Study of the effect of climate variation on irrigated and rainfed rice productivity based on aquacrop crop modelling simulation (Case Study of Java Island)

K Aprilina¹,²*, A Susandi³, A Sopaheluwakan⁴ and H Harsa²

¹Earth Science Graduate Program, Bandung Institute of Technology, Jalan Ganesha No 10, Bandung, Indonesia
²Research and Development Centre, Indonesian Agency for Meteorology Climatology and Geophysics, Jalan Angkasa I No 2 Kemayoran, Jakarta, Indonesia
³Atmospheric Research Group, Bandung Institute of Technology, Jalan Ganesha No 10, Bandung, Indonesia
⁴Center for Applied Climate Services, Indonesian Agency for Meteorology Climatology and Geophysics, Jalan Angkasa I No 2 Kemayoran, Jakarta, Indonesia

*Corresponding E-mail: kharisma.aprilina@bmkg.go.id

Abstract. Aquacrop is a free-licensed Food and Agricultural Organization’s crop modelling tool that requires minimum inputs of climate variables namely rainfall, maximum temperature, minimum temperature variables and geographic information of the area to be simulated (longitude, latitude, altitude). This study aims to measure the difference in irrigated and rainfed rice productivity from the Aquacrop crop modelling simulation to the influence of climate pattern variations in Java Island, Indonesia. The k-means clustering method applied to the rainfall, maximum, and minimum temperature variables from the bias-corrected MERRA2 data resulted in two climate regions. The principal component analysis result showed that the maximum and minimum temperature variables are the variables that most contribute to the determination of the clustering area using the k-means method compared to the rainfall variable. This study has calculated the probability of the irrigated and rainfed rice productivity resulting from the Aquacrop simulation in those climate regions during La Nina [El Nino] years that will be higher [smaller] than the mean value of rice productivity during neutral years. However, the validation between the actual irrigated and rainfed rice productivity with the Aquacrop simulation results from 2001-2014 showed low correlation values that vary between negative and positive values in all climate regions. Meanwhile, the validation on the El Nino composite years generally showed positive correlation values. In addition, the neutral and La Nina composite years resulted in varying correlation values between negative and positive correlation.

Keywords: Aquacrop, Variations of climate patterns, Rice productivity, Clustering

Track Name: Land, Water, Forest, and Food Security
1. Introduction
Java Island is the main rice-producing island in Indonesia [1], although it is not the largest island in the country. However, the population density and population on Java Island is also the highest in this country, where more than half of Indonesia's population lives on this island [2], therefore Indonesia's food security can be represented by the food security status of the people living on Java Island [3]. In order to maintain the stability of food security in Indonesia, various studies that link environmental factors to rice production and productivity have been carried out [4–9]. One of these studies produces a linear regression function that relates climate factors to rice production in Java [9]. However, this regression function only can be applied if there is an assumption of uniformity of climate patterns in all regions in Java, where it has been known that, in general, the climate pattern in Java is monsoon (annual) [10].

A recent study has shown that there is a phenomenon of variations in climate patterns in Java, such as semi-annual variations where some areas experience the peak of the rainy season other than the December–January–February (DJF) period, for example, in the March–April–May (MAM) period [11], this non-uniformity of climate patterns then can be referred as variations in climate patterns. Therefore, it can be assumed that the climate pattern in the Java Island region can also impact variations in the rice productivity in Java Island.

Generally, variations in climate patterns in Indonesia have been delineated using rainfall variables. However, the latest studies had carried out climate classification using rainfall and temperature variables [12], as well as temperature range (maximum temperature–minimum temperature) [13] or maximum temperature, minimum temperature, and relative humidity variables simultaneously [14,15]. Therefore, in this study, variations in climate patterns will be determined from clustering the variables of rainfall, maximum and minimum temperature, considering that these three variables are most often used in climate classification in previous studies.

