

Integrating Spatial Autoregressive Exogenous with Ordinary Kriging for Improved Rainfall Prediction in Java: Enhancing Accuracy with Climate Variables and Spatial Autocorrelation

Sandrina Najwa¹, Dhanti Aurilia Pratiwi², Muhammad Rhafi Ahdian³, Ayu Indriani⁴,
I Gede Nyoman Mindra Jaya⁵, