Vol. 15, No. 3 (2026): 1241 - 1253 http://dx.doi.org/10.23960/jtepl.v15i3.1241-1253

JURNAL TEKNIK PERTANIAN LAMPUNG

ISSN 2302-559X (print) / 2549-0818 (online)
Journal homepage : https://jurnal.fp.unila.ac.id/index.php/JTP

Rainfall Prediction with Radial Basis Function Neural Network and Its Correlation with Bird's Eye Chili (*Capsicum frutescens*) Production in Sawangan Subdistrict Magelang Regency

Azka Sinatrya[1], Bayu Dwi Apri Nugroho[1,✉], Chandra Setyawan[1]

[1] Department of Agricultural and Biosystem Engineering, Universitas Gadjah Mada, Yogyakarta, INDONESIA.

Article History:

Received : 13 October 2025
Revised : 18 May 2026
Accepted : 20 May 2026

Keywords:

Bird's eye chili,
Global climate index,
Modeling,
PCA,
Rainfall.

Corresponding Author:
✉ bayu.tep@ugm.ac.id
(Bayu Dwi Apri Nugroho)

ABSTRACT

Climate change causes rainfall anomalies that directly impact the decline in horticultural crop productivity, particularly bird's eye chili (Capsicum frutescens). This study aimed to analyze the effect of global rainfall indices in Sawangan Subdistrict through the development a prediction model. Modeling was performed using Radial Basis Function Neural Network (RBFNN) method with Principal Component Analysis (PCA) integration to simplify the climate index input variables. Model accuracy was evaluated using the Root Mean Squared Error (RMSE), Nash-Sutcliffe Efficiency (NSE), and R^2. Furthermore, the prediction results were correlated with chili production data to test the relevance of the model to actual conditions. Results showed that model configurations provide varying performance. The best model based on evaluation is the model in the 15-year range using PCA 3 global climate indices and training percentage of 90% (RMSE: 101.39; NSE: 0.7268). However, for validation and correlation with production, it was found that the 15-year range using PCA 5 global climate indices and training percentage of 70% was the best model with highest R^2 value of 0.8572 and correlation value close to actual value. Variations in data period, number of climate indices, and training data proportion affect model performance. Adding data volume and variable complexity does not always improve accuracy, so it is necessary to identify the optimum point to get the most reliable prediction model.

1. INTRODUCTION

As an agrarian country, Indonesia has a large portion of its population that depends on the agricultural sector for their live (Faisal *et al*., 2020; Moeis *et al*., 2020; Santoso & Nasir, 2021). Agricultural activities are greatly influenced by climatic characteristics, such as temperature, season, and especially rainfall. These factors are crucial in determining whether a location is suitable for growing crops (Negara *et al*., 2018). Rainfall is vital because it fulfills the water needs of plants. However, the average annual rainfall in Indonesia varies, these fluctuations depend on the climate and geographical factors of the local area (Duffy *et al*., 2021). In addition, global climate change is also a trigger for erratic rainfall patterns.

Since the industrial revolution, human activities have caused a significant increase in greenhouse gas (GHG) production (Pais *et al*., 2020). The release of these gases into the atmosphere is a major trigger of climate change, the effects of which are clearly visible in changing climate patterns (Azzahra *et al*., 2024; Fawzy *et al*., 2020). Climate change has serious consequences for the agricultural sector, including shifts in planting seasons, changes in the spread of pests and plant diseases, and incompatibility or changes in the suitability of certain crop types with their specific agroecological zones (Edame *et al*., 2011 in Kogo *et al*., 2021).

Climate change, especially in horticultural crops, poses various serious challenges, such as increased salinity problems, droughts, and increased pressure from pests and diseases (Agadinansyah & Suciati, 2023; Lastochkina *et al*., 2022). Salinity and drough can inhibit plant growth and reduce water and nutrient uptake, while the increasing spread of pests and diseases may futher decrease crop yield and quality. These conditions have a direct negative impact on horticultural production.

One horticultural crop that is highly sensitive to climate change is chili plants. For chili plants, adequate water supply is crucial. Water shortages can cause chili plants to wilt, produce small fruits, and even lead to plant death (Padmaningrum *et al*., 2022). Chili peppers themselves are a major commodity with high economic value in Indonesia (Polii *et al*., 2019; Sari *et al*., 2019; Ziaulhaq & Amalia, 2022). Therefore, the surge in chili pepper prices caused by production problems has the potential to affect the inflation rate and national economic stability (Fahmi *et al*., 2023).

Chili peppers are one of the leading horticultural commodities in Magelang District, particularly in Sawangan Sub-district. The sub-district is located on the slopes of Mount Merapi, Indonesia, where the cool climate, fertile volcanic soil, and suitable environmental conditions support chili cultivation, especially bird's eye chili. These favorable conditions contribute to high chili productivity in the area. According to data from the Statistics Agency of Magelang District, bird's eye chili production in Sawangan Sub-district reached 42,975 quintals in 2024 (BPS, 2024).

Global climate indexes are often linked to climate change dynamics because they provide a comprehensive view of the elements that influence weather and climate conditions on a broader scale. Indexes such as the Southern Oscillation Index (SOI) and sea surface temperature (SST) are commonly used tools for identifying climate change patterns. Variations recorded in these pressure and temperature indexes can provide early signals of major climate phenomena, such as El Nino and La Nina, which directly affect global weather patterns (Dabanli *et al*., 2021; Zhang *et al*., 2023). By monitoring trends in these global climate indexes, experts can predict and gain a deeper understanding of ongoing climate change. This makes climate index data an important reference in formulating policies and decision-making related to climate mitigation and adaptation efforts.