Rice productivity as a response variable from the impact of variations in climate patterns in this study was generated from crop modelling simulations using FAO (Food and Agriculture Organization) Aquacrop. Aquacrop's crop modelling algorithm is based on water demand for plants during their growth and development period [16,17]. Aquacrop crop modelling is used in this study because it is easy to use, free license, and requires relatively few inputs of climate variables (rainfall, maximum temperature, minimum temperature) and geographic information of the area to be simulated (longitude, latitude, altitude) [18,19]. In addition, Aquacrop can simulate crop productivity using different irrigation systems, such as irrigated and rainfed systems [18]. So this study aims to determine the impact of variations in climate patterns in Java on rice crop productivity resulting from Aquacrop simulations under different irrigation systems.

2. Data and Method

2.1 Data
2.1.1 Observation Data
The data used in this study are daily rainfall data (precip), maximum temperature (tmax), and minimum temperature (tmin) from 23 stations of the Indonesian Agency for Meteorology, Climatology and Geophysics (BMKG) located on Java Island [20] for 31 years from years 1987-2017.

2.1.2 Supporting data
Supporting data used include:

a. MERRA2 data (The Modern-Era Retrospective analysis for Research and Applications, Version 2) with spatial resolution 0.5° x 0.625° which consists of daily rainfall, maximum and minimum temperature data for 31 years from 1987-2017 [21] which is cropped to Java Island between 9° S-5.5° S and 104° E-117° E.

b. Actual data on irrigated and rainfed rice productivity obtained from the BPS (Indonesia Central Bureau of Statistics) publication in Figures from the official website of the Provincial
BPS in Java, the available data are annual data for each Regency/City ranging from 13-14 years from 2001/2002 - 2014.

c. Southern Oscillation Indices (SOI) data downloaded from http://www.cgd.ucar.edu/cas/catalog/climind/soi.html, for a period of 31 years from 1987-2017.

2.2 Method
Bias correction is carried out using the Quantile Mapping method from the qmap package [22] in the R program. This correction method can capture changes in the average and variability of the data as well as adjust the statistical moment of the corrected data [23]. Clustering was performed simultaneously on the predictor variables (precip, tmax, and tmin) using the k-means method from the package stats [24] in the R program. The k-means method was used because this method resulted in a more stable cluster boundary [25]. Next, the determination of the number of k-groups was conducted using the Silhouette method [26]. The results of grouping with the k-means method, then further analyzed using the Principal Component Analysis (PCA) method. PCA method is proved as an advanced solution to indicate the number of group members in the k-means clustering method [26]. In addition, the PCA method can capture important information from predictor variables and determine the grouped area's characteristics [27]. PCA was conducted on a correlation matrix because the predictor variables, namely annual rainfall, maximum and minimum temperatures, had different units. Thus data standardization was necessary. PCA was performed using the prcomp function in the R programming stats package. PCA was performed to detect the stability of the grouping results by the k-means method [28]. The loadings PC value is used to determine the contribution of each variable to the grouped area.

Aquacrop crop modelling input uses predictor variable segmentation (daily precip, tmax, and tmin from bias-corrected reanalysis data) resulting from clustering. Aquacrop is run with the assumption that apart from predictor variables, other factors such as soil type, soil structure and plant varieties are homogeneous so that the simulation output reflects the influence of the predictor variables [29]. Aquacrop is run with two modes of irrigation systems, namely regular irrigation and rainfed systems. Planting dates are assumed to be the early month of January, May, and September each year from 1987 to 2017.

Table 1. The ENSO composite years correspond to the active mode used in this study.

|                  | El Nino |       |   | La Nina |       |   |
|------------------|---------|-------|---|---------|-------|---|
|                  | weak    | moderate | strong | weak    | moderate | strong |
| 2004             | 1987    | 1991  | 1996 | 1999    | 1988    |
| 2006             | 1994    | 1997  | 2000 | 2005    | 1998    |
| 2014             | 2002    | 2015  | 2008 | 2011    | 2007    |
|                  | 2009    |       | 2016 |        | 2010    |
|                  |         |       | 2017 |        |         |

