the agricultural sector is the main sector in Indonesia.

The plantation sub-sector has a contribution of 3.27% to 2019 GDP or ranks first in the Agriculture, Animal Husbandry, Hunting and Agricultural Services sectors. The plantation sub-sector is a raw materials provider for the industrial sector, absorbing labor and earning foreign exchange. Cocoa is one of the plantation commodities that has a vital role in economic activity in Indonesia and is one of Indonesia's export commodities which is quite important as a source of foreign exchange and oil and gas. Indonesia is the third-largest cocoa producer and exporter in the world after Ghana and Ivory Coast. In Indonesia, in 2018, 98.33% of cocoa plantation area was cultivated by smallholder plantations, 0.89% by large private farms and 0.76% by large state plantations.

Sulawesi Island is one of the cocoa-producing islands in Indonesia. The total area of cocoa plantations cultivated by smallholder plantations is 1,351,988 ha or 85.35% of the total area of cocoa plantations produced by smallholder plantations in Indonesia. In 2019, smallholder plantation cocoa production in Sulawesi reached 751,685 or 62.01% of national smallholder cocoa production. This illustrates that Sulawesi Island has an important role in national cocoa production.

This study aims to determine a spatial regression model between the average cocoa productivity per month with the average drinking temperature per month, the average monthly rainfall and the average length of sunshine per month and the climatic factors that affect cocoa productivity in Sulawesi. The methods being compared are the General Spatial Model, the Spatial Autoregressive Model and the Spatial Error Model. Determination of the best model by looking at the smallest AIC criteria.

MATERIALS AND METHODS

1. Spatial Correlation and Spatial Diversity

Spatial correlation is measured before measuring spatial effects. Spatial autocorrelation is a measure of the similarity of objects in space. The spatial autocorrelation approach can use Moran's Index statistic (Fischer & Wang, 2011) . The statistical hypothesis is as follows:

H_0: I = 0 (no spatial correlation between locations)

H_1: I ≠ 0 (there is a spatial correlation between locations)

The Moran Index equation is as follows:

$$I = \frac{n\sum_{i=1}^{n}\sum_{j=1}^{n} W_{ij}(y_i-\bar{y})(y_j-\bar{y})}{\sum_{i=1}^{n}\sum_{j=1}^{n} W_{ij}\sum_{i=1}^{n}(y_i-\bar{y})^2} \quad (1)$$

where W is the weighted spatial matrix resulting from the standardization of rows, y is the vector of the dependent variable, and n is the number of observations. Moran's Index test statistics can be stated as follows:

$$Z_{hitung} = \frac{I-E(I)}{\sigma(I)} \quad (2)$$

$$E(I) = -\frac{I}{n-1} \quad (3)$$

where I is the Moran's Index, E (I) is the expected value of the Moran's Index, σ (I) is the standard deviation of the Moran's Index and n is the number of observations.

$$\text{Moran's Index test criteria } = \begin{cases} |\mathrm{Z_{Hitung}}| \le \mathrm{Z}\alpha_{/2}, \text{fail to reject } \mathrm{H_0} \\ |\mathrm{Z_{Hitung}}| > \mathrm{Z}\alpha_{/2}, \text{reject } \mathrm{H_0} \end{cases} \quad (4)$$

The Moran's Index value is in the interval −1 ≤ I ≤ 1. If I> E (I), then the neighbouring locations have a positive relationship, namely that neighbours' values are similar. If I <E (I), then the adjacent areas have a negative relationship, that is, the values for the neighbours are not identical to each other (Anselin, 1988). The spatial variation is caused by differences in the characteristics between the points of observation. Detection of spatial diversity can use the Breusch-Pagan (BP) test. God-Frey (1978) and Breusch and Pagan (1979) in Arbia (2006) says that the homogeneity of variety is fulfilled if the equation is as follows:

$$E(\varepsilon_i^2|\mathrm{X}) = \alpha_1 x_{1i} + \alpha_2 x_{2i} + \cdots + \alpha_p x_{pi} \quad (5)$$

with a value of α_j zero (j = 2, 3, ..., p), x_{1i} is the regression constant which is always one and x_{2i}, ..., x_{pi} is the 2nd to the p independent variable. The homogeneity hypothesis of variety is as follows:

H_0: $\alpha_2 = \alpha_3 = \cdots = \alpha_p = 0$

H_1: there is at least one $\alpha_j \neq 0$

Anselin (1988) states that the Breusch-Pagan (BP) test statistics are as follows:

$$BP = \frac{1}{2}\left(\sum_{i=1}^{n} x_i f_i\right)'\left(\sum_{i=1}^{n} x_i x_i'\right)\left(\sum_{i=1}^{n} x_i f_i\right) \quad (6)$$

with $f_i = \left(\frac{\hat{\varepsilon}_i}{\hat{\sigma}} - 1\right), \hat{\varepsilon}_i = \left(y_i - \hat{\beta}' x_i\right)$ and $\hat{\sigma}^2 = \sum_{i=1}^{n} \hat{\varepsilon}_i^2$

$$\text{BP test criteria} = \begin{cases} BP \le \chi^2_{(p-1)}, \text{fail to reject } \mathrm{H_0} \\ BP > \chi^2_{(p-1)}, \text{reject } \mathrm{H_0} \end{cases} \quad (7)$$

where p is the number of regression parameters.

2. Spatial Dependence

The Lagrange multiplier (LM) test is used to test the spatial dependence of the SAR model. The SAR model, Lagrange multiplier test, is calculated using the following formula:

$$LM_\rho = \left(\frac{\varepsilon' W y}{\varepsilon' \varepsilon n^{-1}}\right)^2 \frac{1}{H} \quad (8)$$

with with $H = \{(WX\beta)'[1 - X(X'X)^{-1}X'](WX\beta)\sigma^{-2}\} + tr(W'W + W^2)$ $\boldsymbol{\varepsilon}$ is the residual vector of the classical-sized regression model (nx 1), β obtained from the classical regression model, σ^2 is the remaining middle square of the classical regression model, and tr represents the core operation of the matrix, namely the addition of the diagonal elements of a matrix (Fischer & Wang, 2011).

$$\text{Test criteria} = LM_\rho \begin{cases} LM_\rho \le \chi^2_{(q)}, \text{fail to reject } \mathrm{H_0} \\ LM_\rho > \chi^2_{(q)}, \text{ reject } \mathrm{H_0} \end{cases} \quad (9)$$

and q is the number of spatial parameters. If H_0 rejected then the appropriate regression model is the SAR model.

The spatial dependence of the SEM model can be detected by the lagrange multiplier (LM) test. The SEM model lagrange multiplier test is calculated using the following formula:

$$LM_\lambda = \left(\frac{\varepsilon\prime Wy}{\varepsilon'\varepsilon n^{-1}}\right)^2 \frac{1}{tr(W'W+W^2)} \quad (10)$$

with **ε** is the residual vector of the classical-sized regression model (*nx 1*), and tr states the core operation of the matrix, namely the sum of the diagonal elements of a matrix (Fischer & Wang 2011).

$$\text{Test criteria} = LM_\lambda \begin{cases} LM_\lambda \leq \chi^2_{(q)}, \text{tidak tolak } H_0 \\ LM_\lambda > \chi^2_{(q)}, \text{tolak } H_0 \end{cases} \quad (11)$$

and q is the number of spatial parameters. If H_0 is rejected, so the appropriate regression model is the SEM model.

3. General Spatial Model (GSM)

General spatial model is a linear regression model in which the independent variable x-i is correlated with the dependent variable y to the j as well as the i-dependent dependent error with the j dependent error ($\rho \neq 0$ and $\lambda \neq 0$), the general form is as follows (Anselin, 1988):

$$\begin{aligned} y &= \rho Wy + X\beta + u \\ u &= \lambda Wu + \varepsilon \\ \varepsilon &\sim N(0, \sigma^2 I) \end{aligned} \quad (12)$$

with y is a vector of variable size dependent (*n x 1*), X is a matrix of independent variables of size (n x (p + 1)), β regression coefficient vector sized ((p + 1) x 1), ρ is the spatial lag autoregression coefficient, λ is the spatial error autoregression coefficient which is valued $|\lambda| \leq 1$, u is a random error vector that is assumed to contain autocorrelation of measure (*n* x 1), W is the spatial weighted matrix of size (n x n), and n is the number of objects of observation.