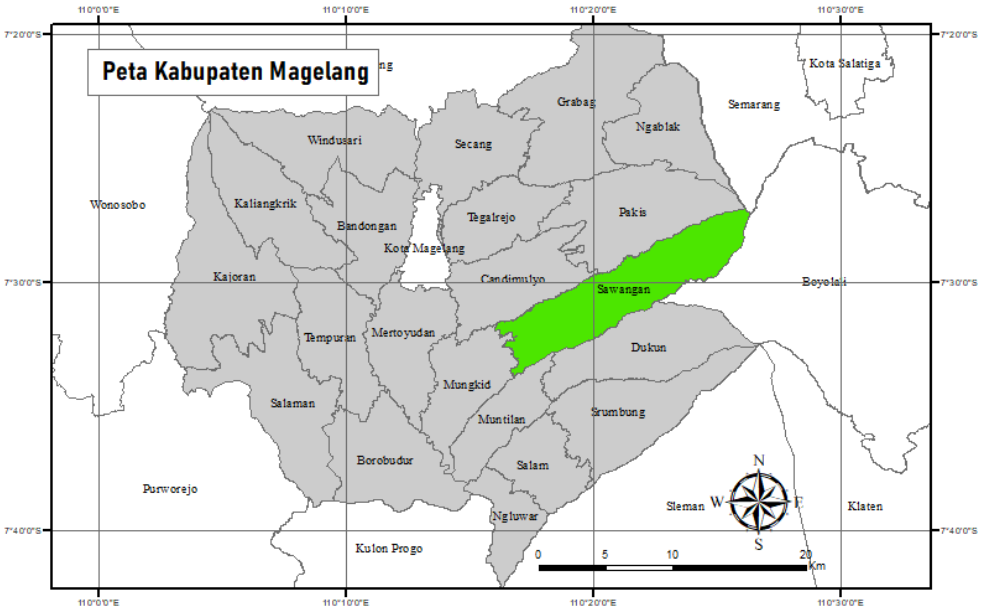
The issue of climate change requires immediate anticipation to address issues related to the availability and price of chili commodities. One effort in this anticipation is through a rainfall prediction model. In developing the model, a method using artificial neural networks can be an alternative. When compared to the Numerical Weather Prediction (NWP) technique, rainfall prediction using Artificial Neural Networks (ANN) shows significant advantages (Thakur *et al*., 2021). One type of ANN is the Radial Basis Function Neural Network (RBFNN), which, according to Liao & Li (2024), is superior to the Back Propagation Neural Network (BPNN) model and interpolation methods.

Many previous studies have used ANN models for rainfall prediction. However, studies combining PCA-based global climate indices with RBFNN and relating the prediction results to bird's eye chili production at the district level are still limited. Therefore, this study developed an RBFNN model using PCA-reduced climate indices and examined its relationship with chili production performance. Prediction models that more closely resemble actual conditions will greatly facilitate the design of agricultural management strategies. These prediction models were developed with the primary objective of reducing the risks associated with decision-making processes. These risks are related to various factors that affect crop yields, both in terms of quantity and quality of agricultural products (Piekutowska *et al*., 2021). This study aims to develop an RBFNN-based rainfall prediction model using PCA-reduced global climate indices and to evaluate its relationship with bird's eye chili productivity in Sawangan Sub-district.

2. MATERIALS AND METHODS

2.1. Research Location and Materials

The location selected for this study was Sawangan Subdistrict, Magelang Regency, Central Java (Figure 1). The materials used in this study included rainfall data for 2004-2024 for Sawangan Subdistrict, Magelang Regency; Central Java, Southern Oscillation Index (SOI) data for 2004-2024; Anomaly of Sea Surface Temperature (ASST) data consisting of Nino 3, Nino West, Nino 3.4, and Nino 4 from 2004 to 2024; and data on bird's eye chili production from 2020 to 2024 in Sawangan District, Magelang Regency, Central Java. The tools used include a laptop, Microsoft Excel software, Minitab software, and SPSS software.



(Source: Author)

Figure 1. Map of Magelang Regency

2.2. Research Procedures

2.2.1. Data Collection

The research procedure began with a data collection stage that focused on secondary data. Information on specific rainfall for the Sawangan District was obtained from data provided by the BMKG (Meteorology, Climatology, and Geophysics Agency). Meanwhile, data related to global climate indexes was obtained from international sources, namely the websites of the Japan Meteorological Agency (JMA) and the National Oceanic and Atmospheric Administration (NOAA) of the United States. Data on chili production was sourced from the Magelang Regency Agricultural Office.

2.2.2. Preliminary Analysis

The next step is preliminary analysis by evaluating the consistency and homogeneity of the rainfall data that has been collected. To test the consistency of the data, the Rescaled Adjusted Partial Sum (RAPS) method is used, which is a method that does not depend on references from other station data (Yuliana *et al*., 2024). Meanwhile, to ensure homogeneity, in example, the absence of significant changes over time, testing is carried out using the Standard Normal Homogeneity Test (SNHT) (Alexandersson, 1986 in Kabbilawsh *et al*., 2024).

RAPS method:

$$S_0^* = 0 \tag{1}$$

$$S_k^* = \sum_{i=1}^{k}(Y_i - \bar{Y}), \text{ with k = 1,2, … , n} \tag{2}$$

$$S_k^{**} = \frac{S_k^*}{D_y}, \text{ with k = 1,2, … , n} \tag{3}$$

$$D_y^2 = \sum_{i=1}^{n}\frac{(Y_i - \bar{Y})^2}{n} \tag{4}$$

$$Q = \max|S_k^{**}|, 0 \le k \le n \tag{5}$$

$$R = \max S_k^{**} - \min S_k^{**}, 0 \le k \le n \tag{6}$$

SNHT method:

$$T_k = k\overline{z_1^2} + (n-k)\overline{z_2^2} \quad k = 1,2,3,4 \ldots n \tag{7}$$

where,

$$\bar{z_1} = \frac{1}{k}\sum_{i=1}^{k}(X_i - X_{mean})/SD \tag{8}$$

$$\bar{z_2} = \frac{1}{n-k}\sum_{i=1+k}^{n}(X_i - X_{mean})/SD \tag{9}$$

Another initial analysis step was to observe the correlation or relationship between the various available global climate indexes. Furthermore, to simplify and reduce the complexity of the various climate index data, both in terms of units and measurement locations, the Principal Component Analysis (PCA) technique was used (Bhargawa & Singh, 2021; Wu, 2021) using SPSS software.

