CAUCHY – Jurnal Matematika Murni dan Aplikasi
Volume 11 (1) (2026), Pages 754-771
p-ISSN: 2086-0382; e-ISSN: 2477-3344

Short- and Long-Run Relationships Between Observed Rainfall and CMIP6 Projections Using the ARDL Approach: Evidence from Majalengka Regency, Indonesia

Sri Nurdiati*, Mohamad Khoirun Najib, and Fathia Rahmaisty

School of Data Science, Mathematics, and Informatics, IPB University, Bogor, Indonesia

Abstract

This study examines the short-run and long-run relationships between observed rainfall and precipitation projections from CMIP6 climate models in Majalengka Regency under three emission scenarios (SSP1–2.6, SSP2–4.5, and SSP5–8.5). The analysis employs the Autoregressive Distributed Lag (ARDL) framework to evaluate both dynamic adjustments and long-run equilibrium relationships between observed and modeled rainfall series. Stationarity was assessed using the Augmented Dickey–Fuller (ADF) test, and the bounds testing procedure was applied to examine the existence of long-run associations. The results indicate a stable adjustment mechanism between observed rainfall and CMIP6 precipitation projections. The estimated long-run precipitation coefficient is -0.561, indicating a weak negative long-term association between modeled and observed rainfall. The error correction term (ECT_{t-1}) is estimated at -0.762 and is statistically significant, suggesting that approximately 76.2% of deviations from the long-run equilibrium are corrected within one period. The best estimated model explains about 84.81% of the variation in observed rainfall. These findings suggest that CMIP6 rainfall projections exhibit a measurable statistical relationship with observed rainfall patterns; however, the results should be interpreted cautiously for long-term climate assessment. The analysis is based on a single observational station and does not incorporate bias correction or statistical downscaling, which may limit the representation of local-scale rainfall variability.

Keywords: ARDL bounds testing; Augmented Dickey–Fuller test; BMKG rainfall data; CMIP6 precipitation projections; cointegration analysis; Majalengka Regency

1. Introduction

Rainfall plays a critical role in hydrological processes [1, 2], agricultural productivity [3], and water resource management [4], particularly in tropical regions where rainfall variability is inherently high [5, 6]. In Indonesia, complex interactions between large-scale atmospheric circulation, regional monsoon systems, and local topography generate pronounced temporal and spatial variability in rainfall patterns [5, 7]. This variability poses significant challenges for reliable rainfall modeling and climate impact assessments, especially at local and regional scales where decision-making in agriculture [8, 9], disaster mitigation [10], and water management [11] is most sensitive to rainfall uncertainty.

*Corresponding author. E-mail: nurdiati@apps.ipb.ac.id

Submitted: February 01, 2026 Reviewed: February 22, 2026 Accepted: March 20, 2026
DOI: https://doi.org/10.18860/cauchy.v11i1.40902

To address these challenges, a wide range of statistical and time-series approaches has been applied to rainfall modeling. Conventional univariate models, such as Autoregressive Integrated Moving Average (ARIMA) [12, 13], have been widely used to capture historical rainfall dynamics and short-term persistence. While these models often perform well in reproducing past observations, they rely solely on internal lag structures and do not explicitly incorporate external climatic drivers. As a result, their ability to represent long-term relationships or to link observed rainfall with climate model outputs remains limited. Multivariate approaches, including multiple linear regression [14] and other empirical models using meteorological predictors [15], have also been explored. However, these methods typically assume static relationships and often fail to distinguish between short-run fluctuations and long-run equilibrium behavior.

Recent advances in climate science have provided extensive climate projection datasets through initiatives such as the Coupled Model Intercomparison Project Phase 6 (CMIP6) [16]. CMIP6 offers multi-scenario precipitation projections under different Shared Socioeconomic Pathways (SSPs) [17], enabling comprehensive assessments of future rainfall behavior under varying emission trajectories. Despite their global importance, CMIP6 outputs are characterized by coarse spatial resolution and structural uncertainty, which may limit their direct applicability at local scales [18, 19]. Consequently, evaluating how CMIP6 rainfall projections relate to observed rainfall—particularly in terms of short- and long-run dynamics—has become an important research frontier.

Analyzing the models separately for each SSP scenario is important because different emission pathways generate distinct climate forcing trajectories, which may influence precipitation variability and long-term rainfall patterns [20, 21]. Consequently, the dynamic relationship between observed rainfall and model-simulated precipitation may vary across scenarios. Estimating the ARDL model for each SSP scenario allows the study to examine whether the strength and adjustment dynamics of the rainfall relationship remain consistent under different climate projection pathways.

From a methodological perspective, the Autoregressive Distributed Lag (ARDL) framework has emerged as a powerful tool for analyzing dynamic relationships between time-series variables [22, 23]. Unlike traditional models, ARDL allows the simultaneous examination of short-run adjustments and long-run equilibrium relationships, even when variables exhibit different orders of integration. This flexibility has led to increasing applications of ARDL in climate [24], hydrology [25], and environmental studies [26]. However, most existing studies focus either on forecasting performance or on long-run associations alone, with limited attention to how climate model outputs and observations interact across different time horizons.

Despite the increasing use of CMIP6 climate projections in regional climate studies, limited research has examined the dynamic relationship between observed rainfall and CMIP6 precipitation outputs using econometric time-series approaches. Most previous studies focus on bias correction [27, 28], statistical downscaling [29], or model performance evaluation [30], while relatively few studies investigate the temporal interaction between observed and modeled precipitation using cointegration-based frameworks. This study contributes to the literature by applying the Autoregressive Distributed Lag (ARDL) approach to analyze both short-run dynamics and long-run equilibrium relationships between observed rainfall from BMKG and CMIP6 precipitation projections at a local station scale.

This gap is particularly evident in local-scale rainfall studies in Indonesia, where few investigations explicitly assess whether climate model projections capture short-term rainfall variability while potentially diverging in long-term behavior. Moreover, comparative analyses across multiple emission scenarios within a unified statistical framework remain scarce. Understanding these distinctions is essential, as misinterpretation of long-run relationships may lead to inappropriate conclusions regarding the reliability of climate projections for local applications.

Motivated by these gaps, this study aims to investigate the short- and long-run relationships between observed monthly rainfall and CMIP6 precipitation outputs in Majalengka Regency, Indonesia, using the ARDL approach. Specifically, this research (i) examines the existence of long-run cointegration between observed and modeled rainfall across different SSP scenarios, (ii) quantifies the contrasting short-run and long-run effects of CMIP6 projections on observed rainfall, and (iii) evaluates model performance under low-, medium-, and high-emission scenarios. By doing so, this study provides new insights into the applicability and limitations of climate model rainfall projections for local-scale rainfall modeling and climate impact analysis.

2. Study Area and Datasets

This section describes the geographical characteristics of the study area and the datasets used in the analysis, including both observational rainfall records and CMIP6 precipitation projections.

2.1. Majalengka Regency

Majalengka Regency is located in the eastern part of West Java Province, Indonesia, and represents a typical tropical monsoon climate region. Geographically, the regency lies between approximately 6.6°–7.4°S and 108.0°–108.4°E (Fig. 1), with elevation ranging from lowland plains to mountainous areas exceeding 1,000 m above sea level. This diverse topography plays an important role in shaping local rainfall patterns through orographic effects and land–atmosphere interactions [31, 32].