The assumption test in spatial regression includes the assumption of homogeneity of variance and normalcy of errors. Anselin (1988) states that to estimate the regression parameters of a spatial general model obtained by the maximum likelihood estimation method, expressed in the form:

$$\beta = (X'(I-\lambda W)'(I-\lambda W)X)^{-1}X'(I-\lambda W)'(I-\lambda W)(I-\rho W)y \quad (13)$$

the variance estimator is as follows:

$$\sigma^2 = \frac{((I-\lambda W)(I-\rho W)y-(I-\lambda W)X\beta)\prime((I-\lambda W)(I-\rho W)y-(I-\lambda W)X\beta)}{n} \quad (14)$$

4. Spatial Autoregressive Model (SAR)

Spatial autoregressive model (SAR) is a spatial model with an area approach that considers the effect of spatial lag only on the response variable (Anselin, 1988). The SAR model is obtained from the equation $y = \rho Wy + X\beta + u$ if the dependent variables are spatially correlated ($\rho \neq 0$ and $\lambda = 0$), then the equation will be as follows:

$$\begin{aligned} y &= \rho Wy + X\beta + \varepsilon \\ \varepsilon &\sim N(0, \sigma^2 I) \end{aligned} \quad (15)$$

with ε_i is the error at location i which is assumed to spread normally with zero mean and constant range σ^2. Parameter estimator β the SAR model is estimated using the maximum likelihood method, expressed in the form:

$$\beta = (X'X)^{-1}X'y - (X'X)^{-1}X'\rho Wy \quad (16)$$

Estimation of ρ is carried out through a numerical iteration to obtain an estimator of ρ which maximizes the log likelihood function. The rest of the SAR model is calculated by the following formula:

$$\begin{aligned} y &= \rho Wy + X\beta + \varepsilon \\ (I - \rho W)y &= X\beta + \varepsilon \\ \varepsilon &= (I - \rho W)y - X\beta \end{aligned} \quad (17)$$

5. Spatial Error Model (SEM)

Spatial error model is a linear regression model in which the error variable has a spatial correlation. This is because there are explanatory variables that are not involved in the linear regression model so that they will be counted as errors and these variables are spatially correlated with errors at other locations. The spatial error model is obtained from the equation $y = \rho Wy + X\beta + u$ if $\rho = 0$ and $\lambda \neq 0$, then the equation will be as follows:

$$y = X\beta + u \quad (18)$$
$$u = \lambda Wu + \varepsilon$$
$$\varepsilon \sim N(0, \sigma^2 I)$$

with ε_i is the error at location i which is assumed to spread normally with zero mean and constant range σ^2. Parameter estimator β SEM models are estimated using the maximum likelihood method, expressed in the form:

$$\beta = [(X - \lambda WX)'(X - \lambda WX)]^{-1}(X - \lambda WX)'(y - \lambda Wy) \quad (19)$$

Estimator for $\sigma^2$is:

$$\sigma^2 = \frac{[(I-\lambda W)(y-X\beta)\prime(I-\lambda W)(y-X\beta)]}{n} \quad (20)$$

Estimation of λ is carried out through a numerical iteration to obtain a λ estimator that maximizes the log likelihood function. The rest of the SEM model is calculated by the following formula:

$$u - \lambda Wu = \varepsilon \quad (21)$$
$$(I - \lambda W)u = \varepsilon$$
$$\varepsilon = (I - \lambda W)(y - X\beta)$$

