Contents lists available at DJM

DESIMAL: JURNAL MATEMATIKA

p-ISSN: 2613-9073 (print), e-ISSN: 2613-9081 (online), DOI 10.24042/djm

http://ejournal.radenintan.ac.id/index.php/desimal/index

Canonical Correlation Analysis of Global Climate Elements and Rainfall in the West Java Regions

Arisya Maulina Bowo, Iin Irianingsih, Budi Nurani Ruchjana*

Universitas Padjadjaran, Indonesia

ARTICLE INFO

Article History

Received : 02-02-2020
Revised : 15-04-2020
Accepted : 06-05-2020
Published : 20-05-2020

Keywords:
Canonical Correlation;
Global Climate Elements;
Rainfall;
Relationship;
Simultaneous.

*Correspondence: E-mail:
budi.nurani@unpad.ac.id

Doi:
10.24042/djm.v3i2.5870

ABSTRACT

Indonesia has a diversity of climate influenced by several global phenomena such as El Nino Southern Oscillation (ENSO), *Indian Ocean Dipole* (IOD)*, and Asian-Australian Monsoon. Continuously climate changing indirectly causes a hydrometeorological disaster. The purpose of this study was to analyze the relationship between global climate elements* (ENSO, IOD, *Asian-Australian Monsoon*) *with rainfall in the West Java regions* (*Bogor Regency, Bandung Regency, Sukabumi Regency, Garut Regency, and Kuningan Regency*) *simultaneously. The selection of the five regions was based on the natural disaster reports of Badan Nasional Penanggulangan Bencana* (BNPB)*. The research method used was a quantitative research method through one of multivariate analysis technique called canonical correlation analysis. The results of this study indicate that there was a simultaneous relationship between global climate elements, with rainfall in the West Java regions by 0.819. The global climate element and rainfall in the West Java regions that most influenced the relationship were Asian-Austalian Monsoon and Kuningan Regency rainfall.*

http://ejournal.radenintan.ac.id/index.php/desimal/index

INTRODUCTION

Indonesia is a country with strategic geographical conditions that causes Indonesia to have a diversity of weather and climate. Climate diversity in Indonesia is influenced by several global phenomena such as El Nino Sourthen Oscillation (ENSO) on the Pacific Ocean, Indian Ocean Dipole (IOD) on the Indian Ocean, and the Monsoon Asia-Australia (Ridwan, 2019).

Climate changes because of interactions between elements of climate, including interactions between global phenomena mentioned earlier. Climate change that occurs continuously indirectly causes hydrometeorological disasters. According to Aldrian *et al.* (2011), hydrometeorological disasters consist of landslides, floods, droughts, hurricanes, tidal waves, abrasion, and forest and land fires. Based on the natural disaster reports of Badan Nasional Penanggulangan Bencana (BNPB), Bogor Regency, Bandung Regency, Sukabumi Regency, Garut Regency, and Kuningan District are five regions prone to hydro-meteorological disasters in West Java over the past 10 years.

Canonical correlation analysis is a multivariate analysis technique that aims to measure the relationship between a set of independent variables with a set of dependent variables simultaneously. It was first introduced by Hotelling (1936). Hotelling measured the correlation between arithmetic speed and arithmetic power to reading speed and reading power.

Canonical correlation analysis has been developed by several researchers in various cases. Chaghooshi et al. (2015) applied canonical correlation analysis to analyze the relationship between supply chain quality management and the competitive advantages of Sahami Alyaf (SA) Company, Iran. Irianingsih *et al.* (2016) analyzed the relationship between learning behavior and learning achievement in students of SMPN 1 Sukasari Purwakarta with canonical correlation analysis. Rustiana *et al.* (2017) conducted a study on the prediction of rainfall in the Cimanuk watershed with canonical correlation analysis. However, based on these studies, no one has tested the canonical correlation analysis assumptions completely, i.e., linearity, multivariate normality, homoscedasticity, and there is no multicollinearity between variables in a set of variables. Therefore, the authors are interested in studying canonical correlation analysis by including all the prerequisite assumption tests of canonical correlation analysis in explaining the relationship between global climate elements, i.e., ENSO, IOD, and the Monsoon Asia-Australia with rainfall in the West Java regions, i.e., Bogor Regency, Bandung Regency, Sukabumi Regency, Garut Regency, and Kuningan Regency.

METHOD

This research uses quantitative research methods through one multivariate analysis technique called canonical correlation analysis. The analysis is used to measure the linear relationship between a set of independent variables with a set of independent variables; canonical variables are formed for each set. A canonical variable is a linear combination of a set of variables. Canonical correlation analysis constructs a canonical function that maximizes the canonical correlation coefficient between two canonical variables (Hair, *et al.,* 2009).