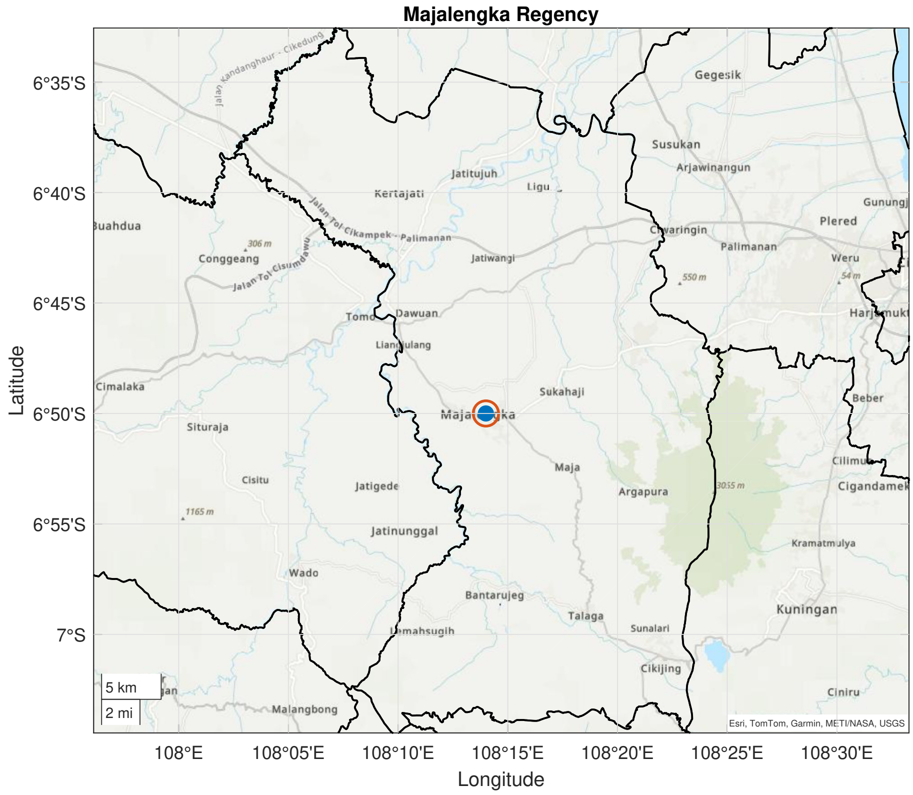

Fig. 1: Location of Majalengka Regency.

2.2. Observational Data

Observed rainfall data were obtained from the Badan Meteorologi Klimatologi dan Geofisika (BMKG) through the Kertajati Meteorological Station, located in Majalengka Regency at approximately 6.73° S latitude, 108.26° E longitude, and an elevation of 85 m above sea level. The dataset consists of daily rainfall measurements expressed in millimeters, covering the period from January 1994 to December 2017.

Prior to analysis, the daily rainfall records were subjected to quality control procedures. Special codes indicating missing or unrecorded values were identified and treated accordingly. Isolated missing observations were infilled using the average of rainfall values from the immediately preceding and following days to preserve temporal continuity. Subsequently, the cleaned daily data were aggregated into monthly total rainfall series by summing daily values within each month. This aggregation was performed to align the observational data with the temporal resolution of the climate model outputs.

The resulting monthly rainfall dataset provides a consistent and reliable representation of local rainfall variability in Majalengka and serves as the dependent variable in the ARDL modeling framework.

2.3. CMIP6 Datasets

Climate model rainfall data were derived from the Coupled Model Intercomparison Project Phase 6 (CMIP6), accessed via datasets produced by the European Centre for Medium-Range Weather Forecasts (ECMWF). The CMIP6 data are provided in gridded NetCDF format and include the precipitation flux variable (pr).

This study utilizes CMIP6 precipitation data under three Shared Socioeconomic Pathway (SSP) scenarios: SSP1-2.6 (low emissions), SSP2-4.5 (medium emissions), and SSP5-8.5 (high emissions). The spatial resolution of the CMIP6 grid is approximately 1.25° × 1.875°, covering the broader region of Indonesia. Since the Kertajati station does not coincide exactly with a model grid point, spatial interpolation was required to extract rainfall estimates representative of the station location. Bicubic interpolation was applied to obtain a continuous rainfall time series at the station coordinates.

The precipitation flux values were converted into monthly rainfall totals to ensure consistency with the observational dataset. The combined CMIP6 time series was divided into a historical or training period (1994–2014) and a testing period (2015–2017) to evaluate model performance across different emission scenarios. These CMIP6 rainfall series were treated as exogenous variables in the ARDL framework, enabling the assessment of both short-run and long-run relationships with observed rainfall.

3. Methods

This section outlines the methodological framework adopted in this study, including data interpolation, stationarity testing, ARDL modeling, cointegration analysis, and model evaluation.

3.1. Bicubic Interpolation

The bicubic interpolation function can be expressed as

$$f(t, x) = \sum_{i=0}^{3} \sum_{j=0}^{3} a_{ij} (t - t_0)^i (x - x_0)^j, \tag{1}$$

where $f(t, x)$ denotes the interpolated rainfall value at time t and location x, t_0 represents the time coordinate of the upper-left corner of the interpolation grid, x_0 denotes the spatial coordinate of the upper-left corner, and a_{ij} are the coefficients determined from the known rainfall values surrounding the point (t_0, x_0).

Bicubic interpolation was applied to resample the precipitation data in order to align the spatial resolution of the CMIP6 model outputs with the observational dataset. This interpolation method provides smooth spatial transitions between grid cells and is commonly used in climate data regridding applications. It should be noted that bicubic interpolation assumes spatial smoothness, which may not always fully represent the localized variability of precipitation fields. Compared to hydrological interpolation methods such as Inverse Distance Weighting (IDW) or kriging, bicubic interpolation prioritizes smooth spatial continuity rather than stochastic spatial dependence. In this study, the interpolation procedure was used solely for spatial resampling, and the resulting dataset was verified to ensure that no negative precipitation values were generated.

3.2. Augmented Dickey–Fuller (ADF) Test

Stationarity is a fundamental requirement in time series analysis because non-stationary data can lead to biased and invalid statistical inference. The Augmented Dickey–Fuller (ADF) test, developed by Dickey and Fuller [33], extends the standard Dickey–Fuller test by incorporating lagged differences to account for residual autocorrelation, thereby improving test reliability [34]. In the ARDL framework, the ADF test is applied to ensure that all variables are integrated of order $I(0)$ or $I(1)$, since the presence of any $I(2)$ variable invalidates the ARDL approach [22]. All tests are conducted at a significance level of $\alpha = 0.05$ with the following hypotheses:

$$H_0 : \text{The time series is non-stationary (contains a unit root)},$$
$$H_1 : \text{The time series is stationary (does not contain a unit root)}.$$

The ADF test can be specified using three alternative regression forms:

1. Without intercept: $\Delta Y_t = \delta Y_{t-1} + u_t$,
2. With intercept: $\Delta Y_t = \beta + \delta Y_{t-1} + u_t$
3. With intercept and time trend: $\Delta Y_t = \beta_1 + \beta_2 t + \delta Y_{t-1} + u_t$,

where Δ denotes the first-difference operator, Y_t represents the time series under consideration, and u_t is a white-noise error term.

The computed ADF test statistic is compared with the MacKinnon critical values. If the test statistic is smaller than the corresponding critical value, the null hypothesis is rejected and the series is concluded to be stationary. Conversely, if the test statistic exceeds the critical value, the null hypothesis cannot be rejected, indicating that the series is non-stationary.