Then, the probabilities of irrigated and rainfed rice productivity in each group are calculated by first determining the distribution pattern of rice productivity data in the composite years experiencing climate variability mode. Determination of the composite year ENSO (El Nino Southern Oscillation) using a standardized SOI index (Table 1). While the determination of weak, medium, and strong ENSO modes refers to [30], namely weak La Nina SOI values range from 0.3 to 0.65, moderate La Nina SOI values range from 0.65 to 0.95, strong La Nina SOI values are more than 0.95, weak El Nino SOI values between -0.6 to -1.2, moderate El Nino SOI values ranged from -1.2 to -1.8, and strong El Nino SOI values less than -1.8.
3. Result and Discussion

3.1 The result of bias correction with the quantile mapping method

The MERRA2 data before being bias-corrected showed a fairly good agreement with the observation data, but for the months of January-May, the MERRA2 data was not able to reach the average monthly maximum value in those months (Figure 1a). Therefore, the bias correction on the MERRA2 rainfall data is useful for reaching the maximum rainfall value that was not previously achieved by the MERRA2 data before the correction.

![Figure 1a](image1.png)

![Figure 1b](image2.png)

![Figure 1c](image3.png)

**Figure 1.** Boxplot of the 1987-2017 (a) aggregate monthly average rainfall, (b) maximum temperature, and (c) minimum temperature from MERRA2 data (red), observations (green), and bias-corrected MERRA2 (blue).

Meanwhile the average monthly maximum temperature data for MERRA2 (Figure 1b) before bias correction looks underestimated compared to the observation data. The corrected data tend to be able to reach the range of the average monthly maximum temperature data from the observation data. The average monthly maximum temperature data for MERRA2 in July-September has a minimum and maximum data distribution range that is longer than other months, so this also affects the distribution range of bias-corrected MERRA2 in July-September.

In contrast to the maximum temperature data, the monthly average minimum temperature on the MERRA2 data (Figure 1c) tends to overestimate the monthly minimum temperature average of the observation data. However, corrected MERRA2 data tends to be able to reach the average monthly minimum temperature range of observation data, although in several months, it still tends to be overestimated but it is better than when it has not been corrected.

3.2 Clustering of Java Island climate regions with the k-means method

The Java Island climate region clustering is conducted by grouping the variables of rainfall, maximum temperature, and minimum temperature from the bias-corrected MERRA2 data simultaneously. The determination of the optimal number of k in this k-means method uses the Silhouette method, which is a method of determining the number of groups based on the average of the Silhouette width, the wider the average, the clearer the differences between different groups and members who are in the same group will be increasingly similar [31].
The results of clustering on predictor variables data from bias-corrected MERRA2 that are aggregated on an annual scale (Figure 2a) produce two clusters or groups of climate regions in Java. The number of optimal k groups from the Silhouette method also shows k=2 (Figure 2b). Silhouette width (Si) > 0 (Figure 2c) indicates that the data group can be clustered well. The value of the Si width is closer to 1, meaning the better the grouping formed [31] [32].

Table 2. Descriptive statistic of predictor variables using annual aggregation.

| Cluster | Variable                              | N Grid | N  | Min   | Q1    | Mean  | Median | Q3    | Max   | SD    |
|---------|---------------------------------------|--------|----|-------|-------|-------|--------|-------|-------|-------|
| 1       | Total rainfall (mm)                    | 33     | 1023 | 893   | 1508  | 1852  | 1787   | 2117  | 4108  | 481   |
| 2       | Total rainfall (mm)                    | 8      | 248  | 1158  | 1819  | 2232  | 2111   | 2561  | 6261  | 653   |
| 1       | Maximum temperature (°C)              | 33     | 1023 | 30.2  | 31.3  | 31.8  | 31.7   | 32.2  | 33.7  | 0.75  |
| 2       | Maximum temperature (°C)              | 8      | 248  | 25.2  | 26.1  | 27.7  | 28.6   | 28.9  | 29.7  | 1.43  |
| 1       | Minimum temperature (°C)              | 33     | 1023 | 20.4  | 23.3  | 23.7  | 23.9   | 24.4  | 26.2  | 1.1   |
| 2       | Minimum temperature (°C)              | 8      | 248  | 17.2  | 18.1  | 18.8  | 18.9   | 19.3  | 20.9  | 0.77  |
| 1       | Irrigated rice productivity (ton/ha)  | 33     | 1023 | 2.84  | 4.38  | 5.14  | 5.10   | 5.89  | 7.15  | 1.02  |
| 2       | Irrigated rice productivity (ton/ha)  | 8      | 248  | 3.23  | 5.02  | 5.66  | 5.77   | 6.36  | 7.10  | 0.89  |
| 1       | Rainfed rice productivity (ton/ha)    | 33     | 1023 | 0.77  | 2.38  | 3.55  | 3.18   | 4.43  | 7.14  | 1.38  |
| 2       | Rainfed rice productivity (ton/ha)    | 8      | 248  | 1.28  | 2.93  | 4.25  | 4.22   | 5.37  | 7.11  | 1.49  |