The analysis begins with testing the prerequisite assumptions in the canonical correlation analysis of each variable. If there are assumptions that are not met, then the variable transformation is needed. If all assumptions are met, then proceed to determining the canonical functions and coefficient estimators. After the canonical correlation coefficient is obtained, the next step is to test the significance of the canonical correlation. If there is a significant canonical correlation coefficient, then it is followed by a redundancy analysis. Then, the final step is the interpretation of canonical variables.

The prerequisite assumptions of canonical correlation analysis

Following are the prerequisite assumptions that must be fulfilled in the canonical correlation analysis.

a. Linearity, a linear relationship (linearity) between the independent variable and the dependent variable. Linearity affects two aspects of canonical correlation results. First, the canonical correlation coefficient of a pair of canonical variables is based on a linear relationship. Second, canonical correlation analysis maximizes linear relationships between two sets of variables (Hair, *et al.*, 2009).

Suppose there are a number of independent variables $X_1, X_2, \dots, X_p$ and a number of dependent variables $Y_1, Y_2, \dots, Y_q$. Linearity assumption testing is done by ANOVA or ANOVA lack of fit test of X_j and Y_l $(j = 1, 2, \dots, p;\ l = 1, 2, \dots, q)$ regression in a univariate manner. The lack of fit test is done if there is repeated data or one variable X value produces several different Y variable values.

- ANOVA
  $\mathrm{H_0}$: $b_1 = 0$ (X_j does not effect Y_l linearly, $j = 1, 2, .., p$; $l = 1, 2, \dots, q$)
  $\mathrm{H_1}$: $b_1 \neq 0$ (X_j effect Y_l linearly, $j = 1, 2, .., p$; $l = 1, 2, \dots, q$)
  Decision criteria: Reject $\mathrm{H_0}$ if $F \geq F_{TABEL}$.
- ANOVA *Lack of Fit*
  $\mathrm{H_0}$: The relationship between X_j and Y_l is linear ($j = 1, 2, \dots, p$; $l = 1, 2, \dots, q$)
  $\mathrm{H_1}$: The relationship between X_j and Y_l is not linear ($j = 1, 2, \dots, p$; $l = 1, 2, \dots, q$)
  Decision criteria: Reject $\mathrm{H_0}$ if $F < F_{TABEL}$ (Walpole, 2011).

b. Multivariate normality. This assumption is needed for statistical tests of the significance of each canonical function. Data that are multivariate normal distributed can be seen through the plot between the Mahalanobis distance (d_i^2) and the quantile of the chi-square distribution $\left(\frac{i-\frac{1}{2}}{n}\right)$ (Johnson and Wichern, 1998).

c. Homoscedasticity. This assumption is said to be important in canonical correlation analysis because it is the opposite of heteroscedasticity which can reduce intervariable correlations (Hair, *et al.*, 2009). Homoscedasticity can be known through Glejser testing.

d. There is no multicollinearity between variables in a set of variables. Gozhali (in Yudiaatmaja, 2013), states that multicollinearity occurs when two or more variables correlate very strongly. A very strong correlation is meant when $\rho \geq 0.9$.

Canonical functions and coefficient estimators determination

Suppose there are a number of independent variables $X_1, X_2, \dots, X_p$ denoted as random vector **X** and a number of dependent variables $Y_1, Y_2, \dots, Y_q$ denoted as random vector **Y**.

$$\mathbf{X} = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{pmatrix} \qquad \mathbf{Y} = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_q \end{pmatrix}$$

The characteristics of the random vectors **X** and **Y** are as follows.

$$E(\mathbf{X}) = \boldsymbol{\mu}_{\mathbf{X}}$$
$$E(\mathbf{Y}) = \boldsymbol{\mu}_{\mathbf{Y}}$$
$$Cov(\mathbf{Y}) = \boldsymbol{\Sigma}_{\mathbf{YY}}$$
$$Cov(\mathbf{X}) = \boldsymbol{\Sigma}_{\mathbf{XX}}$$
$$Cov(\mathbf{X}, \mathbf{Y}) = \boldsymbol{\Sigma}_{\mathbf{XY}} = \boldsymbol{\Sigma}'_{\mathbf{YX}}$$