Standardization:

$$\text{Mean} = \frac{\Sigma x}{n} \tag{11}$$

$$\text{Standard deviation} = \sqrt{\frac{\Sigma(x-mean)^2}{n-1}} \tag{12}$$

$$\text{Standardization} = \frac{value-mean}{standard\ deviation} \tag{13}$$

Calculating the covariance matrix:

$$\text{Covariance} = \frac{\Sigma(\text{x}-(\text{mean of x}))(\text{y}-(\text{mean of y}))}{number\ of\ data\ point} \tag{14}$$

$$Covarian\ matrix = \begin{bmatrix} cov\ (x,x) & cov\ (x,y) \\ cov(y,x) & cov(y,y) \end{bmatrix} \tag{15}$$

Finding the eigenvectors:

$$Eigen\ value\ (\lambda) : \lambda I - A = 0 \tag{16}$$

$$Eigen\ vektor(v) : (\lambda I - A)v = 0 \tag{17}$$

Final dataset:

$$Final\ dataset = Standardization\ Data\ Set\ \times Feature\ Vector \tag{18}$$

2.2.3. Model Creation

The model used is a Radial Basis Function Neural Network (RBFNN). In an RBFNN, the number of neurons in the hidden layer can be adjusted according to specific requirements. RBFNNs feature rapid learning speeds and possess strong approximation capabilities.

The RBFNN model was constructed with an input layer architecture such as PC1 values and monthly rainfall with outliers removed and the number of lags determined based on the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF), a radial basis layer by determining the number of nodes, and an output layer for rainfall prediction. The variations made were the time period (10, 15, and 20 years), the number of climate indexes reduced (3 and 5 indexes), and the training data (70%, 80%, and 90%). The best model was selected based on the lowest RMSE, highest NSE and R^2 values. The RBFNN model development and analysis were performed using SPSS.

2.2.4. Evaluation and Analysis of Correlation with Chili Production

Model performance was evaluated using Root Mean Squared Error (RMSE) and Nash Sutcliffe Efficiency (NSE) (Susanto *et al*., 2022; Wang & Lu, 2018 in Saidah, H., 2021) using data from years not included in the training process. In model validation, the R Square value was observed using predicted and actual rainfall data in 2024.

Finally, the monthly rainfall values, both predicted and actual, were correlated with chili production values, and the values closest to actual conditions were observed.

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}(Xi-Yi)^2}{N}} \tag{19}$$

$$\text{NSE} = 1 - \frac{\sum_{i=1}^{N}(Xi-Yi)^2}{\sum_{i=1}^{N}(Xi-\overline{Xi})^2} \tag{20}$$

$$R^2 = 1 - \frac{\sum(Y_i-\hat{Y}_i)^2}{\sum(Y_i-\bar{Y})^2} \tag{21}$$

3. RESULTS AND DISCUSSION

3.1. Rainfall Consistency and Homogeneity Test

This study uses rainfall data covering a period of 21 years, from 2004 to 2024. The data used is monthly data, so the total data sample consists of 252 observations. In this analysis, data consistency is determined based on the lowest reference value in Table 1. The null hypothesis indicates that the series is consistent, and the alternative hypothesis indicates that the data series is not homogeneous. If the calculated value is greater than the table value, the null hypothesis is rejected, or the rainfall data is inconsistent. Conversely, if the calculated value is less than the table value, the null hypothesis is accepted, or the rainfall data is consistent.

Based on the data presented in Table 2, the calculation results show that the calculated values do not exceed the critical values (table values) at both significance levels, both at α=1% and α=5%. This finding indicates that the rainfall data analyzed has consistent characteristics throughout the observation period. Therefore, the rainfall data is considered to have adequate reliability and integrity, making it representative and suitable for further use in the modeling process.

Table 1. $Q/n^{1/2}$ and $R/n^{1/2}$ value tables

N	$Q/n^{1/2}$			$R/n^{1/2}$		
	90%	95%	99%	90%	95%	99%
10	1.05	1.14	1.29	1.21	1.28	1.38
20	1.10	1.22	1.42	1.34	1.43	1.60
30	1.12	1.24	1.46	1.40	1.50	1.70
40	1.13	1.26	1.50	1.42	1.53	1.74
50	1.14	1.27	1.52	1.44	1.55	1.78
100	1.17	1.29	1.55	1.50	1.62	1.86
□	1.22	1.36	1.63	1.62	1.75	2.00

Table 2. Results of the RAPS rainfall test

	Table (α=1%)	Table (α=5%)	Count	Description
$Q/n^{1/2}$	1.63	1.36	0.2052	Consistent
$R/n^{1/2}$	2.00	1.75	0.2820	Consistent

Table 3. Results of the SNHT rainfall test

T0	T	*p*-value (Two-tailed)	Alpha
5.926	250	0.289	0.05

Note: The *p*-value has been computed using 10000 Monte Carlo simulations.

Table 3 contains a summary of the results of the Standard Normal Homogeneity Test (SNHT) applied to rainfall data. In this test, the null hypothesis (H_0) states that the data is homogeneous, while the alternative hypothesis (H_1) indicates a change or break in the data series. The null hypothesis (H_0) is rejected if the p-value is less than or equal to the α value (significance level), and the null hypothesis (H_0) is not rejected (accepted) if the *p*-value is greater than the

α value. Based on the test results, the *p*-value obtained is 0.289, which is greater than the α value of 0.05 (5% significance level). Thus, the null hypothesis (H_0) fails to be rejected, so the rainfall data used has been verified as homogeneous and meets the requirements to be included in the subsequent modeling process.

3.2. Global Climate Index Analysis

3.2.1. Global Climate Index Correlation Analysis

The strength of a correlation is determined by how close its value is to ±1; values close to zero indicate a weak relationship. A negative sign (-) in the correlation coefficient indicates an inverse relationship. The global climate index correlation matrix (Table 4) shows a very strong positive correlation (value >0.800) between Nino3.4, Nino3, and Nino4. In contrast, Nino West shows a correlation ranging from weak to moderate and is inversely related to other Nino indexes, but has a direct relationship with SOI. The SOI correlation (Table 4) shows a fairly strong and inverse trend with ASST Nino3.4, Nino3, and Nino4. Based on the correlation, it is recommended to initiate PCA using two variations of data sets, namely with 5 climate indexes and 3 climate indexes (Nino3.4, Nino3, and Nino4).