Actually based on altitude map from [33], area of cluster 2 are dominated with highland area, meanwhile cluster 1 are dominated with lowland area. Based on descriptive statistic (Table 2), mean of total annual rainfall in cluster 2 is higher than mean of total annual rainfall in cluster 1. While mean of annual maximum and minimum temperature in cluster 2 is lower than those in cluster 1.

Figure 2. (a) The results of clustering predictor variables with annual aggregates using the k-means method (The numbering grid of MERRA2 only used in this study); (b) The plot of determining the optimal number of k using the Silhouette method; (c) Silhouette width histogram.

3.3 Comparison of the results of the k-means method with the Principal Component Analysis (PCA)

PCA can be used to compare the magnitude of the variance of the predictor variables that affect the clustering results of the k-means method. The loading plot (Figure 3a) can show the magnitude of the
loading value (positive or negative) in dimensions 1 (PC1) and 2 (PC2) on bias-corrected MERRA2 data. Meanwhile, Table 2 provides detailed information on the loading values for each PC dimension on bias-corrected MERRA2. A positive loading value means that the variable contributes positively to the principal component (PC). A negative loading value means that the variable contributes negatively to the PC. Loading value is the significant contribution value given by the variable to each PC [34].

Loading plots (Figures 3a) show that in the corrected MERRA2 data, the maximum and minimum temperature variables are located at an angle of almost 90° with the rainfall variable. In contrast, the maximum temperature variable is adjacent to the minimum temperature variable. This shows that the maximum temperature and minimum temperature variables have a negative relationship with the rainfall variable, but the correlation is low; when the rainfall variable increases, the maximum temperature and minimum temperature variables tend to decrease, while the maximum temperature and minimum temperature variables are strongly correlated with one another.

PC1 on the loading plot of the corrected MERRA2 can explain the variance of 69.9%, while PC2 can explain the variance of 24.8%. According to Table 3, the maximum and minimum temperature variables in the clustering of MERRA2 contributed negatively and strong to PC1. In contrast, the rainfall variable contributed positively and relatively strong to PC1, then on PC2, the rainfall variable contributed negatively and strong, while the maximum and minimum temperatures contribute negatively and weak. Based on the plot and the loading value of PCA on the corrected MERRA2 data, it can be seen that the maximum temperature variable and the minimum temperature variable are variables that have more influence or contribute strongly to PC1 than the rainfall variable, and this affects the grouping results in the k-means method.

Table 3. The predictor variable loading value in each dimension of the bias-corrected MERRA2

| Predictor variables | Bias-corrected MERRA2 |
|---------------------|-----------------------|
|                     | Dimension 1 | Dimension 2 | Dimension 3 |
| Rainfall            | 0.648      | -0.759      | 0.064       |
| Maximum temperature | -0.941     | -0.170      | 0.293       |
| Minimum temperature | -0.890     | -0.373      | -0.264      |

Figure 3. (a) The loading plot of predictor variables on bias-corrected MERRA2, and (b) Score plot on bias-corrected MERRA2 grid.