Canonical correlation is obtained by measuring the linear relationship between the linear combination of **X** and the linear combination of **Y**. To determine the linear combination, the two variable sets can be arranged into

$$U = a_1 X_1 + a_2 X_2 + \cdots + a_p X_p = \mathbf{a}'\mathbf{X}$$
$$\boldsymbol{V} = \boldsymbol{b_1 Y_1} + \boldsymbol{b_2 Y_2} + \cdots + \boldsymbol{b_q Y_q} = \mathbf{b}'\mathbf{Y}$$

so

$$Var(U) = \mathbf{a}' Cov(\mathbf{X}) \mathbf{a} = \mathbf{a}' \boldsymbol{\Sigma}_{\mathbf{XX}} \mathbf{a}$$
$$Var(V) = \mathbf{b}' Cov(\mathbf{Y}) \mathbf{b} = \mathbf{b}' \boldsymbol{\Sigma}_{\mathbf{YY}} \mathbf{b}$$
$$\boldsymbol{Cov}(\boldsymbol{U}, \boldsymbol{V}) = \mathbf{a}' \boldsymbol{Cov}(\mathbf{X}, \mathbf{Y}) \mathbf{b} = \mathbf{a}' \boldsymbol{\Sigma}_{\mathbf{XY}} \mathbf{b}$$

The number of linear combination pairs formed by **X** and **Y** is defined as $\boldsymbol{k} =$

$\boldsymbol{min}\,(\boldsymbol{p},\boldsymbol{q})$. Canonical correlation is obtained by

$$\boldsymbol{\rho_{UV}} = \frac{\mathbf{a}'\boldsymbol{\Sigma}_{\mathbf{XY}}\mathbf{b}}{\sqrt{\mathbf{a}'\boldsymbol{\Sigma}_{\mathbf{XX}}\mathbf{a}}\sqrt{\mathbf{b}'\boldsymbol{\Sigma}_{\mathbf{YY}}\mathbf{b}}} = \frac{\boldsymbol{Cov(U,V)}}{\sqrt{\boldsymbol{Var(U)}}\sqrt{\boldsymbol{Var(V)}}} \quad (1)$$

Johnson and Wichern (1998) state that the kth pair of canonical variables, is the pair of linear combinations U_k and V_k having unit variances, which maximize the correlation among all choices uncorrelated with the previous $k-1$ canonical variable pairs.

Therefore, using the Lagrange Multiplier method, the coefficient vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ that maximize the correlation between U and V can be obtained by determining the eigenvectors of the matrix $\boldsymbol{\Sigma}_{\mathbf{XX}}^{-1}\boldsymbol{\Sigma}_{\mathbf{XY}}\boldsymbol{\Sigma}_{\mathbf{YY}}^{-1}\boldsymbol{\Sigma}_{\mathbf{YX}}$ and $\boldsymbol{\Sigma}_{\mathbf{YY}}^{-1}\boldsymbol{\Sigma}_{\mathbf{YX}}\boldsymbol{\Sigma}_{\mathbf{XX}}^{-1}\boldsymbol{\Sigma}_{\mathbf{XY}}$. The root of the eigenvalues corresponding to the two matrices is the correlation coefficient between U and V (Qiu, *et al.*, 2016).

The characteristics of the canonical variables are as follows.

a. $Var(U_r) = Var(V_r) = 1$

b. $Cov(U_r, U_s) = Cov(U_s, U_r) = 0$, $r \neq s$

c. $Cov(V_r, V_s) = Cov(V_r, V_s) = 0$, $r \neq s$

d. $Cov(U_r, V_s) = Cor(U_r, V_s) = 0$, $r \neq s$

for $r, s = 1,2, \dots, k$ (Johnson dan Wichern, 1998).

The significance test of canonical correlation

There are two hypotheses tested in canonical correlation analysis, namely:

1. Overall Canonical Correlation Test

   $H_0: \rho_1 = \rho_2 = \cdots = \rho_k = 0$ (All canonical correlations are not significant)

   $H_1: \rho_r \neq 0$ ($r = 1,2, \dots, k$, there is at least one significant canonical correlation)

   Statistics test:

   $$\boldsymbol{B} = -\left[\boldsymbol{n} - \mathbf{1} - \frac{\mathbf{1}}{\mathbf{2}}(\boldsymbol{p} + \boldsymbol{q} + \mathbf{1})\right] \boldsymbol{ln\,\Lambda} \quad (2)$$

   where

   $\Lambda = \prod_{r=1}^{k}(1 - \rho_r^2)$

   n : the number of the observations

   Decision criteria: Reject H_0 if $B > \chi_\alpha^2$ with pq degrees of freedom at α level of significant.

2. Individual Canonical Correlation Test

   $H_0: \rho_m = 0$ (The mth canonical correlation is not significant, $m = 1, 2, \dots, k$)

   $H_1: \rho_m \neq 0$ (The mth canonical correlation is significant)

   Statistics test:

   $$\boldsymbol{B_m} = -\left[\boldsymbol{n} - \mathbf{1} - \frac{1}{2}(\boldsymbol{p} + \boldsymbol{q} + \mathbf{1})\right] \boldsymbol{ln\,\Lambda_m} \quad (3)$$

   where

   $\Lambda_m = \prod_{m=r}^{k}(1 - \rho_m^2)$

   n : the number of the observations

   Decision criteria: Reject H_0 if $B_m > \chi_\alpha^2$ with $(p - m + 1)(q - m + 1)$ degrees of freedom at α level of significant.