Table 4. Correlation matrix of global climate index for 2004–2024

	Nino3.4	Nino3	Nino4	NinoWest	SOI
Nino3.4	1.000	.949	.919	-.409	-.698
Nino3	.949	1.000	.805	-.351	-.631
Nino4	.919	.805	1.000	-.345	-.726
NinoWest	-.409	-.351	-.345	1.000	.473
SOI	-.698	-.631	-.726	.473	1.000

3.2.2. Principal Component Analysis (PCA) of the Global Climate Index

The analytical output of the PCA process is the cumulative proportion represented by the principal components generated for various time periods, namely 10 years (2014-2023), 15 years (2009-2023), and 20 years (2004-2023). The results are presented in Table 5 for five global climate indexes and Table 6 for three global climate indexes.

Table 5. PCA results percentage for 5 global climate indexes

Component	Initial Eigenvalues Cumulative (%)		
	2014 -2023 (10 Years)	2009 – 2023 (15 Years)	2004 – 2023 (20 Years)
1	76.29	75.45	72.23
2	88.98	89.51	88.44
3	96.47	96.63	96.19
4	99.69	99.68	99.65
5	100.00	100.00	100.00

Table 6. PCA results percentage for 3 global climate indexes

Component	Initial Eigenvalues Cumulative (%)		
	2014 -2023 (10 Years)	2009 - 2023 (15 Years)	2004 – 2023 (20 Years)
1	93.75	93.68	92.80
2	99.47	99.40	99.35
3	100.00	100.00	100.00

The cumulative eigenvalue must reach above 70% to be considered capable of adequately explaining the total variance of the data (Uddin *et al*., 2019 in Azam & Rahman, 2022). When the five global climate indexes are considered, the cumulative percentage of variance explained by the first principal component in the three time variations exceeds 70%. This finding indicates that a single component effectively summarizes the majority (more than 70%) of the overall variability found in the five indexes. More optimal results were obtained from the analysis of

three strongly correlated global climate indexes (Nino3.4, Nino3, and Nino4). In this scenario, the first principal component successfully explained a much higher proportion of variance, namely above 90%. This high figure confirms that most of the variability in the three indexes is highly correlated and can be reduced and represented very well using only one principal component.

3.3. Determining Lag and Outliers

With Minitab, ACF, PACF, and outlier analyses can be performed. To determine significant lags on the ACF and PACF plots, a 5% significance limit is used, marked by two red lines. An autocorrelation value at a certain lag is considered significant if the autocorrelation bar exceeds the upper or lower limit of the significance zone. If this occurs, there is an indication that the data value at the current time period is significantly influenced by the value at the previous time period. Outliers in the data are identified using boxplot visualization by looking at points outside the interquartile range.

3.3.1. Determination of Rainfall Lag and Outliers

The analysis results show that there are no outliers (see Table 7) that identify extreme rainfall values in various periods, so there is no data that needs to be removed from the rainfall row and other factors (PC1). When viewed from the same lag between ACF and PACF (see Table 7), values from 1, 6, and 17 months ago affect current rainfall values for a 10 years period. For the 15 years period, values from 1, 4, 6, and 17 months prior influence current rainfall values. Meanwhile, for the 20 years period, values from 1, 4, 5, 6, 11, and 12 months prior influence current rainfall values. These values are used as modeling inputs in each period.

Table 7. Results of ACF and PACF lag for rainfall

Period	Lag ACF	Lag PACF	The Same Lag	Outlier
10 Years (2014-2023)	1, 2, 5, 6, 7, 11, 12, 13, 17, 18	1, 3, 4, 6, 17	1, 6, 17	-
15 Years (2009-2023)	1, 2, 4, 5, 6, 7, 8, 11, 12, 13, 17, 18, 19, 24, 29, 30	1, 3, 4, 6, 10, 17, 28	1, 4, 6, 17	-
20 Years (2004-2023)	1, 2, 4, 5, 6, 7, 8, 11, 12, 13, 17, 18, 19, 23, …	1, 3, 4, 5, 6, 10, 11, 12	1, 4, 5, 6, 11, 12	-

Table 8. ACF and PACF lag results for PC1 5 global climate indexes

Period	Lag ACF	Lag PACF	The Same Lag	Outlier
10 Years (2014-2023)	1, 2, 3, 4, 5	1, 2, 3, 4	1, 2, 3, 4	-
15 Years (2009-2023)	1, 2, 3, 4, 5	1, 2, 3	1, 2, 3	-
20 Years (2004-2023)	1, 2, 3, 4, 5, 6	1, 2, 3, 4, 10	1, 2, 3, 4	-

3.3.2. Determination of Lag and Outlier in the Global Climate Index

The analysis results show no outliers (see Table 8), which identifies the PC1 5 value of the global extreme climate index across various periods, meaning that no data needs to be removed from the PC1 5 global climate indexes and rainfall rows. When viewed from the same lag between ACF and PACF (see Table 8), the values in the previous 1, 2, 3, and 4 months affect the current PC1 5 global climate indexes value for the 10 years period. For the 15 years period, values from 1, 2, and 3 months prior influence the current PC1 5 global climate indexes value. Meanwhile, for the 20 years period, values from 1, 2, 3, and 4 months prior influence the current PC1 5 global climate indexes value. These values are used as modeling inputs in each period.