3.3. Autoregressive Distributed Lag (ARDL) Model

The Autoregressive Distributed Lag (ARDL) model combines lagged values of the dependent variable and lagged values of independent variables to capture both short-run dynamics and long-run relationships in time series data [35]. A key advantage of the ARDL framework is its ability to accommodate variables integrated of order $I(0)$ or $I(1)$, making it suitable for heterogeneous datasets. Moreover, ARDL allows cointegration testing without requiring all variables to have the same order of integration, which enhances its flexibility and applicability in empirical analysis [36].

3.3.1. Autoregressive (AR) Model

The Autoregressive (AR) model describes the influence of past values of a dependent variable on its current value. In an AR model, the present value of the dependent variable is explained by its own lagged values. In general, an autoregressive model of order p is expressed as

$$y_t = \sum_{i=1}^{p} a_i y_{t-i} + \varepsilon_t, \tag{2}$$

where y_t denotes the value of the dependent variable at time t, y_{t-i} represents the i-th lag of the dependent variable, a_i is the regression coefficient associated with the i-th lag, and ε_t is the error term at time t.

3.3.2. Distributed Lag (DL) Model

The Distributed Lag (DL) model captures the effect of an independent variable observed at multiple lag periods on a dependent variable. This model allows the analysis of impacts that occur not only contemporaneously but also over previous periods. In general, a distributed lag model of order q can be written as

$$y_t = \sum_{i=0}^{q} c_i x_{t-i} + \varepsilon_t, \tag{3}$$

where y_t is the dependent variable at time t, x_{t-i} denotes the i-th lag of the independent variable, c_i represents the regression coefficient associated with the lagged independent variable, and ε_t is the error term.

3.3.3. Autoregressive Distributed Lag (ARDL) Model

By combining the AR and DL components, the general form of an Autoregressive Distributed Lag (ARDL) model of order (p, q) can be expressed as

$$y_t = \sum_{i=1}^{p} a_i y_{t-i} + \sum_{i=0}^{q} c_i x_{t-i} + \varepsilon_t, \tag{4}$$

where y_t is the dependent variable at time t, y_{t-i} represents the lagged values of the dependent variable, x_{t-i} denotes the lagged values of the independent variable, a_i and c_i are the corresponding regression coefficients, and ε_t is the error term. This formulation enables the simultaneous examination of short-run dynamics and long-run relationships between variables [35].

3.3.4. Unrestricted Error Correction Representation

To analyze the short-run dynamics and the adjustment toward the long-run equilibrium, the ARDL model can be reparameterized into its Unrestricted Error Correction Model (UECM) form. The UECM representation is the basis of the bounds testing procedure proposed by [37].

The UECM specification can be written as

$$\Delta Y_t = \alpha + \sum_{i=1}^{p-1} \phi_i \Delta Y_{t-i} + \sum_{j=0}^{q-1} \beta_j \Delta X_{t-j} + \lambda Y_{t-1} + \delta X_{t-1} + \varepsilon_t \tag{5}$$

where Δ denotes the first-difference operator, α is the intercept, and ε_t is the error term.

The coefficients of the lagged level variables (Y_{t-1}, X_{t-1}) capture the long-run relationship between the variables. The joint significance of these coefficients is evaluated using the bounds testing procedure.

Once the existence of a long-run relationship is confirmed, the model can be expressed in the Error Correction Model (ECM) form:

$$\Delta Y_t = \alpha + \sum_{i=1}^{p-1} \phi_i \Delta Y_{t-i} + \sum_{j=0}^{q-1} \beta_j \Delta X_{t-j} + \lambda ECT_{t-1} + \varepsilon_t \tag{6}$$

where

$$ECT_{t-1} = Y_{t-1} - \theta_0 - \theta_1 X_{t-1} \tag{7}$$

The parameter λ represents the speed of adjustment toward the long-run equilibrium. A negative and statistically significant value of λ indicates that deviations from the long-run equilibrium are corrected over time.

3.4. Model Selection Criteria

The lag structure of the Autoregressive Distributed Lag (ARDL) model was determined based on out-of-sample predictive performance. Several candidate lag specifications were estimated, and their forecasting accuracy was evaluated using the Root Mean Squared Error (RMSE) on the validation dataset. The model with the lowest RMSE was selected as the optimal specification. The RMSE is defined as

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(P_{\mathrm{obs},i} - P_i\right)^2}, \tag{8}$$

where $P_{\mathrm{obs},i}$ denotes the observed rainfall value at the i-th observation, P_i represents the predicted rainfall value at the i-th observation, and n is the total number of observations.

This approach prioritizes predictive accuracy and ensures that the selected model generalizes well to unseen data. The RMSE criterion provides a direct measure of forecast error and is widely used in time-series forecasting studies to select the most reliable predictive model [Ref: Cottonyield].

3.5. Cointegration Test

Cointegration describes a condition in which two or more non-stationary time series move together toward a stable long-run equilibrium despite short-run fluctuations, making it essential for validating long-run relationships in the ARDL framework. In this study, cointegration is examined using the bounds testing approach proposed by [36], which tests the joint significance of long-run parameters through an F-statistic under the hypotheses

$$\begin{aligned} &H_0: \text{No long-run relationship exists (no cointegration)}, \ \alpha_1 = \alpha_2 = \cdots = \alpha_n = 0, \\ &H_1: \text{A long-run relationship exists (cointegration), at least one } \alpha_i \neq 0. \end{aligned}$$

The F-statistic, defined as

$$F = \frac{\mathrm{SSR}/k}{\mathrm{SSE}/(n-k-1)}, \tag{9}$$

is compared with lower $I(0)$ and upper $I(1)$ critical bounds to determine the presence of cointegration: values above $I(1)$ indicate a long-run relationship, values below $I(0)$ indicate no cointegration, and intermediate values are inconclusive [38]. Establishing cointegration allows the ARDL model to be used for analyzing both short-run dynamics and long-run equilibrium relationships between rainfall variables.

3.6. Evaluation Metrics

Evaluation metrics are used to assess the performance of a model in predicting or estimating a given variable. In this study, the performance of the Autoregressive Distributed Lag (ARDL) model is evaluated using three statistical metrics: Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and the coefficient of determination (R^2). These metrics are employed to quantify how closely the model predictions approximate the observed rainfall data. The RMSE is defined as in Eq. (8), while the MAE is expressed as

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|P_{\mathrm{obs},i} - P_i\right|, \tag{10}$$

where $P_{\mathrm{obs},i}$ is the observed rainfall value at the i-th observation, P_i is the corresponding predicted value, and n is the total number of observations. The R^2 statistic is defined as

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(P_{\mathrm{obs},i} - P_i\right)^2}{\sum_{i=1}^{n}\left(P_{\mathrm{obs},i} - \overline{P}_{\mathrm{obs}}\right)^2}, \tag{11}$$

where $\overline{P}_{\mathrm{obs}}$ denotes the mean of the observed rainfall values. These evaluation metrics provide complementary information regarding model accuracy and reliability, enabling a comprehensive assessment of the ARDL model performance in representing observed rainfall variability [39].

4. Results and Discussion

This section presents the empirical findings of the study, beginning with the extracted datasets and followed by stationarity testing, ARDL estimation, cointegration analysis, model evaluation, and interpretation of the results.

4.1. Data Extraction

The first stage of this study is data extraction. The datasets are classified into two categories: observational data and model data. The visual representations of both datasets used in this study are presented in Fig. 2.