Redundancy Analysis

Redundancy is a value that calculates the proportion of total variance that can be explained by the canonical variables of the dependent variable and the independent variable.

- The redundancy **Y** explained by **U** is defined as

  $$\boldsymbol{Rd}_{(\mathbf{Y|U})} = \frac{\sum_{r=1}^{k}\sum_{l=1}^{q}\boldsymbol{\rho}_{Y_l V_r}^2\,\boldsymbol{\rho}_r^2}{\boldsymbol{q}} \quad (4)$$

  where $\mathbf{U} = (U_1 \quad U_2 \quad \cdots \quad U_k)'$.

- The redundancy **X** explained by **V** is defined as

  $$\boldsymbol{Rd}_{(\mathbf{X|V})} = \frac{\sum_{r=1}^{k}\sum_{j=1}^{p}\boldsymbol{\rho}_{X_j U_r}^2\,\boldsymbol{\rho}_r^2}{\boldsymbol{p}} \quad (5)$$

  where $\mathbf{V} = (V_1 \quad V_2 \quad \cdots \quad V_k)'$ (Rencher, 1998).

- The redundancy index **Y** explained by **U** is defined as

$$\boldsymbol{RI}_{(\mathbf{Y}|\mathbf{U})} = \boldsymbol{Rd}_{(\mathbf{Y}|\mathbf{U})}\boldsymbol{\rho}_{\boldsymbol{r}}^{2} \quad (6)$$

where $r = 1,2,\dots,k$.

- The redundancy index **X** explained by **V** is defined as

$$\boldsymbol{RI}_{(\mathbf{X}|\mathbf{V})} = \boldsymbol{Rd}_{(\mathbf{X}|\mathbf{V})}\boldsymbol{\rho}_{\boldsymbol{r}}^{2} \quad (7)$$

where $r = 1,2,\dots,k$ (Hair *et al*., 2009).

The redundancy between **X** and **Y** is measured by squaring the canonical correlation coefficient and is called the redundancy coefficient (Rencher, 1998).

$$\boldsymbol{R}_{\boldsymbol{r}}^{2} = \boldsymbol{\rho}_{\boldsymbol{r}}^{2}, \boldsymbol{r} = \mathbf{1},\mathbf{2},\dots,\boldsymbol{k} \quad (8)$$

Interpretation of canonical variables

Hair, *et al.* (2009) explained that there are three methods for interpreting canonical variables, namely:

1. Canonical Weights
   Canonical weights are canonical coefficients **a** and **b** that multiplied by the standard deviation of the corresponding variables so that they become standard. Canonical weights are interpreted as the contribution of origin variables to canonical variables (Rencher, 1998).
2. Canonical Loadings
   Canonical loadings is referred to as canonical structure correlation or simple linear correlation between the original variables and each of its canonical variables.
   The canonical load of variable **X** is obtained by the following formula.
   $$\mathbf{R}_{XU} = \mathbf{R}_{XX}\mathbf{c} \quad (9)$$
   where $\mathbf{R}_{XX}$ is the correlation matrix of the vector variables **X** and **c** is the standardized canonical coefficient **a**.
   The canonical load of variable **Y** is obtained by the following formula.
   $$\mathbf{R}_{YV} = \mathbf{R}_{YY}\mathbf{d} \quad (10)$$
   where $\mathbf{R}_{YY}$ is the correlation matrix of the variable vector **Y** and **d** is the standardized canonical coefficient **b** (Rencher, 1998).
3. Canonical Cross-Loadings
   The canonical cross-loadings can be calculated from the correlation between the origin variable and the canonical variable which is incompatible with the origin variable.
   The canonical cross-loadings of variable **X** is obtained by the following formula.
   $$\mathbf{R}_{XV} = \mathbf{R}_{XU}\boldsymbol{\rho}_{\boldsymbol{k}} \quad (11)$$
   The canonical cross-loadings of variable **Y** is obtained by the following formula.
   $$\mathbf{R}_{YU} = \mathbf{R}_{YV}\boldsymbol{\rho}_{\boldsymbol{k}} \quad (12)$$
   (Hair, *et al.*, 2009)

Data

Data used in this study are rainfall index of five West Java regions (Bogor Regency, Bandung Regency, Sukabumi Regency, Garut Regency, and Kuningan District), Nino3.4 index, DMI (Dipole Mode Index), and AUSMI (Australian Monsoon Index) from January 2010 - December 2018 with amount 108 data.