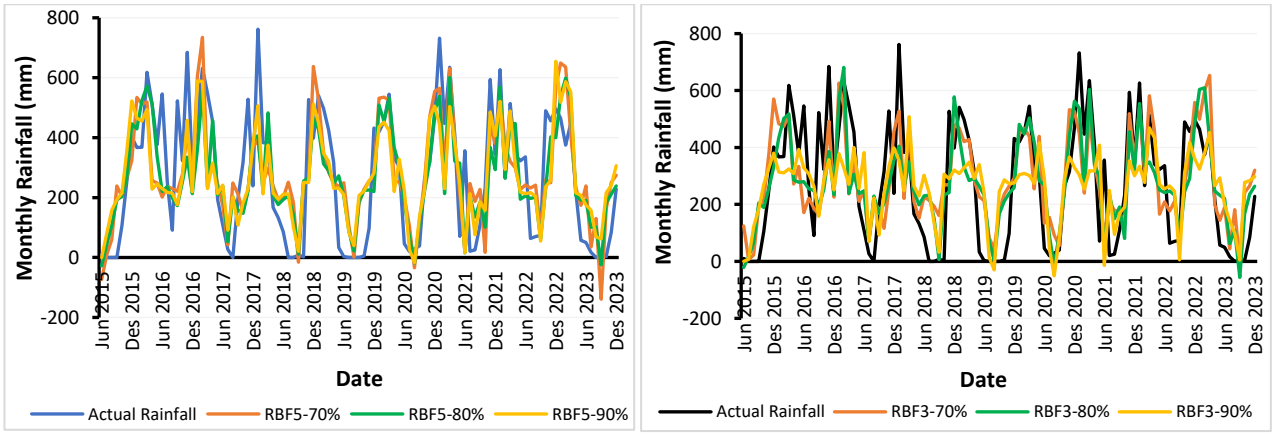
Outlier points (see Table 9) appear in the 15 years period, in row 83 (November 2015) and in the 20 years period in row 143 (November 2015). This indicates the need to remove these rows or months from PC1 3 global climate indexes and rainfall in modeling for these periods. When viewed from the same lag between ACF and PACF (see Table 9), the values in the previous 1 and 2 months affect the current value of PC1 3 global climate indexes for the 10 years period. For the 15 years period, the values from 1, 2, and 3 months prior influence the current value of PC1 3 of the global climate indexes. Meanwhile, for the 20 years period, the values from 1, 2, 3, and 4 months prior influence the current value. These values are used as modeling inputs in each period.

Table 9. ACF and PACF lag results for PC1 3 global climate indexes

Period	Lag ACF	Lag PACF	The Same Lag	Outlier
10 Years (2014-2023)	1, 2, 3, 4, 5	1, 2, 28	1, 2	-
15 Years (2009-2023)	1, 2, 3, 4, 5	1, 2, 3, 28	1, 2, 3	Row 83 (November 2015)
20 Years (2004-2023)	1, 2, 3, 4, 5, 6	1, 2, 3, 4, 28	1, 2, 3, 4	Row 143 (November 2015)

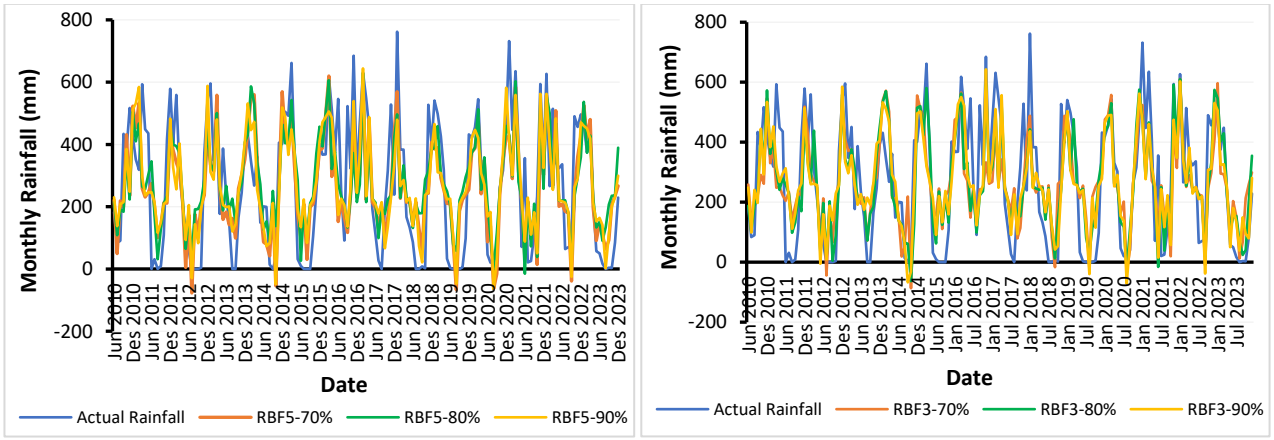
3.4. Modeling Using RBFNN

The comparison between actual rainfall and the RBFNN prediction results for all data periods (10-year, 15-year, and 20-year) showed generally similar performance patterns for both the five climate indices and the three PCA-based climate indices (Figure 3 to 5). Overall, the models were able to follow the general rainfall pattern, but limitations were still observed in capturing extreme rainfall fluctuations, particularly during periods of high rainfall where predicted values tended to be lower than the observed data. In contrast, during low rainfall periods, several models produced higher prediction values than the actual rainfall. These results indicate that some rainfall variability in the study area may not have been fully represented by the selected global climate indices and the RBFNN model configuration.



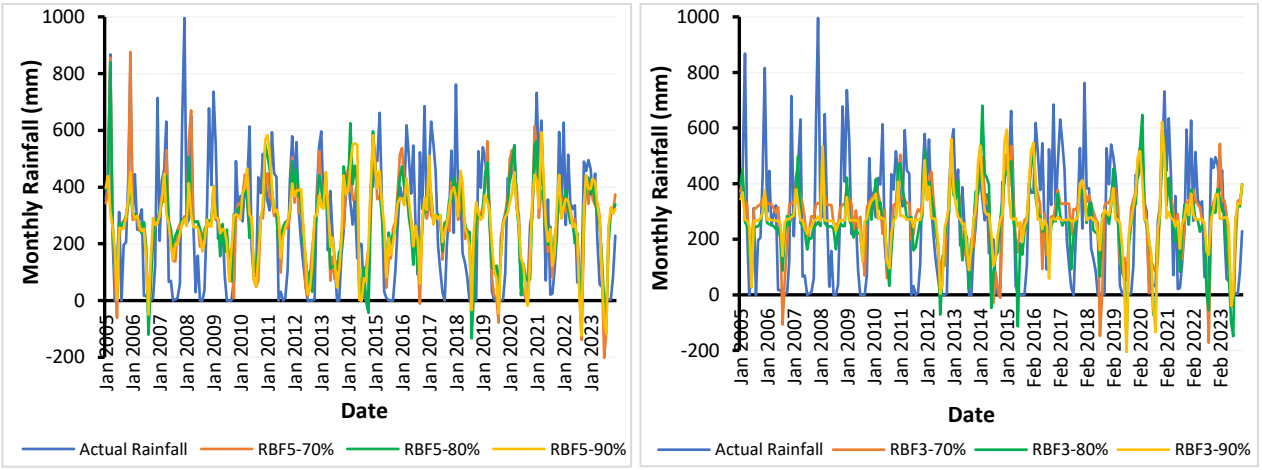
5 Climate Indexes 3 Climate Indexes

Figure 2. Actual rainfall vs. predictions for 2014-2023 (10-year period)