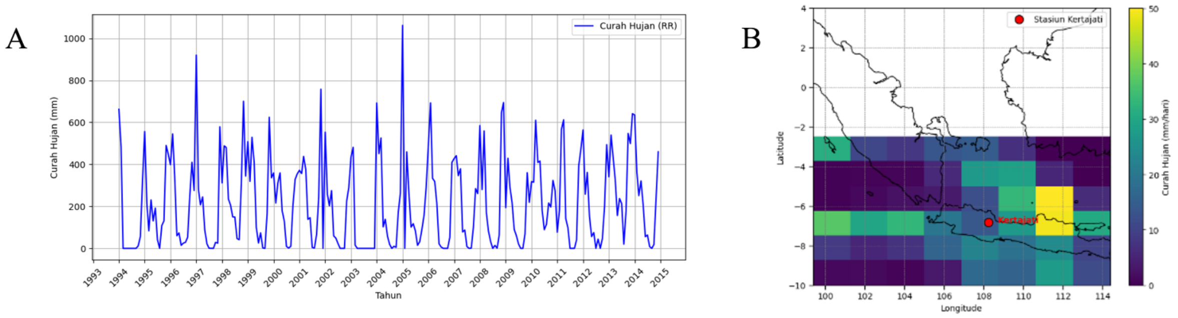

Fig. 2: Monthly rainfall data overview. (A) Observed rainfall data from the BMKG Kertajati Station for January 1994–December 2017. (B) Grid-Averaged CMIP6 model rainfall data for January 1994–December 2014.

As shown in Fig. 2, the observed rainfall data exhibit pronounced monthly variability with several high rainfall peaks, reflecting the strong influence of seasonal and local climatic processes. In contrast, the CMIP6 model data display smoother rainfall patterns, which can be attributed to the coarser spatial resolution of global climate models. This discrepancy indicates that climate model outputs require further evaluation to assess their ability to represent the observed rainfall characteristics recorded at the BMKG Kertajati Station.

4.1.1. Observed Data

Daily rainfall data were obtained from the Kertajati Meteorological Station (6.73° S latitude, 108.26° E longitude, 85 m elevation) for the period January 1994–December 2017 from BMKG, with measurements recorded in mm/day. Special codes (8888 for unmeasured and 9999 for missing values) were identified and treated, and missing observations were infilled using the average of the preceding and following days to preserve temporal continuity. The cleaned daily data were then aggregated into monthly totals, producing a time series that exhibits strong seasonal variability and several extreme rainfall peaks, notably around 2005 (Fig. 2A).

4.1.2. Model Data

CMIP6 precipitation data were obtained from the ECMWF portal in gridded NetCDF format using the precipitation flux variable (pr), which was converted from $\mathrm{kg\,m^{-2}\,s^{-1}}$ to $\mathrm{mm\,day^{-1}}$ by multiplying by 86,400. Because the Kertajati station (6.73° S, 108.26° E) does not coincide with CMIP6 grid points, rainfall values at the station location were obtained using bicubic spatial interpolation implemented with the `griddedInterpolant` function in MATLAB, with missing values replaced by averages of neighboring grids.

As shown in Fig. 3, all scenarios exhibit similar seasonal patterns with rainfall peaks between January and March and drier conditions in mid-year, but rainfall intensity and variability increase with emission level. The high-emission SSP5-8.5 scenario shows the most extreme and irregular rainfall, while SSP1-2.6 is the most stable and SSP2-4.5 exhibits intermediate behavior, indicating that stronger greenhouse gas forcing amplifies local rainfall variability.

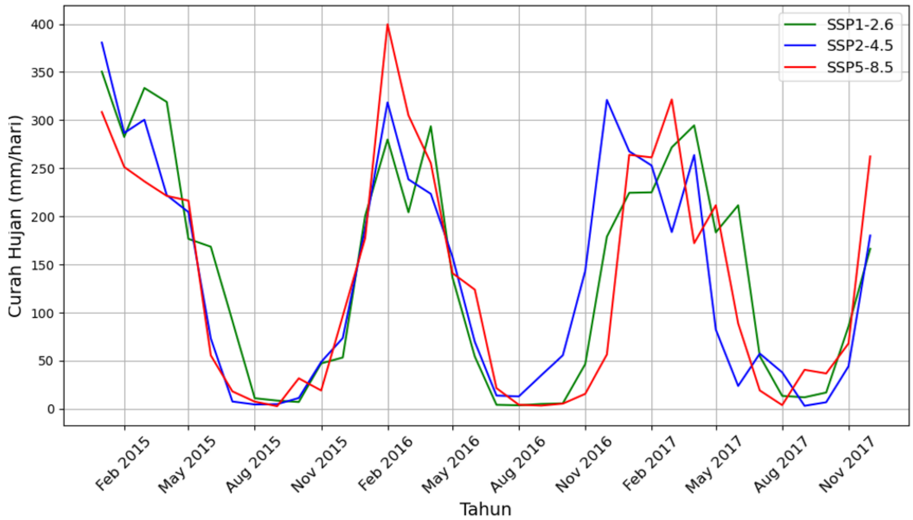

Fig. 3: Monthly rainfall projections from CMIP6 under different emission scenarios (2015–2017).

4.2. Stationarity Test

Following the extraction of observational and model datasets, the next step is to examine the stationarity properties of the time series to determine the presence of unit roots. Stationarity testing is a crucial prerequisite prior to ARDL modeling, as the ARDL framework assumes that the variables are stationary at level or integrated of order $I(0)$ or $I(1)$. The stationarity test was conducted using the Augmented Dickey–Fuller (ADF) test at a significance level of $\alpha = 0.05$ for two main variables: observed rainfall (rr) from BMKG and CMIP6 model rainfall (*precipitation*). The results of the stationarity tests are presented in Table 1.

Table 1: Results of the Augmented Dickey–Fuller (ADF) stationarity test.

Data Source	Scenario	ADF Statistic	Critical Value (5%)	p-value	Conclusion
Observed data	–	−10.055	−3.4270	0.001	Stationary
Model data	Historical + SSP1-2.6	−7.6689	−3.4270	0.001	Stationary
	Historical + SSP2-4.5	−7.3897	−3.4270	0.001	Stationary
	Historical + SSP5-8.5	−7.7091	−3.4270	0.001	Stationary

As shown in Table 1, all variables exhibit ADF statistics that are smaller than the MacKinnon critical value at the 5% significance level (−3.4270). For instance, the observed rainfall series yields an ADF statistic of −10.055 with a p-value of 0.001, indicating rejection of the null hypothesis and confirming stationarity at level. Similarly, the CMIP6 rainfall series under all emission scenarios are stationary at level, with ADF statistics of −7.6689, −7.3897, and −7.7091 for SSP1-2.6, SSP2-4.5, and SSP5-8.5, respectively, all accompanied by p-values of 0.001.

These results indicate that both the observed rainfall data and CMIP6 model outputs satisfy the stationarity requirement. Consequently, all variables can be directly incorporated into the ARDL modeling framework without the need for further differencing or transformation. The ARDL framework can still be applied to stationary variables because it provides a flexible structure for modeling distributed lag relationships [40].

4.3. Autoregressive Distributed Lag (ARDL) Modeling

The Autoregressive Distributed Lag (ARDL) model was employed to analyze the relationship between observed rainfall (rr) and CMIP6-projected rainfall (*precipitation*) by accounting for lagged effects of both variables. This modeling framework is designed to capture short-run dynamics as well as long-run relationships between the two series. The modeling procedure began with the identification of optimal lag structures for the ARDL specification. Initially, candidate models were selected based on the minimum value of the Akaike Information Criterion (AIC).

However, because the AIC-selected model did not yield satisfactory predictive performance when evaluated on the testing dataset, the final model selection was based on the lowest Root Mean Squared Error (RMSE) obtained from out-of-sample testing.