The rainfall index data is obtained from calculations on the Global Satellite Mapping of Precipitation (GSMap) rainfall data using the equation defined by Mulyana in Ningsih and Putranto (2019) as follows.

$$\textit{Rainfall index} = \frac{\boldsymbol{x}_{\boldsymbol{i}} - \bar{\boldsymbol{x}}}{\boldsymbol{s}_{\boldsymbol{x}}} \quad (13)$$

where x_i is the *ith* monthly rainfall data; $\bar{x}$ and s_x are respectively the average and standard deviation of the rainfall at a certain time period. Nino3.4 index data was obtained from the National Oceanic and Atmospheric Administration (NOAA) website. DMI data were obtained from the Japan Agency for Marine-Earth Science and Technology (JAMSTEC) website. AUSMI data were obtained from Lembaga Penerbangan dan Antariksa Nasional (LAPAN) Bandung.

The independent variables in this study are Nino3.4 index (X_1), DMI (X_2), and AUSMI (X_3). The dependent variables are Bogor Regency Rainfall Index (Y_1), Bandung Regency Rainfall Index (Y_2), Sukabumi Regency Rainfall Index (Y_3), Garut Regency Rainfall Index (Y_4), and Kuningan Regency Rainfall Index (Y_5).

RESULTS AND DISCUSSION

The prerequisite assumption test results of correlation analysis for data

a. Linearity Test Results

Linearity testing of variables X_1 and X_2 on each variable $Y_l\ (l = 1, 2, 3, 4, 5)$ is done by ANOVA lack of fit test because there is repeated data on variables X_1 dan X_2. Based on the linearity test procedure assisted with SPSS 23, the results of the linearity test can be seen in Table 1, Table 2, and Table 3.

Table 1. The ANOVA *Lack of Fit* Test Results of X_1

	F
$Y_1 * X_1$	1.254
$Y_2 * X_1$	1.214
$Y_3 * X_1$	1.263
$Y_4 * X_1$	0.961
$Y_5 * X_1$	1.164

Based on Table 1, all values of $F < F_{0.01(87,19)} = 2.62$, means that H0 is accepted. Thus, it can be concluded that there is linearity between X_1 with each variable $Y_l\ (l = 1, 2, 3, 4, 5)$.

Table 2. The ANOVA *Lack of Fit* Test Results of X_2

	F
$Y_1 * X_2$	36.981
$Y_2 * X_2$	2.013
$Y_3 * X_2$	4.485
$Y_4 * X_2$	1.569
$Y_5 * X_2$	3.086

Based on Table 2, all values of $F < F_{0.01(105,1)} = 6336$, means that H0 is accepted. Thus, it can be concluded that there is linearity between X_2 with each variable $Y_l\ (l = 1, 2, 3, 4, 5)$.

Table 3. The ANOVA Test Results of X_3

	F
$Y_1 * X_3$	47.239
$Y_2 * X_3$	150.779
$Y_3 * X_3$	38.177
$Y_4 * X_3$	97.483
$Y_5 * X_3$	193.551

For the ANOVA test in Table 3, it shows the linear effect test and obtained $F > F_{0.01(1,106)} = 6.88$, it means H0 is rejected, so it can be concluded that there is a linear effect between the variables X_3 with each variable $Y_l\ (l = 1, 2, 3, 4, 5)$.