5 Climate Indexes 3 Climate Indexes

Figure 3. Actual rainfall vs. predictions for 2009-2023 (15-year period)



5 Climate Indexes 3 Climate Indexes

Figure 4. Actual rainfall vs. predictions for 2004-2023 (20-year period)

The prediction results across all data periods showed similar limitations in capturing extreme rainfall fluctuations. During periods of high rainfall, the predicted values generally tended to be lower than the observed rainfall, while higher predictions were often produced during low rainfall periods. This condition may indicate that short-term local climate variability was not fully represented by the selected global climate indices and the RBFNN model configuration.

Mean error (ME) was used to evaluate the average difference between observed and predicted rainfall values. Positive ME values indicate that the model tended to underestimate rainfall, while negative values indicate overestimation. For the 10-year period, the ME values for the models using five PCA-based climate indexes were -8.10, 0.95, and -3.80, while the models using three indexes produced values of -19.69, -5.98, and 8.41. In the 15-year period, the ME values for the five-index models were 8.60, -9.13, and -1.61, whereas the three-index models produced values of 3.60, 2.05, and 1.85. Meanwhile, in the 20-year period, the ME values for the five-index models were -1.81, -7.67, and -6.55, while the three-index models produced values of -13.27, 7.18, and -3.04. Overall, the models using three climate indexes tended to produce larger fluctuations in bias compared to the five-index models. This result indicates that reducing the number of climate indexes may limit the model's ability to capture rainfall variability, particularly under different training data proportions and longer observation periods.

To examine whether there were significant differences between the RBFNN models developed using three and five PCA-based climate indexes, the Wilcoxon signed-rank test was performed using SPSS for each training data proportion and time period. The results generally produced p-values greater than 0.05, indicating that the prediction performances of the compared models were statistically similar. A significant difference was only observed in the 10 years period with 80% training data, where the p-value obtained was 0.005. Overall, the reduction in the number of climate indexes did not substantially affect the prediction performance of the developed RBFNN models. This finding suggests that the PCA reduction process was able to retain the main climate variability information needed by the RBFNN model, even when fewer climate indexes were used. Therefore, the use of three PCA-based climate indexes may provide a simpler model configuration while maintaining comparable prediction performance.

Table 10. Results of Wilcoxon signed-rank test for RBFNN models using PCA 3 and PCA 5 indexes

Period	*p*-values		
	70%	80%	90%
10 years (2014-2023)	0.322	0.005	0.967
15 years (2009-2023)	0.464	0.814	0.078
20 years (2004-2023)	0.812	0.629	0.367

3.5. Sensitivity Testing and Validation

Based on the results presented in Table 11, it can be concluded that the Radial Basis Function Neural Network (RBFNN) model configuration has generally not achieved optimal performance across all variations of time periods and training data proportions explored. Model performance, measured in terms of accuracy and efficiency, is highly dependent on the interaction between the number of climate indices reduced through PCA and the proportion of data allocated for training. Given the complex dynamics over longer periods, a more adaptive modeling approach may be necessary. Nevertheless, overall, the best model performance in terms of highest accuracy (smallest RMSE) and highest efficiency (highest NSE) was found in the configuration with a 15 years time period (2009-2023), using PCA 3 Indexes, and a training data proportion of 90%.

Table 11. RMSE and NSE of prediction models

Period (Years)	Global Climate Index	RMSE			NSE		
		Training Percentage					
		70%	80%	90%	70%	80%	90%
10	PCA 5 Indexes	155.26	125.32	125.82	0.37236	0.55909	-1.19290
10	PCA 3 Indexes	141.96	128.56	150.28	0.55744	0.55769	-0.18228
15	PCA 5 Indexes	133.78	151.55	115.14	0.59121	0.44293	0.64764
15	PCA 3 Indexes	138.63	136.65	101.39	0.55826	0.57906	0.72678
20	PCA 5 Indexes	147.23	144.40	144.51	0.50412	0.51777	0.39151
20	PCA 3 Indexes	158.76	149.69	163.65	0.46046	0.47460	0.20257

Table 12. R square model prediction

Period (Years)	Global Climate Index	R^2		
		Training Percentage		
		70%	80%	90%
10	PCA 5 Indexes	0.5946	0.4518	0.7214
10	PCA 3 Indexes	0.8294	0.5186	0.5034
15	PCA 5 Indexes	0.8572	0.8056	0.5957
15	PCA 3 Indexes	0.7118	0.5983	0.6702
20	PCA 5 Indexes	0.4607	0.4726	0.6631
20	PCA 3 Indexes	0.5068	0.4389	0.4064

Table 12 presents the R^2 values obtained from model validation using rainfall data outside the training period. In several scenarios, models using five PCA-based climate indexes showed higher R^2 values than models using three indexes, indicating a better ability to explain rainfall variability. However, the overall prediction performance based on RMSE and NSE remained highest in the model using PCA 3 indexes with 90% training data. These results indicate that the optimal model configuration depends on the evaluation criteria used, where RMSE and NSE were considered the main indicators for selecting the final prediction model.