Score is the new coordinate value on the PC axis, which is formed from PCA results. Visually (Figures 3b), the clustering result is caused by differences in the score on PC1; this means that the clustering occurs because of the influence of the predictor variable that contributes the most to PC1 (Figures 3a), namely the maximum temperature variable and minimum temperature. So the bias-
corrected MERRA2 grid (Figure 3b), which has a lower average maximum and minimum temperature tend to be on the negative score side on PC1.

3.4 Probability of Irrigated and Rainfed Rice Productivity Results of Aquacrop Simulation in ENSO Year.

3.4.1 Irrigated Rice Productivity

Determination of the distribution pattern is useful for calculating the probability of the impact of ENSO on the results of the Aquacrop simulation, with a neutral year as the control year. Determination of the distribution pattern of rice productivity data is carried out by two stages of empirical data distribution tests. The first stage of the test is the Shapiro-Wilk normality test (hereinafter referred as the SW test), and if the results of the data SW test conclude that the data is not normally distributed, then a further test is carried out using the Kolmogorov-Smirnov Test (hereinafter referred as the KS test) [35,36]. Determination of the distribution of the empirical distribution of irrigated rice productivity simulation data is needed to calculate the estimated productivity probability when the ENSO mode is active (La Nina and El Nino) compared to the average productivity during a neutral year.

In this study, several data were produced whose distribution patterns could not be identified by the results of the empirical distribution test thus the determination of the distribution pattern was subjectively determined by visual of the histogram plot and the empirical cumulative density (ECDF) plot.

![Figure 4](image-url)  
**Figure 4.** Probability of (a) irrigated and (b) rainfed rice productivity from Aquacrop simulation (Note: in the upper panel the probability of rice productivity during La Nina > average of rice productivity during neutral years, while the lower panel probability of rice productivity during El Nino < average of rice productivity neutral years), left panel: cluster 1, right panel: cluster 2).

The distribution pattern of the simulation results of irrigated rice ranges from the lognormal, normal, gamma, and weibull distributions. The probability of irrigated rice productivity during El Nino with productivity opportunities lower than the average productivity in a neutral year (El Nino productivity probability < average of neutral years productivity) increases with the increasing El Nino mode (Figure 4a), both in cluster 1 and cluster 2. While the probability during La Nina with the opportunity for irrigated rice productivity that is greater than the average productivity in a neutral year (probability of La Nina rice > average of neutral years productivity) is different for clusters 1 and 2. In cluster 2, the probability level increases with the increase of La Nina mode, while in cluster 1 the probability in weak La Nina mode is higher than the probability of moderate La Nina mode.

3.4.2 Rainfed Rice Productivity

The distribution pattern of the simulation results of rainfed rice productivity ranges from the lognormal, normal, gamma, and weibull distributions. Similar to the results of probability calculations
for irrigated rice, the probability of rainfed rice productivity during El Nino with productivity opportunities lower than the average productivity in a neutral year (probability of El Nino rice productivity < average of neutral years productivity) increases with increasing El Nino mode (Figure 4b), both in cluster 1 and cluster 2. While the probability during La Nina with rainfed rice productivity opportunities is greater than the average productivity during a neutral year (probability of La Nina productivity > average of neutral years productivity) is different for clusters 1 and 2. In cluster 2, the probability level increases with increasing La Nina mode, while in cluster 1 the probability in weak La Nina mode is higher than the probability of medium La Nina mode, this is due to during the 2016 weak La Nina accompanied by a strong negative IOD phenomenon which amplifies rainfall intensity during September-October-November 2016 [37,38]. Therefore, because the Aquacrop crop modeling is more sensitive to changes in rainfall variables, the probability of weak La Nina mode is higher than the probability of moderate La Nina mode.

The probability range (probability of La Nina productivity > average of neutral years productivity) on rainfed rice productivity from Aquacrop simulation during La Nina ranges from ~45-60 % in cluster 1 area and ~80-82 % in cluster 2. While the probability range (P El Nino < neutral year mean) in the cluster 1 region ranged from ~78-97 % and the cluster 2 region ranged from ~40-92 %.