ARDL modeling was conducted separately for each CMIP6 emission scenario—SSP1-2.6, SSP2-4.5, and SSP5-8.5—to assess the sensitivity of model performance to different emission pathways. After selecting the optimal model for each scenario, parameter estimation was performed using the Ordinary Least Squares (OLS) method. The resulting models were then evaluated using three performance metrics: Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and the coefficient of determination (R^2), to assess predictive accuracy and goodness-of-fit relative to the observed rainfall data.

4.3.1. Optimal Lag Selection

The determination of optimal lag lengths is a critical step in ARDL modeling, as the selected lags define how many previous time periods are used to predict current values. In this study, an extensive search over lag combinations ranging from 1 to 96 months (equivalent to 8 years) was conducted for both the dependent and independent variables. This wide lag range was adopted to explore potential long-memory effects and to identify models capable of capturing complex rainfall dynamics.

The lag selection process initially involved computing the AIC for all possible lag combinations, with the model exhibiting the lowest AIC value selected as the best candidate. Using the training dataset, the ARDL$(83, 82)$ model produced the minimum AIC value of -321.2696. This AIC-optimal specification was identical across all scenarios, as the same historical training data were used.

Nevertheless, when the AIC-selected model was evaluated using scenario-specific testing data, its predictive performance varied considerably. Consequently, RMSE computed on the testing dataset was adopted as the primary criterion for final model selection. The optimal lag structures based on minimum RMSE differed across scenarios, as follows:

- SSP1-2.6: ARDL$(7, 21)$ with RMSE $= 135.19$,
- SSP2-4.5: ARDL$(9, 18)$ with RMSE $= 137.24$,
- SSP5-8.5: ARDL$(2, 6)$ with RMSE $= 140.80$.

Although AIC-based selection yields the same model for all scenarios, the RMSE-based evaluation reveals scenario-dependent differences in predictive performance. These differences indicate that rainfall dynamics in Majalengka Regency respond differently under varying emission scenarios, emphasizing the importance of incorporating out-of-sample validation when selecting ARDL model specifications.

4.3.2. ARDL Model Estimation

After determining the optimal lag combinations, the next step involved estimating the parameters of the ARDL models using the Ordinary Least Squares (OLS) method. This estimation aims to examine how past rainfall conditions in Majalengka influence current rainfall, based on both observational data from BMKG and projected rainfall from CMIP6 climate models. The estimation procedure was conducted separately for each CMIP6 emission scenario—SSP1-2.6, SSP2-4.5, and SSP5-8.5. Each ARDL specification includes regression coefficients that quantify the influence of lagged rainfall values on current observed rainfall.

ARDL Model under the SSP1-2.6 Scenario The optimal model for the SSP1-2.6 scenario is ARDL$(7, 21)$, indicating that current observed rainfall is influenced by rainfall observed over the previous seven months and CMIP6-projected rainfall over the previous 21 months. The estimated model can be expressed as

$$\begin{aligned}
y_t = {} & 244.8135 + 0.0478\,y_{t-1} + 0.0662\,y_{t-2} + 0.0756\,y_{t-3} - 0.0129\,y_{t-4} - 0.0170\,y_{t-5} \\
& + 0.0271\,y_{t-6} - 0.0065\,y_{t-7} + 0.0480\,x_t + 0.3263\,x_{t-1} - 0.1158\,x_{t-2} - 0.2278\,x_{t-3} \\
& - 0.1030\,x_{t-4} + 0.0783\,x_{t-5} + 0.0051\,x_{t-6} + 0.0298\,x_{t-7} - 0.5057\,x_{t-8} + 0.1963\,x_{t-9} \\
& - 0.1452\,x_{t-10} + 0.0640\,x_{t-11} + 0.3187\,x_{t-12} + 0.2403\,x_{t-13} - 0.0063\,x_{t-14} \\
& - 0.0682\,x_{t-15} - 0.0413\,x_{t-16} - 0.2298\,x_{t-17} - 0.2373\,x_{t-18} - 0.2388\,x_{t-19} \\
& + 0.2263\,x_{t-20} - 0.1579\,x_{t-21}, \qquad (12)
\end{aligned}$$

where y_t denotes observed rainfall at month t, y_{t-i} represents observed rainfall lagged by i months, x_t denotes CMIP6 rainfall at month t, and x_{t-i} represents CMIP6 rainfall lagged by i months.

The estimation results indicate that not all lagged terms exert the same influence. Several lags exhibit positive effects, such as the first to third lags of observed rainfall and selected lags of CMIP6 rainfall (e.g., lags 1, 11, and 13), while others show negative effects. This pattern reflects the complex rainfall dynamics in which past conditions may enhance or suppress current rainfall depending on prevailing climatic processes.

ARDL Model under the SSP2-4.5 Scenario For the SSP2-4.5 scenario, the optimal specification is ARDL$(9, 18)$. This model suggests that rainfall conditions up to nine months in the past and CMIP6 projections up to eighteen months earlier continue to influence current rainfall. The estimated model is given by

$$\begin{aligned}
y_t = {} & 315.8312 + 0.0253\,y_{t-1} + 0.0468\,y_{t-2} + 0.0797\,y_{t-3} - 0.0295\,y_{t-4} - 0.0235\,y_{t-5} \\
& + 0.0128\,y_{t-6} - 0.0219\,y_{t-7} + 0.0595\,y_{t-8} + 0.0786\,y_{t-9} - 0.0678\,x_t + 0.3641\,x_{t-1} \\
& - 0.1291\,x_{t-2} - 0.2237\,x_{t-3} - 0.1494\,x_{t-4} + 0.0906\,x_{t-5} - 0.0420\,x_{t-6} - 0.0019\,x_{t-7} \\
& - 0.4907\,x_{t-8} + 0.1775\,x_{t-9} - 0.1063\,x_{t-10} + 0.0943\,x_{t-11} + 0.2750\,x_{t-12} + 0.2203\,x_{t-13} \\
& - 0.0542\,x_{t-14} - 0.0572\,x_{t-15} - 0.0380\,x_{t-16} - 0.2492\,x_{t-17} - 0.2178\,x_{t-18}. \qquad (13)
\end{aligned}$$

These results indicate a relatively long climatological memory in the SSP2-4.5 scenario, where rainfall conditions from several preceding months continue to affect current rainfall. This suggests the presence of persistent rainfall dynamics influenced by both observed conditions and climate model projections.

ARDL Model under the SSP5-8.5 Scenario Under the high-emission SSP5-8.5 scenario, the optimal model is ARDL$(2, 6)$, representing the most parsimonious specification among the three scenarios. The estimated model is expressed as

$$\begin{aligned}
y_t = {} & 228.6394 + 0.0884\,y_{t-1} + 0.1526\,y_{t-2} + 0.0852\,x_t + 0.6073\,x_{t-1} + 0.0123\,x_{t-2} \\
& - 0.2739\,x_{t-3} - 0.2769\,x_{t-4} - 0.1675\,x_{t-5} - 0.1719\,x_{t-6}. \qquad (14)
\end{aligned}$$

In this scenario, only the two most recent months of observed rainfall and six months of CMIP6 rainfall contribute to current rainfall variability. The largest coefficients are associated with the first lag of both observed and modeled rainfall, indicating that under high-emission conditions, rainfall dynamics tend to be more short-term and less persistent compared to lower-emission scenarios.