b. Multivariate Normality Test Results

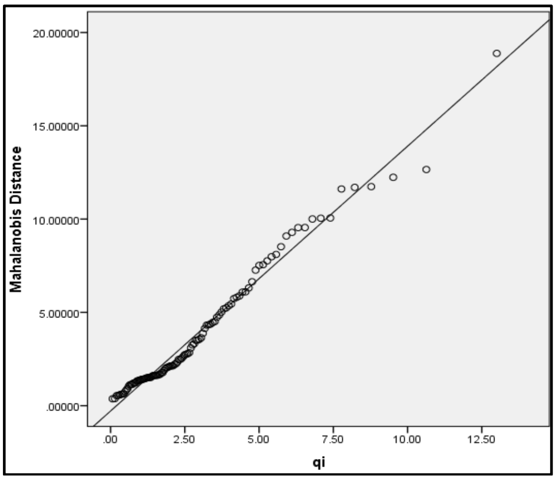

Figure 1. A Chi-*Square* Plot for Dependent Variables

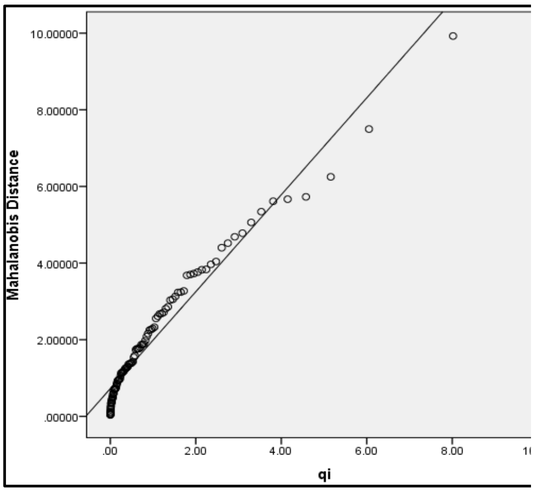

Figure 2. A Chi-*Square* Plot for Independent Variables

Based on Figure 1 and Figure 2, the plot between the Mahalanobis distance and the quintile of the chi-square distribution in the set of dependent variables and independent variables formed resembles a straight line. It means that the data in the set of dependent variable and the data in the set of independent variables are multivariate normal distributed.

c. Heteroscedasticity Test Results

Homoscedasticity assumption testing is done through the Glejser test, which is regressing the independent variable with the absolute value of the residual dependent variable. Out of five dependent variables and three independent variables, the Glejser test results obtained prove the existence of heteroscedasticity at Y_2 and Y_5 because of the value $F > F_{TABEL}$. These results can be seen in Table 4 and Table 5.

Table 4. The Regression of Independent Variables and The Absolute Value of Residual Y_2

	Sum of Squares	df	*Mean Square*	F
Regression	1.822	3	0.607	4.658
Residual	13.560	104	0.130	
Total	15.382	107		

Table 5. The Regression of Independent Variables and The Absolute Value of Residual Y_5

	Sum of Squares	df	*Mean Square*	F
Regression	2.363	3	0.788	7.469
Residual	10.967	104	0.105	
Total	13.330	107		

To overcome the heteroscedasticity, the authors carry out the neglog transformation that is defined as

$$nl(x_i) = \begin{cases} \log(x_i + 1) & x_i \geq 0 \\ -\log(-x_i + 1) & x_i < 0 \end{cases} \quad (14)$$

where x_i is the *ith* data observation of variable X for $i = 1, 2, \dots, n$ (Whittaker, *et al.*, 2005) on all variables. After the transformation is done, the linearity, multivariate normality, and homoscedasticity tests are repeated on new variables. The results of the three tests indicate that all three assumptions were satisfied.

d. Multicollinearity Test Results

Based on the results of calculations with SPSS 23, the correlation between variables Y_1 and Y_3, Y_2 and

Y_4, Y_2 and Y_5, with Y_3 and Y_4 are more than 0.9. It means there are multicollinearity. To overcome the multicollinearity in the dependent variable group, the authors chose to eliminate the variables Y_2 dan Y_3. Testing the assumption of linearity, multivariate normal distribution, homoscedasticity, and multicollinearity is done again and the results obtained indicate that all four assumptions are satisfied.

The results of canonical functions and coefficient estimators determination

Following are the results of the calculation of canonical correlation with SPSS 23.

Table 6. Canonical Correlation

***kth* Function**	**Canonical Coefficient**
1	0.819
2	0.169
3	0.093

Based on calculations assisted with SPSS 23, the canonical variables formed are

$U_1 = 0.084X_1 - 0.315X_2 - 0.699X_3$
$U_2 = 0.409X_1 + 2.526X_2 + 0.138X_3$
$U_3 = 1.805X_1 - 1.501X_2 + 0.022X_3$
$V_1 = -0.362Y_1 - 0.08Y_4 + 1.964Y_5$
$V_2 = 1.467Y_1 - 4.613Y_4 + 3.039Y_5$
$V_3 = 3.014Y_1 - 1.429Y_4 - 0.835Y_5$

where U_r and V_r $(r = 1, 2, 3)$ respectively are the linear combinations of the set of independent variables **X** (global climate elements) and the linear combinations of the set of dependent variables **Y** (rainfall in West Java regions). The pair of variables U_r and V_r form three canonical functions, namely

$$(U_1, V_1) \qquad (15)$$
$$(U_2, V_2) \qquad (16)$$
$$(U_3, V_3) \qquad (17)$$

Furthermore, based on Table 6, the first canonical correlation coefficient is 0.819, which means there is a real and strong relationship between canonical variables on canonical functions (15), while the second and third correlation coefficients are only 0.169 and 0.093.

The canonical correlation significance test results

1. Overall Canonical Correlation Test Results
   Overall canonical correlation test using equation (2) gives a result of $B = 118.884 > \chi^2_{(0.01;9)} = 21.67$, it means H_0 on the hypotheses test is rejected or at least there is one significant canonical correlation.