6. Weighted Matrix

The relationship between regions can be presented in the form of neighbors and weighted by distance (Fischer & Wang, 2011). The weighting matrix is a matrix that states the relationship between regions in the form of a symmetrical matrix and the main diagonal always has a zero value. The i-th row of the weighted matrix shows the relationship of the i-region to other regions. The closer the distance between the locations, the greater the weight given. This is because locations that are close together generally have similar characteristics, in contrast to locations that are far away, generally the features between these locations will be more varied, so the weight given will be smaller. The spatial weighting matrix can be determined based on two categories, namely based on contiguity and distance.

a. Neighbors Weighted

The Queen weighted matrix is defined as $W_{ij} = 1$ for the area which is adjacent (common side) or the vertex (common vertex) meets the area of concern while $W_{ij} = 0$ for other regions. The illustration below shows the formation of Queen's spatial weighting matrix. In Table 1 is the queen's step, which is to determine the area directly adjacent to Figure 1. Based on Table 1, the weighting matrix (W) is obtained by standardizing the matrix arrangement, namely the number of rows is equal to one, so that a weighting matrix is obtained as in formula (22).

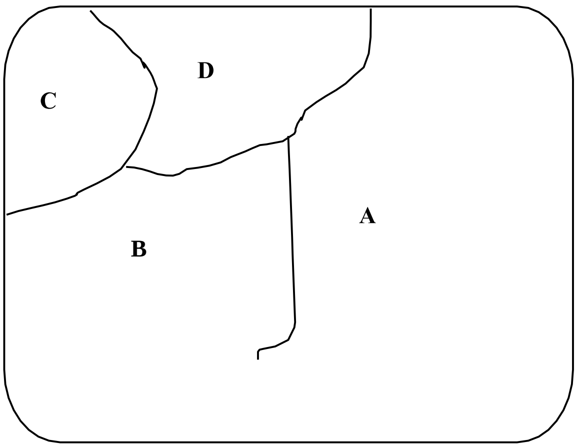

Figure 1. Illustration of Spatial Weighting with Nearest Neighbors

Table 1. Queen's Step Weighting Based on Nearest Neighbor

	A	B	C	D	
A	0	1	0	1	2
B	1	0	1	1	3
C	0	1	0	1	2
D	1	1	1	0	3
	2	3	2	3	

$$W_{queen} = \begin{bmatrix} 0 & 1/2 & 0 & 1/2 \\ 1/3 & 0 & 1/3 & 1/3 \\ 0 & 1/2 & 0 & 1/2 \\ 1/3 & 1/3 & 1/3 & 0 \end{bmatrix} \tag{22}$$

b. Weighted by distance

Determination of the elements of the spatial weighting matrix can be represented in the form of a distance function. In principle, the weight of the distance between a location and its surroundings is determined by the distance between the two areas. Several types of determining the weighting matrix based on the distance: 1) K- the closest neighbor, in this method the researcher can determine the j-location itself, as much as the k-location which is the closest location around the i-location, 2) The weighting is based on the distance formulated as following:

$$w_{ij} = \begin{cases} 1, \ jika \ \ d_{ij} < d(d > 0) \\ 0, \ jika \ \ d_{ij} \geq d(d > 0) \end{cases} \tag{23}$$

where d is the limit of the distance specified and is the distance between location i and location j. The distance between locations is calculated by the following formula:d_{ij}

$$d_{ij} = \sqrt{\sum_{i=1}^{p} \sum_{j=1}^{p} (X_i - X_j)^2} \ \ , i \neq j \tag{24}$$

7. Type of Data

The data used in this study is secondary data on cocoa production in smallholder plantations in 2019.

Table 2. Research Variables

Variable	Variable Name
Y	Average Productivity (tonnes / ha / month)
X_1	Average Minimum Temperature (0C / month)
X_2	Average Rainfall (mm / month)
X_3	Average Duration of Sunlight (hour / day / month)

8. Data source

The data used is sourced from the 2019 Indonesian Cocoa Statistics (Badan Pusat Statistik, 2020), while data on minimum temperature, rainfall and duration of sun exposure were downloaded from the bmkg.go.id site (20BM). Cocoa production data used are production data from the Provinces of North Sulawesi, Gorontalo, Central Sulawesi, West Sulawesi, South Sulawesi and Southeast Sulawesi.