4.4. Cointegration Test

After estimating the Autoregressive Distributed Lag (ARDL) models for each CMIP6 emission scenario (SSP1–2.6, SSP2–4.5, and SSP5–8.5), the next step was to examine the potential long-run association between observed rainfall from BMKG and rainfall simulated by CMIP6 models. The ARDL framework allows the incorporation of lagged dynamics and enables the evaluation of both short-run and long-run relationships between variables. To explore the existence of a long-run relationship, the bounds testing procedure within the ARDL framework was initially implemented. This approach evaluates the joint significance of long-run parameters using an F-statistic. The null hypothesis assumes the absence of a long-run relationship, and the calculated F-statistic is compared against the lower and upper critical bounds at the 5% significance level. If the F-statistic exceeds the upper bound, the null hypothesis is rejected, indicating evidence of a long-run association. Conversely, if it falls below the lower bound, no long-run relationship is inferred, while values between the bounds lead to inconclusive results [22, 36]. The results of the bounds testing procedure for all scenarios are presented in Table 2.

Table 2: Results of ARDL bounds testing.

Scenario	F-statistic	Upper bound (5%)	Lower bound (5%)
SSP1–2.6	9.5798	3.79	2.45
SSP2–4.5	10.0727	3.79	2.45
SSP5–8.5	22.5596	3.79	2.45

The Augmented Dickey–Fuller (ADF) test results indicate that all variables are stationary at level, $I(0)$. In such circumstances, the use of the ARDL bounds testing approach as a formal cointegration test becomes less essential, since any linear combination of $I(0)$ variables is inherently stationary. Therefore, the bounds test results in this study are interpreted as supporting evidence of potential long-run association rather than a strict confirmation of cointegration. The ARDL framework is retained primarily because of its flexibility in modeling dynamic lag structures and capturing short-run adjustments between observed rainfall and CMIP6-simulated rainfall. In this context, the ARDL specification is interpreted as a distributed lag model that allows the estimation of both immediate and lagged effects.

As shown in Table 2, the F-statistics for all CMIP6 scenarios are substantially higher than the upper critical bound at the 5% significance level. Although the variables are stationary at level, these results suggest a strong and stable association between observed rainfall and CMIP6-simulated rainfall across all scenarios.

The error correction representation of the ARDL model indicates that the coefficient of the lagged level term of observed rainfall is estimated at -0.762 and is statistically significant at the 1% level. The negative sign confirms the existence of a stable adjustment mechanism toward the long-run equilibrium. Specifically, approximately 76.2% of deviations from the long-run equilibrium are corrected within one period, indicating a relatively rapid adjustment process following short-run shocks.

Following the estimation of the ARDL models, both long-run and short-run parameters were derived using the training dataset (all scenario use same training dataset, i.e. historical datasets). Long-run coefficients were estimated to assess the persistent influence of time and CMIP6 rainfall projections on observed rainfall. Estimation was conducted using the multiplier approach implemented in R. The long-run estimation results are summarized in Table 3.

Table 3: Long-run estimation results.

Variable	Coefficient	Std. Error	t-value	p-value
Intercept	299.9399	64.9350	4.6191	0.000006
Time (t)	0.1159	0.1861	0.6229	0.5339
CMIP6 rainfall	-0.5609	0.3139	-1.7869	0.0752

The long-run estimation results indicate a negative coefficient for CMIP6 rainfall (-0.5609), suggesting that, in the long term, increases in rainfall projected by climate models are associated with a relative decrease in observed rainfall in Majalengka Regency. This finding implies that long-term CMIP6 projections may not fully capture local-scale rainfall dynamics, potentially due to spatial resolution limitations and structural biases inherent in global climate models [41].

Short-run parameter estimates are presented in Table 4 and provide insight into the immediate response of observed rainfall to changes in CMIP6 rainfall projections.

Table 4: Short-run estimation results.

Variable	Coefficient	Std. Error	t-value	p-value
Intercept	228.6393	47.2021	4.8438	0.000002
Time (t)	0.0883	0.1427	0.6188	0.5366
CMIP6 rainfall	0.6073	0.1352	4.4892	0.000011

As shown in Table 4, CMIP6 rainfall exhibits a positive and statistically significant short-run effect on observed rainfall, with a coefficient of 0.6073. This result indicates that a 1 mm increase in CMIP6 rainfall is associated with an average increase of approximately 0.61 mm in observed rainfall in Majalengka in the short term. Thus, while long-run relationships reveal notable discrepancies, short-run rainfall fluctuations are relatively well captured by climate model projections. Overall, these findings suggest that CMIP6 models are more reliable in representing short-term rainfall variability than long-term equilibrium behavior at the local scale [41].

4.5. Model Performance Evaluation

Model performance evaluation was conducted to assess the extent to which the ARDL model represents monthly rainfall in Majalengka Regency based on BMKG observational data and ARDL model outputs. The evaluation focused on the testing period from 2015 to 2017 to measure out-of-sample predictive performance and to compare prediction accuracy across different climate change scenarios. Three evaluation metrics were employed: Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and the coefficient of determination (R^2). These metrics were computed for three CMIP6 scenarios: SSP1-2.6, SSP2-4.5, and SSP5-8.5.

Table 5: Comparison of evaluation metrics across CMIP6 scenarios.

Scenario	RMSE	MAE	R^2
SSP1-2.6	134.75	104.75	0.8481
SSP2-4.5	136.91	107.86	0.8376
SSP5-8.5	140.80	114.53	0.8335

As shown in Table 5, the ARDL model exhibits comparable performance across all scenarios, with R^2 values ranging from 0.8335 to 0.8481. The SSP1-2.6 scenario yields the lowest RMSE and MAE values (134.75 and 104.75, respectively) and the highest R^2 (0.8481), indicating slightly better predictive performance. In contrast, the SSP5-8.5 scenario shows the largest prediction errors, with an RMSE of 140.80 and an MAE of 114.53, as well as the lowest R^2 value (0.8335). These results suggest that, among the three scenarios, SSP1-2.6 provides the most accurate representation of monthly rainfall in Majalengka Regency under the respective climate change pathways.

4.6. Discussion

The results of this study reveal a clear distinction between the short-run and long-run relationships between observed rainfall and CMIP6-simulated rainfall in Majalengka Regency. In the short run, CMIP6 rainfall exhibits a positive and statistically significant relationship with observed rainfall, indicating that climate model outputs are capable of capturing monthly rainfall variability and seasonal patterns. This finding suggests that large-scale atmospheric processes represented in CMIP6, such as monsoonal circulation and seasonal moisture transport, remain influential at the local scale over short temporal horizons.

In contrast, the long-run relationship between observed and simulated rainfall is found to be negative and relatively weak. This divergence between short-run and long-run dynamics highlights the limitations of global climate models when applied directly to local-scale rainfall assessment over extended periods. While CMIP6 models are designed to represent large-scale climate behavior, their coarse spatial resolution constrains their ability to accurately reproduce localized rainfall processes, including orographic effects, land–atmosphere interactions, and convective rainfall characteristics that are prominent in tropical regions such as West Java. Consequently, biases that are negligible at monthly scales may accumulate over time, leading to discrepancies in long-term rainfall representation.

The differences in model performance across emission scenarios further underscore the sensitivity of rainfall dynamics to climate forcing intensity. The ARDL model exhibits the best predictive performance under the low-emission SSP1-2.6 scenario, while prediction errors increase under the higher-emission SSP2-4.5 and SSP5-8.5 scenarios. This pattern suggests that higher greenhouse gas emissions introduce greater nonlinearity and variability in rainfall behavior, reducing the stability of statistical relationships calibrated using historical observations. As climate forcing intensifies, the assumption of stationarity underlying statistical models becomes increasingly challenged, contributing to higher uncertainty in long-term projections.