3.6. Correlation with Chili Pepper Production

Overall (Figure 5), the correlation between bird's eye chili production and actual rainfall and modeled rainfall tends to be weak. The weak correlation indicates that chili productivity is influenced by multiple factors, not only rainfall and global climate indices. Local climatic conditions, such as temperature and humidity, as well as agricultural management practices, may also affect productivity. Therefore, additional climate variables could improve the model performance and better explain productivity variations. In addition, the production dataset used in this study was limited to the 2020–2024 period, resulting in a relatively small number of observations for correlation analysis. However, a comparative analysis between actual correlation and model correlation can still be performed. The highest correlation identified was the model with a 20 years period, using 5 global climate indexes, and 70% training data, with a correlation value of 0.3694. This value far exceeds the correlation produced by actual rainfall, which only reached 0.1768. This significant difference indicates the risk of overfitting in the modeling process. This suspicion of

overfitting is supported by the fact that the model has an R square value that is among the five lowest of all scenarios tested. Conversely, the correlation between bird's eye chili production and rainfall in the model that most closely approximates the actual correlation is the model with a 15 years period, using 5 climate indexes, and 70% training data. This model also has the highest R square value, indicating the best balance between prediction accuracy and representation of the actual relationship.

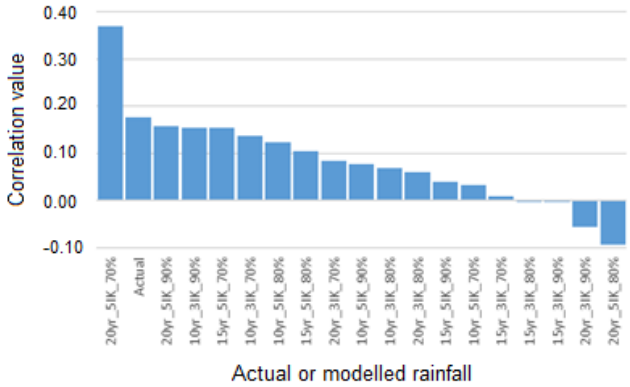


Figure 5. Correlation values of chili production vs. actual and modeled rainfall

4. CONCLUSION

The RBFNN model combined with PCA-based global climate indexes was able to represent rainfall patterns in Sawangan District under several modeling scenarios. Among the tested configurations, the model using the 15 years period with PCA 3 indexes and 90% training data produced the most consistent prediction performance. The results indicate that reducing the dimensionality of climate variables through PCA can still maintain the ability of the model to predict rainfall variability.

The prediction results may provide useful information for agricultural planning in areas vulnerable to climate variability, particularly for bird's eye chili cultivation. However, the relationship between rainfall prediction and chili productivity remained relatively weak, suggesting that crop productivity is influenced by many factors beyond rainfall conditions alone. This study was also limited by the relatively short production dataset available for productivity analysis. Future research should involve longer production records and additional environmental variables, such as temperature, humidity, soil conditions, and crop management factors, to improve the reliability and practical relevance of the prediction model.

AUTHOR CONTRIBUTION STATEMENT

Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
AS	✓	✓		✓	✓	✓		✓	✓	✓	✓			
BDAN	✓	✓		✓						✓		✓		
CS	✓	✓		✓						✓		✓		

C: Conceptualization	Fo: Formal Analysis	O: Writing - Original Draft	Fu: Funding Acquisition
M: Methodology	I: Investigation	E: Writing - Review & Editing	P: Project Administration
So: Software	D: Data Curation	Vi: Visualization	
Va: Validation	R: Resources	Su: Supervision	

ACKNOWLEDGMENTS

The author would like to thank the Central Java Meteorology, Climatology, and Geophysics Agency (BMKG) and the Magelang Regency Agricultural Office for their support in providing data. Thanks are also extended to the anonymous reviewers for their valuable comments.

REFERENCES

Agadinansyah, P.R. & Suciati, L.P. (2023). Adaptasi dan mitigasi petani cabai besar di Desa Sumberejo dalam menghadapi perubahan iklim. *Jurnal Pemikiran Masyarakat Ilmiah Berwawasan Agribisnis*, ***9***(2), 3016-3026. http://dx.doi.org/10.25157/ma.v9i2.10813

Azam, M.G. & Rahman, M.M. (2022). Assessing spatial vulnerability of bangladesh to climate change and extremes: A geographic information system approach. *Mitigation and Adaptation Strategies for Global Change*, ***27***(6), 38. https://doi.org/10.1007/s11027-022-10013-w

Azzahra, F., Samriana, S., & Ferdin, F. (2024). Pengaruh perubahan iklim terhadap pola hujan di Indonesia. *Sindoro Cendikia Pendidikan*, ***3***(1), 41-55.

Bhargawa, A. & Singh, A.K. (2021). Perceiving the trend of terrestrial climate change during the past 40 year (1978-2018). *Journal of Atmospheric Science Research*, ***4***(1), 1-15. https://doi.org/10.30564/jasr.v4i1.2488

BPS (Badan Pusat Statistik). (2024). *Produksi Tanaman Sayur-Sayuran Menurut Kecamatan di Kabupaten Magelang (Kuintal)*. Badan Pusat Statistik, Magelang.

Dabanli, I., Şişman, E., Güçlü, Y.S., Birpınar, M.E., & Şen, Z. (2021). Climate change impacts on sea surface temperature (SST) trend around Turkey seashores. *Acta Geophysica*, ***69***, 295–305. https://doi.org/10.1007/s11600-021-00544-2

Duffy, C., Toth, G.G., Hagan, R.P.O., McKeown, P.C., Rahman, S.A., Widyaningsih, Y., Sunderland, T.C.H., & Spillane, C. (2021). Agroforestry contributions to smallholder farmer food security in Indonesia. *Agroforestry Syst*ems, ***95***, 1109–1124. https://doi.org/10.1007/s10457-021-00632-8

Fahmi, M., Siregar, A., & Effendi, I. (2023). Analysis of the supply and needs of red chili in North Sumatra Province. *Jurnal Ekonomi*, ***12***(1), 596–602.

Faisal, H.N. (2020). Peran penyuluhan pertanian sebagai upaya peningkatan peran kelompok tani (Studi kasus di Kecamatan Kauman Kabupaten Tulungagung). *Jurnal Agribis*, ***6***(1), 1-13.