9. Stages of the Spatial Analysis Procedure

The stages of data analysis carried out are:

1) Data exploration
2) Spatial Correlation Test
3) Dependency test with lagrange multiplier test
4) Estimating parameters for the spatial regression model equations for SAR, SEM and GSM using the maximum likelihood estimation method

5) To test the spatial regression model assumptions which include normality and homogeneity of variance with the Breusch-Pagan test, if it is not normal and the variance is not homogeneous then the data will be transformed then repeat steps 3 and 4
6) The selection of the best model uses the smallest Akaike Information Criterion (AIC) value criteria

RESULTS AND DISCUSSION

1. Data Exploration

The average productivity of smallholder cocoa plantations in the provinces on Sulawesi Island is 37.26 ton/ha/month, Gorontalo Province has the lowest average cocoa productivity of 25.62 ton/ha/month and the highest productivity is produced by the Southern Province as much as 45.61 ton/ha/month. North Sulawesi Province is a province with the lowest average minimum temperature of 22.62 ^{0}C / month and the highest average minimum temperature is in West Sulawesi Province at 24.65 ^{0}C / month.

The average rainfall in Sulawesi Island is 116.30 mm/month. The lowest average rainfall is in Central Sulawesi Province as much as 53.10 mm/month and the highest is in South Sulawesi Province at 277 mm/month. The average length of sun exposure is 6.16 hours/day/month with the lowest average length of sun exposure in Sulawesi Selatan Province at 5.16 hours/day/month, and the highest is in West Sulawesi Province at 7.07 hours/day/month:

Table 3. Distribution of Concentration Measures

Variable	Average	Standard Deviation	Minimum Value	Maximum Value
Y	37.26	8.14	25.62	45.61
X_1	23.56	0.72	22.79	24.65
X_2	116.3	82.80	53.10	277.00
X_3	6.16	0.75	5.16	7.07

The linear relationship between the response and explanatory variables can be seen using the Pearson correlation value in Table 4. The correlation value between the Y variable and the X1 and X2 variables is positive, which means that any increase in the explanatory variable will cause an increase the response variable or the relationship is unidirectional. The correlation value between the Y variable and the X3 variable is negative, which means that any increase in the explanatory variable will cause a decrease the response variable or the relationship is reversed.

Table 4. Pearson Correlation Test

	Y	X_1	X_2
X_1	0.546		
X_2	0.580	-0.006	
X_3	-0,330	0.561	-0,650

Multicollinearity examination uses the VIF (Variance Inflation Factor) value of each variable. Multicollinearity occurs when the VIF value is greater than 10. The VIF value of X_1 is 2.166327, X_2 is 2.571064 and X_3 is 3.754199, which illustrates no multicollinearity between variables.

2. Cocoa Production Distribution Map

The first highest average productivity of smallholder cocoa is in South Sulawesi Province. The second highest is in Southeast Sulawesi Province where the annual rainfall is in the ideal rainfall interval for cocoa. In general, the minimum average temperature in all provinces in Sulawesi Island is below the optimum temperature range, as shown in Figure 2.