The probability of a decrease in irrigated and non-irrigated rice productivity during El Nino in cluster 1 is higher than in cluster 2. Meanwhile, the probability of irrigated and non-irrigated rice productivity experience increasing productivity during La Nina in cluster 2 is higher than in cluster 1. It means that cluster 1 is more vulnerable to El Nino and La Nina than cluster 2. Besides that, the rainfed system is also more vulnerable to El Nino and La Nina than the irrigated yield rice system.

3.5 Validation between Actual Productivity and Simulation Productivity of Aquacrop

3.5.1 Validation for Irrigated Rice Productivity

The actual productivity data for irrigated and rainfed rice available on the official website of BPS for each province is in the form of annual productivity data per city/district within the province and the time span is only 13-14 years from 2001/2002-2014. The validation results of the correlation value of the simulation results of the Aquacrop application with the actual productivity of irrigated rice show a fairly low correlation, ranging from -0.01 to 0.58 with a data length of 13-14 years (Figure 5a) which means the Aquacrop application with default input on variables other than variables climate (daily rainfall, daily maximum and minimum temperature) and geographic location (longitude, latitude, and altitude) derived from the corrected MERRA2 grid data provide estimation results that have low reliability on the actual productivity of irrigated rice, this is as a result of input parameters used is the minimum input parameters required in running the Aquacrop application. However, the Aquacrop application which is inputted using the input parameters that have been calibrated such as the parameters of canopy cover, root depth effectiveness, and rice seed sowing rate provide a better estimate for example by [39] in Subang, West Java, Aquacrop can estimate the productivity of irrigated rice with a correlation coefficient of 0.8 (11 years of data from 2000-2010), Root Mean Square Error (RMSE) of 0.53, and a Mean Square Error (MSE) of 9%.

Because the correlation results for all available years yielded a fairly low correlation value, then the correlation value for the ENSO composite years were calculated. Calculation of the correlation value between the actual productivity and the simulated productivity of irrigated rice in the La Nina composite years resulted in a correlation value that varied between negative and positive correlations (-0.82 to 0.98) with a data length of 5 years (2005, 2007, 2008, 2010 and 2011) with the strongest negative correlation in Boyolali Regency, Central Java and the strongest positive correlation in Probolinggo City, East Java (Figure 5c).
The whole years
Neutral years
La Nina years
El Nino years

Figure 5. The correlation plot between actual productivity and simulated productivity of irrigated rice Aquacrop per city/district in Java in (a) the whole year (2001/2002-2014); (b) composite of neutral years; (c) composite of La Nina years; and (d) composite of El Nino years.

Then the range of correlation values between actual productivity and simulated productivity of irrigated rice in the El Nino composite ranged from 0.045-0.98, with a data length of 5 years (2002, 2004, 2006, 2009 and 2014) where the highest positive correlation value was in Brebes Regency, Central Java (Figure 5b). While the neutral year composite produces a range of correlation values from negative correlation values and positive correlations (between -1 and 1) but with relatively very short data lengths between 3 and 4 years (2001, 2003, 2012, and 2013) (Figure 5d). The correlation values of the ENSO composite years are relatively higher than the correlation value in the whole years, therefore although the correlation values are relatively high, it becomes less reliable because the length of the data is inadequate since it does not have a long period of time.

3.5.2 Validation for Rainfed Rice Productivity
The correlation results for the whole year (Figure 6a) between actual productivity and non-irrigated rice productivity from Aquacrop simulation have a better correlation value than the validation results for irrigated rice, ranging from -0.40 to 0.78, with the highest correlation being in Mojokerto Regency, East Java. The correlation results which tend to be better in the validation results from the simulation of rainfed rice productivity compared to the validation of irrigated rice prove that the Aquacrop application is sensitive to rainfall variables.