2. Individual Canonical Correlation Test Results
   Individual canonical correlation test using equation (3) gives the results: $B_1 = 118.884 > \chi^2_{(0.01;9)} = 21.67$ or canonical correlation of the function (15) is significant, $B_2 = 3.898 < \chi^2_{(0.01;4)} = 13.28$ or canonical correlation of the function (16) is not significant, and $B_3 = 0.899 < \chi^2_{(0.01;1)} = 6.64$ or canonical correlation of the function (17) is not significant. Thus, further analysis is done only on function (15).

Redundancy analysis results

Table 7. The Proportion of Variance ***Y*** and ***X*** Explained by U and V

	$\boldsymbol{U_1}$	$\boldsymbol{V_1}$
Set of The Dependent Variables	0.471	0.702
	$\boldsymbol{U_1}$	$\boldsymbol{V_1}$
Set of The Independent Variables	0.346	0.232

Based on Table 7, through calculations with equations (6) and (7), the proportion of variance of the set of dependent variables (rainfall in West Java regions) that can be explained by the set of independent variables (global climate elements) is 0.3857 or 38.57 % and the proportion of variance of the set of independent variables (global climate elements) that can be explained by the set of independent variables (rainfall in West Java regions) is 0.19 or 19%.

Table 8. Redundancy Coefficient (R^2)

	R^2
1	0.670761
2	0.028561
3	0.008649

Based on Table 8, the redundancy coefficient obtained is 0.670761, it means that the canonical correlation of function (15) can explain the relationship between the **X** and **Y** of 67.08%.

The results of canonical variables interpretation of data

Based on the previous canonical correlation significance test, the results obtained indicate that the significant correlation is only the canonical correlation of function (15). Therefore, interpretation of canonical variables is only done on functions (15).

1. Canonical Weights

Table 9. The Canonical Weight of The Set of Dependent Variables

Variable	Function (15)	Function (16)	Function (17)
Y_1	−0.231	0.936	1.923
Y_4	−0.051	−2.963	−0.918
Y_5	1.212	1.875	−0.515

Table 10. The Canonical Weight of The Set of Independent Variables

Variable	Function (15)	Function (16)	Function (17)
X_1	0.048	0.234	1.035
X_2	−0.113	0.903	−0.536
X_3	0.991	0.195	0.032

Based on Table 9 and Table 10, the canonical variable in function (15) becomes

$U_1 = 0.048X_1 - 0.113X_2 + 0.991X_3$

$V_1 = -0.231Y_1 - 0.051Y_4 + 1.212Y_5$

In the canonical variable U_1, the order of relative contributions from the largest to the smallest of the original variables is X_3 (AUSMI), X_2 (DMI), and X_1 (Nino3.4 index). In the canonical variable V_1, the order of relative contributions from the largest to the smallest of the original variables is Y_5 (Kuningan Regency rainfall index), Y_1 (Bogor Regency rainfall index), dan Y_4 (Garut Regency rainfall index).

2. Canonical Loadings

Table 11. The Canonical Loadings of The Set of Dependent Variables

Variable	Function (15)	Function (16)	Function (17)
Y_1	0.662	−0.232	0.713
Y_4	0.833	−0.451	0.319
Y_5	0.987	−0.063	0.149

Table 12. The Canonical Loadings of The Set of Inependent Variables

Variable	Function (15)	Function (16)	Function (17)
X_1	−0.125	0.495	0.860
X_2	−0.183	0.961	−0.209
X_3	0.994	0.085	−0.065

Based on Table 11 and Table 12, the canonical variable in function (15) becomes

$U_1 = -0.125X_1 - 0.183X_2 + 0.994X_3$

$V_1 = 0.662Y_1 + 0.833Y_4 + 0.987Y_5$

Canonical loadings states the correlation of the origin variable with its canonical variable. In the canonical variable U_1, the origin variable that has the strongest relationship is X_3 (AUSMI). In the canonical variable V_1, the origin variable that has the strongest relationship is Y_5 (Kuningan Regency rainfall index).

3. Canonical Cross-Loadings

Table 13. The Canonical Cross-Loadings of The Set of Dependent Variables

Variable	Function (15)	Function (16)	Function (17)
Y_1	0.542	−0.039	0.066
Y_4	0.682	−0.076	0.030
Y_5	0.808	−0.011	0.014

Table 14. The Canonical Cross-Loadings of The Set of Independent Variables

Variable	Function (15)	Function (16)	Function (17)
X_1	−0.102	0.084	0.080
X_2	−0.150	0.162	−0.019
X_3	0.814	0.014	−0.006

Based on Table 13 and Table 14, the canonical variable in function (15) becomes

$$U_1 = -0.102X_1 - 0.15X_2 + 0.814X_3$$
$$V_1 = 0.542Y_1 + 0.682Y_4 + 0.808Y_5$$

Canonical cross-loadings states the correlation of the origin variable in a canonical variable with its other canonical variables. In the canonical variable U_1, the origin variable that has the strongest relationship with the canonical variable V_1 is X_3 (AUSMI) and in the canonical variable V_1, the origin variable that has the strongest relationship with the canonical variable U_1 is Y_5 (Kuningan Regency rainfall index). It is because the variables X_3 dan Y_5 have the greatest coefficient values.