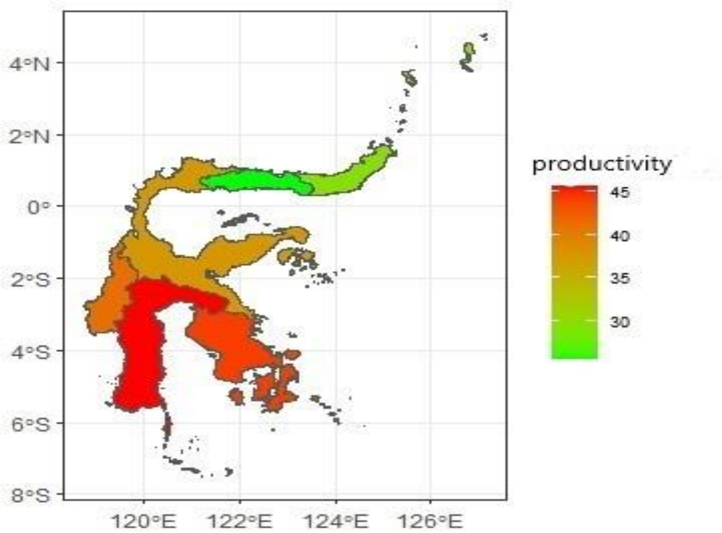

Figure 2. Map of Cocoa Productivity Distribution for Smallholders in Sulawesi Island

3. Moran's Index Analysis

Testing the spatial correlation using neighbor weighting results in a p-value of 0.007723 which is smaller than the significant level of 0.05, so it can be concluded that there is a spatial correlation in smallholder cocoa plantations' productivity data in Sulawesi Island. The results of this correlation test are in accordance with the local moran plots where the provinces of North Sulawesi and Gorontalo which have low productivity are in quadrant III, while other provinces that have higher productivity are in the same quadrant as South Sulawesi Province, namely quadrant I (Figure 3).

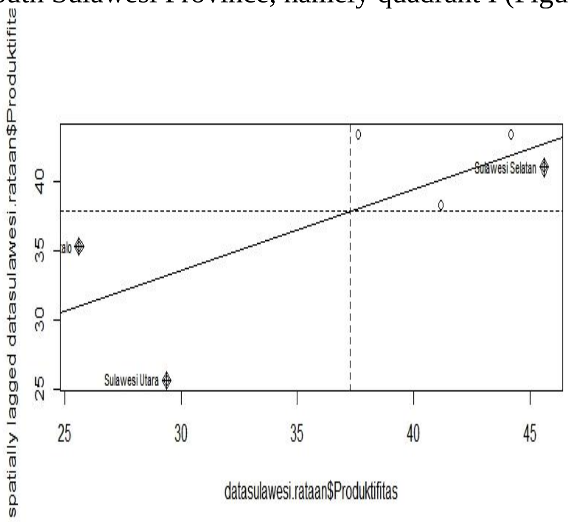

Figure 3. Moran's Local Plot

From the results of testing the error variety of the classical regression model, the BP value of 5.7724 and the p-value of 0.1232 is smaller than the value at the 0.05 level of 5.991, so it can be concluded that there is no spatial variation in the average productivity value. Smallholder cocoa plantations in 2019 so that the SAR, SEM or GSM models are considered sufficient to describe the model $\chi^2_{table} = 5.991$.

4. Spatial Dependency Test

The spatial lag dependency test results and spatial error with the neighbor weight. The spatial lag LM value is 0.036 and the p-value of 0.85 are greater than the significant level of 0.05, so it can be concluded that there is no spatial lag effect. The results of the spatial error dependence test with the neighbor weighting result in a spatial error LM value of 4.25 and a p-value of 0.04 which is smaller than the significant level of 0.05 so that it can be concluded that there is an effect of spatial error, so it is necessary to proceed to make a spatial error model (SEM). The results of the lag and spatial error dependency test with the neighbor weighting result in the LM lag and spatial error value of 4.78 and a p-value of 0.09 which is smaller than the real level of 0.10 so that it can be concluded that there is an effect of spatial lag and error, so it is necessary to proceed to model mixed or spatial autoregressive moving average (SARMA) model.

Table 5. Spatial Dependency Test for SAR, SEM and GSM Models

Spatial Dependency Test	**LM statistics**	***p-value***
Lagrange Multiplier (lag)	0.03635	0.8487
Lagrange Multiplier (error)	4,25298	0.0391
Lagrange Multiplier (SARMA)	4,77543	0.0918

5. SAR model

From Table 6, it is found that at the 5% real level that affects the average productivity of cocoa in Sulawesi Island is the intercept, ρ, X_1 (Average Minimum Temperature) and X_3 (Average Duration of Sun lighting). The models built are:

$$\hat{y} = -1.57x10^6 +0.42421\rho +1.01x10^5X_1-7.48x10^1X_2-9.57x10^4X_3 \quad (25)$$

Coefficient ρ it is 0.42. It means that the i-th province is correlated by 0.42 with the surrounding provinces, for example, Southeast Sulawesi Province is correlated 0.42 with the province of South Sulawesi. The negative sign on the variable coefficients X_2 and X_3 mean that the increase in average rainfall and the length of sun exposure causes the productivity to decrease.