From a practical perspective, these findings imply that CMIP6 outputs are more reliable for analyzing short-term rainfall variability and seasonal dynamics at the local scale than for assessing long-term rainfall equilibrium without additional processing. Direct application of CMIP6 rainfall projections for long-term local planning should therefore be approached with caution. The incorporation of bias correction, statistical downscaling, or hybrid modeling frameworks may be necessary to bridge the scale mismatch between global climate simulations and local hydrological conditions [42].

Several limitations of this study should be acknowledged. The analysis is based on rainfall observations from a single meteorological station, which may not fully capture spatial variability across Majalengka Regency. In addition, the ARDL framework assumes linear relationships and does not explicitly account for other large-scale climate drivers such as the El Niño–Southern Oscillation (ENSO) or the Indian Ocean Dipole (IOD), which are known to influence rainfall variability in Indonesia. Despite these limitations, the study provides robust insights by explicitly distinguishing between short-run and long-run rainfall dynamics using a statistically rigorous modeling framework.

Future research may extend this work by incorporating multi-station observations, integrating additional climate indices, and applying hybrid approaches that combine statistical and machine learning models. Such efforts would further enhance the understanding of local rainfall responses under changing climate conditions and improve the reliability of climate model applications for regional-scale impact assessments.

5. Conclusion

This study examined the dynamic relationship between observed rainfall and CMIP6 precipitation projections in Majalengka Regency using the Autoregressive Distributed Lag (ARDL) framework. The empirical analysis reveals the presence of a stable adjustment mechanism between observed rainfall and model-simulated precipitation.

The estimated long-run precipitation coefficient is -0.561, indicating a weak negative long-term association between CMIP6 precipitation projections and observed rainfall. The error correction term (ECT_{t-1}) is estimated at -0.762 and is statistically significant, suggesting that approximately 76.2% of short-run disequilibrium is corrected within one period. This result indicates the presence of a stable adjustment process toward the long-run equilibrium. In addition, the model explains approximately 84.81% of the variation in observed rainfall, suggesting a moderate explanatory power of the estimated model.

Overall, the findings suggest that CMIP6 rainfall projections exhibit a measurable statistical relationship with observed rainfall patterns, although the representation of local-scale variability remains limited. These results highlight the importance of applying bias correction or statistical downscaling techniques when using climate model projections for local rainfall analysis.

Several limitations should be noted. The analysis relies on rainfall observations from a single station, which may not fully represent spatial rainfall variability in the region. In addition, the model does not explicitly incorporate other climate drivers that may influence rainfall dynamics. Future studies may extend this work by incorporating multiple observation stations, additional climate variables, and bias-corrected or downscaled climate projections.

CRediT Authorship Contribution Statement

Sri Nurdiati: Conceptualization, Methodology, Writing–Original Draft. **Mohamad Khoirun Najib:** Data Curation, Formal Analysis, Writing–Review & Editing. **Fathia Rahmaisty:** Software, Validation, Visualization.

Declaration of Generative AI and AI-assisted technologies

The authors used ChatGPT to assist with English language editing and translation. The authors take full responsibility for the content, accuracy, and integrity of the manuscript.

Declaration of Competing Interest

The authors declare no competing interests.

Funding and Acknowledgments

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data and Code Availability

The observational rainfall data used in this study are available from the Indonesian Agency for Meteorology, Climatology, and Geophysics (BMKG) upon request. The CMIP6 climate model data are publicly available. The processed datasets and analysis codes used to support the findings of this study are available from the corresponding author upon reasonable request.

References

[1] Elena Cristiano, Marie-Claire ten Veldhuis, and Nick Van De Giesen. "Spatial and temporal variability of rainfall and their effects on hydrological response in urban areas–a review". In: *Hydrology and Earth System Sciences* 21.7 (2017), pp. 3859–3878. DOI: 10.5194/hess-21-3859-2017.

[2] Renato Morbidelli et al. "A review on rainfall data resolution and its role in the hydrological practice". In: *Water* 13.8 (2021), p. 1012. DOI: 10.3390/w13081012.

[3] Sintayehu Eshetu Abebaw. "A Global Review of the Impacts of Climate Change and Variability on Agricultural Productivity and Farmers' Adaptation Strategies". In: *Food Science & Nutrition* 13.5 (2025), e70260. DOI: 10.1002/fsn3.70260.

[4] AV Senthil Kumar et al. "Rainfall prediction using machine learning". In: *Advancements in Climate and Smart Environment Technology*. IGI Global Scientific Publishing, 2024, pp. 100–113. DOI: 10.4018/979-8-3693-3807-0.ch009.

[5] Melly Ariska et al. "Spatio-temporal variations of Indonesian rainfall and their links to Indo-Pacific modes". In: *Atmosphere* 15.9 (2024), p. 1036. DOI: 10.3390/atmos15091036.

[6] Maulana Putra, Mohammad Syamsu Rosid, and Djati Handoko. "A review of rainfall estimation in Indonesia: data sources, techniques, and methods". In: *Signals* 5.3 (2024), pp. 542–561. DOI: 10.3390/signals5030030.

[7] MD Syaifullah et al. "Spatiotemporal analysis of rainfall variability using satellite precipitation in watersheds". In: *Global Journal of Environmental Science and Management* 11.3 (2025), pp. 1143–1162. DOI: 10.22034/gjesm.2025.03.15.

[8] Nourma Al Viandari et al. "Sustainable development strategies of rainfed paddy fields in Central Java, Indonesia: A review". In: *Caraka Tani: Journal of Sustainable Agriculture* 37.2 (2022), pp. 275–288. DOI: 10.20961/carakatani.v37i2.58242.

[9] Heri Mulyanti, Istadi Istadi, and Rahmat Gernowo. "Historical, recent, and future threat of drought on agriculture in East Java, Indonesia: a review". In: *E3S Web of Conferences*. Vol. 448. EDP Sciences. 2023, p. 03016. DOI: 10.1051/e3sconf/202344803016.

[10] Riani Muharomah and Budi Indra Setiawan. "Identification of climate trends and patterns in South Sumatra". In: *Agromet* 36.2 (2022), pp. 79–87. DOI: 10.29244/j.agromet.36.2.79-87.

[11] Ryco Farysca Adi et al. "Designing climate change adaptation action for Bogor City". In: *International Conference on Radioscience, Equatorial Atmospheric Science and Environment*. Springer, 2023, pp. 807–820. DOI: 10.1007/978-981-97-0740-9_72.

[12] MMH Khan et al. "Short-term and long-term rainfall forecasting using ARIMA model". In: *International Journal of Environmental Science and Development* 14.5 (2023), pp. 292–298. DOI: 10.18178/ijesd.2023.14.5.1447.

[13] Ichwana Ramli et al. "Forecasting of rainfall using Seasonal Autoregressive Integrated Moving Average (SARIMA) Aceh, Indonesia". In: *Mathematical Modelling of Engineering Problems* 10.2 (2023). DOI: 10.18280/mmep.100216.

[14] Mohammad Dian Purnama and Mutia Eva Mustafidah. "Relationship Between Temperature and Humidity on Rainfall: A Multiple Linear Regression Analysis". In: *Engineering, Mathematics and Computer Science Journal (EMACS)* 6.2 (2024), pp. 151–156. DOI: 10.21512/emacsjournal.v6i2.11466.

[15] Ichwana Ramli et al. "Linear regression analysis using log transformation model for rainfall data in water resources management Krueng Pase, Aceh, Indonesia". In: *International Journal of Design & Nature and Ecodynamics* 17.1 (2022), pp. 79–86. DOI: 10.18280/ijdne.170110.

[16] Veronika Eyring et al. "Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization". In: *Geoscientific Model Development* 9.5 (2016), pp. 1937–1958. DOI: 10.5194/gmd-9-1937-2016.