In the ENSO composite years (Figure 6b, c, d), for the neutral years, the correlation range is between -1 and 1 this is because the available data range is short, resulting in a high correlation value. Then in the La Nina years, the correlation ranges from -0.88 to 0.95, each of which comes from validation in Jombang Regency and Mojokerto Regency, East Java. Correlation results for the El Nino years produce a correlation range ranging from 0 – 0.94, with the highest correlation value being in the Pekalongan Regency, Central Java.
4. Conclusion
The k-means clustering method applied to the rainfall, maximum, and minimum temperature variables from the bias-corrected MERRA2 data resulted in two climate regions. Although the loading value from the PCA results shows that the maximum and minimum temperature variables contribute the most to the clustering produced by the k-means method. However, the rainfall variable also contributes to the climate region cluster on Java Island, as indicated by the probability value of cluster 1 area to experience a decline in productivity during El Nino from the results of the Aquacrop simulation which is generally higher than the probability in cluster 2 both in irrigated and rainfed rice productivity. Likewise, the probability value for cluster 1 area to experience an increase in productivity during La Nina from the results of the Aquacrop simulation is generally lower than the probability in cluster 2 for both irrigated and rainfed rice productivity. This is because the Aquacrop simulation tends to be sensitive to rainfall variables and the cluster 2 area has a higher average rainfall than the cluster 1 area.

In general, the validation results show that between the actual productivity and the productivity of the Aquacrop simulation results, the correlation value is quite low, both in the simulation results for irrigated and rainfed rice productivity. Then it is known that the validation of the ENSO composite years show that in the El Nino years, the correlation between actual productivity and the productivity of the Aquacrop simulation results shows a positive correlation value, while in the neutral and La Nina years the correlation value varies between negative and positive values.

References
[1] Panuju D R, Mizuno K and Trisasongko B H 2013 The dynamics of rice production in Indonesia 1961–2009 J. Saudi Soc. Agric. Sci. 12 27–37
[2] Sumarsono D S D and Ningsih S 2018 The Effect of Own-Source Revenue, Population, and Population Density on the Capital Expenditure of City and District Governments on Java Island: Proceedings of the Journal of Contemporary Accounting and Economics Symposium 2018 on Special Session for Indonesian Study Journal of Contemporary Accounting and Economics Symposium 2018 on Special Session for Indonesian Study (Bali, Indonesia: SCITEPRESS - Science and Technology Publications) pp 581–7
(35) Razali N M and Wah Y B 2011 Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests J. Stat. Model. Anal. 2 21–33
(36) Yap B W and Sim C H 2011 Comparisons of various types of normality tests J. Stat. Comput. Simul. 81 2141–55
(37) Lim E-P and Hendon H H 2017 Causes and Predictability of the Negative Indian Ocean Dipole and Its Impact on La Niña During 2016 Nat. Sci. Rep. 7 1–11
(38) Nur’utami M N and Hidayat R 2016 Influences of IOD and ENSO to Indonesian Rainfall Variability: Role of Atmosphere-ocean Interaction in the Indo-pacific Sector Procedia Environ. Sci. 33 196–203
(39) Santikayasa I, Amdan W, Perdinan and Sugiarito Y 2018 Quantification the Impact of Climate Change on Paddy Field Yield Production in Subang, Indonesia IOP Conf. Ser. Earth Environ. Sci. 187 012059

Acknowledgments
This paper is part of first author master thesis in Faculty of Earth Science and Technology, Bandung Institute of Technology (ITB) that funded by the scholarship from the Center of Training and Education, Indonesian Agency for Meteorology Climatology and Geophysics (BMKG). We thank to advanced comment and suggestion from anonymous reviewers.

Supplementary data

Table 4. Determination of empirical distribution with Shapiro-Wilk Test (SW) and Kolmogorov-Smirnov Test (KS) on Irrigated Rice Productivity.