Table 6. Estimation and Testing of SAR Model Parameters

Variable	Coefficient	*Z value*	*P-value*
Intercept	$-1.57x10^6$	-9.4545	$<2x10^{-16}$
X_1	$1.01x10^5$	10,7044	$<2x10^{-16}$
X_2	$-7.48x10^1$	-0.9277	0.3536
X_3	$-9.57x10^4$	-8.8538	$<2x10^{-16}$
ρ	0.42421	11.13	0.00084921

6. SEM Model

From Table 7, it is found that at the fundamental level of 10% that affects the average productivity of cocoa in Sulawesi Island is the intercept, λ, X_1 (Average Minimum Temperature) and X_3 (Average Duration of Sun lighting). The models built are:

$$\hat{y} = -1.22x10^2 + 06 +0.88184\rho + 8.59X_1 - 3.12x10^4X_2 - 7.29X_3 \quad (26)$$

Coefficient λ amounting to 0.88 and real means that the i-th province is correlated by 0.88 with the surrounding provinces, for example, Southeast Sulawesi Province correlates 0.88with South Sulawesi Province. The negative sign on the variable coefficients X_2 and X_3 mean that the increase in the average rainfall and the length of sun exposure causes the productivity to decrease.

Table 7. Estimation and Testing of SEM Model Parameters

Variable	Coefficient	*Z value*	*P-value*
Intercept	$-1.22x10^2$	-5.7412	0.0000
X_1	8.59	7.9540	0.0000
X_2	$-3.12x10^{-4}$	-0.0420	0.9665
X_3	-7.29	-6.8883	0.0000
λ	0.88184	3.6688	0.0555

7. GSM model

From Table 9, it is found that at the 5% real level that affects the average productivity of cocoa in Sulawesi Island is the intercept, ρ, λ, X_1 (Average Minimum Temperature), X_2 (Average Rainfall) and X_3 (Average Duration of Sunlighting). The model built is:

$$\hat{y} =9.97x10^1 + 0.84755\rho +0.88184\rho + 5.41X_1+1.53x10^{-2}X_2-3.99X_3 \quad (27)$$

Coefficient ρ amounting to 0.85 and real means that the i-th province is correlated by 0.85 with the surrounding provinces, for example, Southeast Sulawesi Province correlates 0.88with South Sulawesi Province. The negative sign on the variable coefficient X_3 means that the increase in the length of sun exposure causes cocoa productivity to decrease.

Table 8. Estimation and Testing of GSM Model Parameters

Variable	Coefficient	*Z value*	*P-value*
Intercept	$-9.97x10^1$	-33,552	0.0000
X_1	5.41	21,092	0.0000
X_2	$1.53x10^{-2}$	6,310	0.0000
X_3	-3.99	10,702	0.0000
ρ	0.84755	102.95	$<2.22x10^{-16}$
λ	-1.1973	102.95	$<2.22x10^{-16}$

8. Best Model Selection

R^2 in the spatial regression model is pseudo R^2. The spatial regression model uses the maximum likelihood method to be compared with R^2 from the classical regression model using the OLS method. The best model estimation uses the AIC value; the best model has the smallest AIC value. Based on Table 10, it is found that the SARMA spatial regression is the best model compared to the other three models because it has the smallest AIC value, which is equal to -61.145. The assumption test is carried out for all models, at the 5% real level, it is found that the remainder produced by the SEM model is not independent. In contrast, the rest's normality test cannot be carried out because it requires several at least seven observations. At the 5% real level, all the models used produce a homogeneous residue.