[17] Kashif Ullah et al. "Multi-Scenario Flood Susceptibility Projections in the Eastern Hindu Kush: Integrating Machine Learning, CMIP6, and Land Use Change". In: *Earth Systems and Environment* (2025), pp. 1–20. DOI: 10.1007/s41748-025-00793-x.

[18] Amal John et al. "Quantifying CMIP6 model uncertainties in extreme precipitation projections". In: *Weather and Climate Extremes* 36 (2022), p. 100435. DOI: 10.1016/j.wace.2022.100435.

[19] Guillaume Evin, Aurélien Ribes, and Lola Corre. "Assessing CMIP6 uncertainties at global warming levels". In: *Climate Dynamics* 62.8 (2024), pp. 8057–8072. DOI: 10.1007/s00382-024-07323-x.

[20] Matthew J Gidden et al. "Global emissions pathways under different socioeconomic scenarios for use in CMIP6: a dataset of harmonized emissions trajectories through the end of the century". In: *Geoscientific Model Development* 12.4 (2019), pp. 1443–1475. DOI: 10.5194/gmd-12-1443-2019.

[21] Malte Meinshausen et al. "The shared socio-economic pathway (SSP) greenhouse gas concentrations and their extensions to 2500". In: *Geoscientific Model Development* 13.8 (2020), pp. 3571–3605. DOI: 10.5194/gmd-13-3571-2020.

[22] Emeka Nkoro and Aham Kelvin Uko. "Autoregressive Distributed Lag (ARDL) cointegration technique: application and interpretation". In: *Journal of Statistical and Econometric Methods* 5.4 (2016), pp. 63–91.

[23] Abu Hayat Md Saiful Islam et al. "Beyond Rainfall and Fertilizer: Autoregressive Distributed Lag Insights on Population, Greenhouse Gas Emissions, and Wheat Sustainability in Bangladesh". In: *Environmental and Sustainability Indicators* (2025), p. 101098. DOI: 10.1016/j.indic.2025.101098.

[24] Pervez Zamurrad Janjua, Ghulam Samad, and Nazakatullah Khan. "Climate change and wheat production in Pakistan: an autoregressive distributed lag approach". In: *NJAS-Wageningen Journal of Life Sciences* 68 (2014), pp. 13–19. DOI: 10.1016/j.njas.2013.11.002.

[25] Duaa B Telfah, Nawal Louzi, and Tala M AlBashir. "Water demand time series forecast by autoregressive distributed lag (ARDL) co-integration model". In: *Journal of Water and Land Development* 50 (2021), pp. 195–206. DOI: 10.24425/jwld.2021.138175.

[26] Isaac Adjei Mensah et al. "Potential economic indicators and environmental quality in African economies: new insight from cross-sectional autoregressive distributed lag approach". In: *Environmental Science and Pollution Research* 28.40 (2021), pp. 56865–56891. DOI: 10.1007/s11356-021-14598-8.

[27] Zhongfeng Xu et al. "Bias-corrected CMIP6 global dataset for dynamical downscaling of the historical and future climate (1979–2100)". In: *Scientific Data* 8.1 (2021), p. 293. DOI: 10.1038/s41597-021-01079-3.

[28] Xinyi Li and Zhong Li. "Evaluation of bias correction techniques for generating high-resolution daily temperature projections from CMIP6 models". In: *Climate Dynamics* 61.7 (2023), pp. 3893–3910. DOI: 10.1007/s00382-023-06778-8.

[29] Fasil M Rettie et al. "High-resolution CMIP6 climate projections for Ethiopia using the gridded statistical downscaling method". In: *Scientific Data* 10.1 (2023), p. 442. DOI: 10.1038/s41597-023-02337-2.

[30] Mohammed Sanusi Shiru and Eun-Sung Chung. "Performance evaluation of CMIP6 global climate models for selecting models for climate projection over Nigeria: Performance Evaluation of CMIP6 Global Climate Models for Selecting Models for Climate Projection over Nigeria". In: *Theoretical and Applied Climatology* 146.1 (2021), pp. 599–615. DOI: 10.1007/s00704-021-03746-2.

[31] S Fathaya, E Kusratmoko, and R Saraswati. "Characteristics of Landslide and Rainfall Areas in Majalengka Regency, West Java Province". In: *IOP Conference Series: Earth and Environmental Science*. Vol. 884. 1. IOP Publishing. 2021, p. 012054. DOI: 10.1088/1755-1315/884/1/012054.

[32] Fransiska Yustiana and Najib Ibrahim. "Analisis Deret Waktu Curah Hujan dan Karakteristik Iklim di Kota Majalengka". In: *Jurnal Teknik Sipil* 12.2 (2023), pp. 156–163. DOI: 10.24815/jts.v12i2.32097.

[33] David A Dickey and Wayne A Fuller. "Distribution of the estimators for autoregressive time series with a unit root". In: *Journal of the American Statistical Association* 74.366a (1979), pp. 427–431. DOI: 10.2307/2286348.

[34] Dina Intan Fitria and Abd Jamal. "Analisis faktor-faktor yang mempengaruhi ekspor daerah antar provinsi di Indonesia". In: *Jurnal Ilmiah Mahasiswa Ekonomi Pembangunan* 2.4 (2017), pp. 509–516.

[35] Uwe Hassler and Jürgen Wolters. "Autoregressive distributed lag models and cointegration". In: *Allgemeines Statistisches Archiv* 90.1 (2006), pp. 59–74. DOI: 10.1007/s10182-006-0221-5.

[36] M Hashem Pesaran and Yongcheol Shin. *An autoregressive distributed lag modelling approach to cointegration analysis*. Tech. rep. 9514. Faculty of Economics, University of Cambridge, 1995.

[37] M Hashem Pesaran, Yongcheol Shin, and Richard J Smith. "Bounds testing approaches to the analysis of level relationships". In: *Journal of Applied Econometrics* 16.3 (2001), pp. 289–326. DOI: 10.1002/jae.616.

[38] A. D. Wahyuningtyas, H. Sasana, and R. R. Sugiarti. "Analysis of the Influence of Digital Economic Development on Economic Growth in Indonesia Year 1996–2019". In: *Dinamika Ekonomi: Jurnal Ekonomi dan Bisnis* 3.1 (2019), pp. 269–281.

[39] R. Mahindraguna, R. H. Dananjaya, and G. Chrismaningwang. "Validasi Data Hujan Satelit IMERG Terkalibrasi dengan Metode Geographically Weighted Regression terhadap Data Hujan Stasiun". In: *Proceeding Civil Engineering Research Forum*. Vol. 3. 1. 2023, pp. 149–159.

[40] M Hashem Pesaran and Yongcheol Shin. "An autoregressive distributed lag modelling approach to cointegration analysis". In: *Econometrics and Economic Theory in the 20th Century: The Ragnar Frisch Centennial Symposium*. Cambridge University Press, 1999, pp. 371–413. DOI: 10.1017/CCOL521633230.011.

[41] Ramandeep Kumar Sharma et al. "Modelling the climate change and cotton yield relationship in Mississippi: Autoregressive distributed lag approach". In: *Ecological Indicators* 166 (2024), p. 112573. DOI: 10.1016/j.ecolind.2024.112573.

[42] Mohamad Khoirun Najib et al. "Quantile mapping using the Alpha Power Transformed X-Lindley (APTXL) distribution for bias correction of Coupled Model Intercomparison Project Phase 6 (CMIP6) outputs: a case study over the Toba Lake region". In: *Theoretical and Applied Climatology* 157.1 (2026), p. 57. DOI: 10.1007/s00704-025-05975-1.