| Kel. | Var. | Mode | SW p-value | Decision | Kolmogorov-Smirnov one sample test p-value | Distribution |
|------|------|------|------------|----------|-------------------------------------------|--------------|
| 1    | EN   | weak | 0.014**   | reject H0 | lnorm = 0.4 gamma= 0.5 weibull=0.2        | gamma        |
| 1    | EN   | moderate | 0.000*** | reject H0 | lnorm = 1 gamma= 0.8 weibull=0.1          | lognormal    |
| 1    | EN   | strong  | 0.007*** | reject H0 | lnorm = 0.7 gamma= 0.6 weibull=0.6        | lognormal    |
| 1    | LN   | weak  | 0.000*** | reject H0 | lnorm = 0.07 gamma= 0.1 weibull=0.04      | gamma        |
| 1    | LN   | moderate | 0.021**  | reject H0 | lnorm = 0.7 gamma= 0.8 weibull=0.7        | gamma        |
| 1    | LN   | strong | 0.000*** | reject H0 | lnorm = 0.006 gamma= 0.01 weibull=0.05    | weibull      |
| 1    | N    | weak  | 0.001*** | reject H0 | lnorm = 0.3 gamma= 0.2 weibull=0.2        | normal       |
| 2    | EN   | weak  | 0.524    | accept H0 |                                           | normal       |
| 2    | EN   | moderate | 0.305    | accept H0 |                                           | normal       |
| 2    | EN   | strong  | 0.879    | accept H0 |                                           | normal       |
| 2    | LN   | weak  | 0.002**  | reject H0 | lnorm = 0.3 gamma= 0.2 weibull=0.2        | lognormal(1) |
| 2    | LN   | moderate | 0.213    | accept H0 |                                           | normal       |
| 2    | LN   | strong | 0.006**  | reject H0 | lnorm = 0.1 gamma= 0.1 weibull=0.3        | weibull      |
| 2    | N    | accept H0 | 0.101    | accept H0 |                                           | normal       |

Notes: (1): probability is determined subjectively by visual of the plot

Table 5. Determination of empirical distribution with Shapiro-Wilk Test (SW) and Kolmogorov-Smirnov Test (KS) on Rainfed Rice Productivity.

| Kel. | Var. | Mode | SW p-value | Decision | Kolmogorov-Smirnov one sample test p-value | Distribution |
|------|------|------|------------|----------|-------------------------------------------|--------------|
| 1    | EN   | weak  | 0.000*** | reject H0 | lnorm = 0.04 gamma= 0.07 weibull=0.02     | lognormal    |
| 1    | EN   | moderate | 0.000*** | reject H0 | lnorm = 0.000 gamma= 0.000 weibull=0.000 | lognormal(1) |
| 1    | EN   | strong  | 0.000*** | reject H0 | lnorm = 0.06 lnorm= 0.08 gamma=0.02       | lognormal    |
| 1    | LN   | weak  | 0.000*** | reject H0 | gamma=0.05 lnorm= 0.2 weibull=0.002       | lognormal    |
| 1    | LN   | moderate | 0.000*** | reject H0 | gamma=0.5 lnorm= 0.6 weibull=0.4          | lognormal    |
| 1    | LN   | strong  | 0.000*** | reject H0 | gamma=0.1 lnorm= 0.03 weibull=0.09        | gamma        |
| 1    | N    | weak  | 0.000*** | reject H0 | gamma=0.03 lnorm= 0.01 weibull=0.01       | gamma        |
| 2    | EN   | weak  | 0.015*   | reject H0 | lnorm = 0.6 gamma= 0.7 weibull=0.07       | gamma(4)     |
| 2    | EN   | moderate | 0.52     | accept H0 |                                           | normal       |
| 2    | EN   | strong  | 0.047*   | reject H0 | lnorm = 0.7 gamma= 0.6 weibull=0.3        | lognormal    |
| 2    | LN   | weak  | 0.015*   | reject H0 | lnorm = 0.4 gamma= 0.6 weibull=0.6        | gamma(4)     |
| 2    | LN   | moderate | 0.026*   | reject H0 | lnorm = 0.9 gamma= 0.9 weibull=0.07       | gamma(4)     |
| 2    | LN   | strong  | 0.002**  | reject H0 | lnorm = 0.1 gamma= 0.1 weibull=0.2        | weibull      |
| 2    | N    | weak  | 0.076*   | reject H0 | lnorm = 0.1 gamma= 0.2 weibull=0.7        | weibull      |

Notes: (4): probability is determined subjectively by visual of the plot