Table 9. Selection of the Best Model

	Classic Regression	SAR	SEM	SARMA / GSM
AIC	37.80	28.67	36.14	-61,145
Rho		0.42	0.88	0.85
Lambda				-1.20
Normality	Normal	The sample unit is less than 7		
Homogeneous	Homogeneous	Homogeneous	Homogeneous	Homogeneous
Freedom	Free	Free	Not free	Free

The interpretation of the results of the SARMA model is as follows:

The average minimum temperature has a significant positive effect on the average cocoa production with a coefficient of 5.41, meaning that an increase of 1°C / month will increase the average cocoa productivity by 5.41 ton/ha/month assuming the other variables are constant. This is in line with Rubiyo & Siswanto's research (2012), which states that cocoa plants require high temperatures and the optimum temperature for cocoa growth is 30-32 ^{0}C). This temperature effect can influence the growth of cocoa because it is closely related to the availability of water, sunlight and humidity (Safuan, Kandari, & Natsir, 2013).

Average rainfall has a significant positive effect on the average cocoa production with a coefficient of 0.015, meaning that an increase of 1 mm/month of rainfall will increase the average cocoa productivity by 0.015ton / ha/month assuming the other variables are constant. The original habitat for cacao plants is wet tropical forest. It grows under the shade of tall trees, high rainfall, relatively the same year-round temperature and relatively constant high humidity. In useful cultivation techniques, some of the original cocoa habitat characteristics are still preserved by providing shade (Pusat Penelitian dan Perkebunan., 2010). Wibawa and Baon (2008) obtained the same research for improvement rainfall will increase cocoa productivity. The ideal rainfall for cocoa growth is 1,500-2,500 mm / year with an even distribution of rain throughout the year.

The average length of sun exposure has a significant negative effect on the average cocoa production with a coefficient of -3.99, meaning that an increase in 1 hour/day/month of exposure time will reduce the average cocoa productivity by 3.99 ton/ha/month assuming the other variables are constant. This is due to the cacao plant that requires optimal radiation because excess radiation will reduce its productivity.

CONCLUSION

The best model estimation uses the AIC value; the best model has the smallest AIC value. From the three models used, it was found that the SARMA spatial regression was the best model compared to the other two models because it had the smallest AIC value, which was -61.145. The SARMA models built are $\hat{y}$ = -9.97+ 0.84755ρ +0.88184ρ + 5.41X_1+0.0153X_2-3.99$X_{3.}$

REFERENCES

Anselin, L. (1988). *Spatial Econometrics: Methods and Models.* Netherlands: Kluwer Academic Publishers.

Arbia, G. (2006). *Spatial Econometrics: Statistical foundation and Application to Regional Convergence*. Berlin (GB): Springer.

Badan Pusat Statistik. (2020). *Statistik Kakao Indonesia 2019.* Jakarta: Badan Pusat Statistik.

BMKG. Retrieved Januari 2, 2020, from https://dataonline.bmkg.go.id/home

Fischer, M., & Wang, J. (2011). *Spatial Data Analysis: Models, Methods and Techniques.* New York: Springer Heidelberg Dordrecht.

Pusat Penelitian dan Perkebunan. (2010). *Budidaya dan pasca panen kakao.* Jakarta: Eska Media.

Rubiyo, & Siswanto. (2012). Peningkatan produksi dan pengembangan kakao (Theobroma cacao L.) di Indonesia. *Buletin Riset Tanaman Rempah dan Aneka Tanaman Industri, 3*(1), Rubiyo, & Siswanto. (2012). .

Safuan, L., Kandari, A. M., & Natsir, M. (2013). Evaluasi kesesuaian lahan tanaman kakao (Theobroma cacao L.) berdasarkan aplikasi analisis data iklim menggunakan aplikasi sistem informasi geografi. *Jurnal Agroteknos, 3*(2), 80-85.

Wibawa, A., & Baon, J. (2008). *Kesesuaian lahan. Panduan lengkap kakao: Manajemen agribisnis dari hulu hingga hilir.* Jakarta: Penebar Swadaya.