Fawzy, S., Osman, A.I., Doran, J., & Rooney, D.W. (2020). Strategies for mitigation of climate change: A review. *Environmental Chemistry Letters*, ***18***, 2069–2094. https://doi.org/10.1007/s10311-020-01059-w

Kabbilawsh, P., Kumar, D.S., & Chithra, N.R. (2024). Assessment of temporal homogeneity of long-term rainfall time-series datasets by applying classical homogeneity tests. *Environment, Development and Sustainability*, ***26***, 16757–16801. https://doi.org/10.1007/s10668-023-03310-0

Kogo, B.K., Kumar, L., & Koech, R. (2021). Climate change and variability in Kenya : A review of impacts on agriculture and food security. *Environment, Development and Sustainability*, ***23***, 23–43. https://doi.org/10.1007/s10668-020-00589-1

Lastochkina, O., Aliniaeifard, S., SeifiKalhor, M., Bosacchi, M., Maslennikova, D., & Lubyanova, A. (2022). Novel approaches for sustainable horticultural crop production: Advances and prospects. *Horticulturae*, ***8***(10), 910. https://doi.org/10.3390/horticulturae8100910

Liao, Z. & Li, M. (2024). Comparing rainfall prediction at various time scales and rainfall interpolation at the regional scale using artificial neural networks. *Theoretical and Applied Climatology*, ***155***, 9929–9940. https://doi.org/10.1007/s00704-024-05205-0

Moeis, F.R., Dartanto, T., Moeis, J.P., & Ikhsan, M. (2020). A longitudinal study of agriculture households in Indonesia: The effect of land and labor mobility on welfare and poverty dynamics. *World Development Perspectives*, ***20***, 100261. https://doi.org/10.1016/j.wdp.2020.100261

Negara, H.R.P., Irzani, I., & Ripai, R. (2018). Konstruksi model matematika pola curah hujan menggunakan artificial neural network (ANN) dengan metode backpropagation. *Jurnal Sains dan Teknologi*, ***1***(1), 10-18.

Padmaningrum, D., Suminah, S., Utami, B.W., Ihsaniyati, H.,& Widiyanti, E. (2022). Pemberdayaan kelompok tani melalui budidaya cabai sebagai upaya peningkatan pendapatan petani lahan kering di Kabupaten Sukoharjo. *Jurnal Pengabdian kepada Masyarakat*, ***13***(1), 158-167. https://doi.org/10.26877/e-dimas.v13i1.7001

Pais, I.P., Reboredo, F.H., Ramalho, J.C., Pessoa, M.F., Lidon, F.C., & Silva, M.M. (2020). Potential impacts of climate change on agriculture – A review. *Emirates Journal of Food and Agriculture*, ***32***(6), 397-407. https://doi.org/10.9755/ejfa.2020.v32.i6.2111

Piekutowska, M., Niedbała, G., Piskier, T., Lenartowicz, T., Pilarski, K., Wojciechowski, T., Pilarska, A.A., & Czechowska-Kosacka, A. (2021). The application of multiple linear regression and artificial neural network models for yield prediction of

very early potato cultivars before harvest. *Agronomy*, ***11***(5), 885. https://doi.org/10.3390/agronomy11050885

Polii, M.G.M., Sondakh, T.D., Raintung, J.S.M., Doodoh, B., & Titah, T. (2019). Kajian teknik budidaya tanaman cabai (*Capsicum annuum* L.) Kabupaten Minahasa Tenggara. *Eugenia*, ***25***(3), 73-77.

Saidah, H., Setiawan, A., Hanifah, L., Pradjoko, E., & Suroso, A. (2021). Koreksi bias data hujan luaran GCM ECHAM5 untuk prediksi curah hujan bulanan dan musiman Pulau Lombok. *Jurnal Sains Teknologi & Lingkungan*, ***7***(2), 209-219. https://doi.org/10.29303/jstl.v7i2.289

Santoso, A., & Nasir, M. (2021). Pemetaan lahan dan komoditas pertanian berbasis WebGIS di Kabupaten OKU Timur. *Jurnal Ilmiah Betrik*, ***12***(2), 129–139. https://doi.org/10.36050/betrik.v12i2.320

Sari, I., Yanti, N.D., & Hidayat, T. (2019). Faktor-faktor yang mempengaruhi usahatani cabai rawit (*Capsicum fretescens* L.) di Kabupaten Tabalong. *Frontier Agribisnis*, ***3***(4), 23-30. https://doi.org/10.20527/frontbiz.v3i4.1937

Susanto, A., Oetomo, W., & Wulandari, E. (2022). Analysis of satellite rain data usage on the rationalization activities of the rain post network (Case study: Rationalization of the jelai watershed rain post network). *Journal of Research and Community Service*, ***3***(14), 23-30. https://doi.org/10.36418/dev.v3i14.327

Thakur, N., Karmakar, S., & Soni, S. (2021). Rainfall forecasting using various artificial neural network techniques – A review. *International Journal of Scientific Research in Computer Science, Engineering, dan Information Technology*, ***7***(3), 506-526. https://doi.org/10.32628/CSEIT2173159

Wu, T. (2021). Quantifying coastal flood vulnerability for climate adaptation policy using principal component analysis. *Ecological Indicators*, ***129***, 108006. https://doi.org/10.1016/j.ecolind.2021.108006

Yuliana, A., Sujono, J., & Karlina. (2024). Analysis of extreme rainfall in the Mt. Merapi Area. *Journal of the Civil Engineering Forum*, ***10***(8), 73-84. https://doi.org/10.22146/jcef.10084

Zhang, L., Li, Y., Yu, S., & Wang, L. (2023). Risk transmission of el niño-induced climate change to regional green economy index. *Economic Analysis and Policy*, ***79***, 860-872. https://doi.org/10.1016/j.eap.2023.07.006

Ziaulhaq, W., & Amalia, D.R. (2022). Pelaksanaan budidaya cabai rawit sebagai kebutuhan pangan masyarakat. *Indonesian Journal of Agriculture and Environmental Analytics (IJAEA)*, ***1***(1), 27-36. https://doi.org/10.55927/ijaea.v1i1.812