CONCLUSIONS AND SUGGESTIONS

Based on the results and discussion in this study, the conclusions obtained are by using Canonical Analysis, it can be explained that there is a simultaneous relationship between the global climate elements, i.e., ENSO, IOD, and the Monsoon Asia-Australia with rainfall in the West Java regions, i.e., Bogor Regency, Bandung Regency, Sukabumi Regency, Garut Regency, and Kuningan Regency by 0.819. In this relationship, the global climate elements and rainfall in the West Java regions which had the most powerful influence were the Monsoon Asia-Australia and the Kuningan Regency rainfall. This result can be used as a recommendation for relevant instances such as BMKG in the prediction of climate phenomena, especially in the West Java regions.

Based on the results and discussion in this study, the authors suggest for further research to simultaneously test linearity and heteroscedasticity using the MANOVA (Multivariate Analysis of Variance) approach.

ACKNOWLEDGMENT

The authors thank Rector Universitas Padjadjaran for supporting the research through Academic Leadership Grant 2019 with contract number: 3330/UN6.D/ LT/2019. Also thanks to the anonymous reviewers for valuable review of this paper.

REFERENCES

Aldrian, E., M. Karmini & Budiman. (2011). *Adaptasi dan Mitigasi Perubahan Iklim di Indonesia*. Jakarta: Pusat Perubahan Iklim dan Kualitas Udara Kedeputian Bidang Klimatologi, BMKG.

Chaghooshi, A.J., M.S. Neshan & M.M. Moghadam. (2015). Canonical Correlation Analysis Between Supply Chain Qualitu Management and Competitive Advantages. *Foundations of Management, 7,* 83-92.

Hair, J.F., Jr.W.C. Black, B.J. Babin & R.E. Anderson. (2009). *Multivariate Data Analysis. (7th ed.)*. New York: Prentice Hall International, Inc.

Hotelling, H. (1936). Relations Between Two Sets of Variates. *Biometrika*, *28* (3-4), 321-377.

Irianingsih, I., N. Gusriani, S. Kulsum & K. Parmikanti. (2016). Analisis Korelasi Kanonik Perilaku Belajar terhadap Prestasi Belajar Siswa SMP (Studi Kasus Siswa SMPN 1 Sukasari Purwakarta). *Prosiding Seminar Matematika dan Pendidikan Matematika*, 693-703.

Johnson, R.A. & D.W. Wichern. (1998). *Applied Multivariate Statistical Analysis. (4th ed.)*. New York: Prentice Hall.

Ningsih, B.H.S & W.W. Putranto. (2019). Dampak Kejadian La Niña terhadap Curah Hujan di Kabupaten Jembrana. *Kappa Journal, 3,* 18-30.

Qiu, Q. K. Shi, X.J. Qiao & K. Jiang. (2016). Determining the Dominant Environmental Parameters for Greenhouse Tomato Seedling Growth Modeling Using Canonical Correlation Analysis. *IFAC-PapersOnLine, 49* (16), 387-391.

Rencher, A.C. (1998). *Multivariate Statistical Inference and Applications*. New York: John Wiley & Sons, Inc.

Ridwan, M. (2019). Prakiraan Musim Kemarau Tahun 2019 di Indonesia. https://www.bmkg.go.id/berita/?p=prakiraan-musim-kemarau-tahun-2019-di-indonesia&lang=ID&s=detil. (Diakses pada tanggal 14 Juli 2019 20:13).

Rustiana, S., B.N. Ruchjana, A.S. Abdullah, E. Hermawan, S.B. Sipayug, I.G.N.M. Jaya & Krismianto. (2017). Rainfall Prediction of Cimanuk Watershed Regions with Canonical Correlation Analysis (CCA). *J. Phys.: Conf. Ser. 893 012021.*

Walpole, R.E., R.H. Myers, S.L. Myers & K. Ye. (2011). *Probablility & Statistics for Engineers & Scientics (9th ed.)*. Boston: Prentice Hall.

Whittaker, J., C. Whitehead & M. Somers. (2005). The Neglog Transformation and Quantile Regression for the Analysis of a Large Credit Scoring Database. *Applied Statistics*, *54* (5), 863-878.

Yudiaatmaja, F. (2013). *Analisis Regresi dengan Menggunakan Aplikasi Komputer Statistik SPSS*. Jakarta: Gramedia Pustaka Utama.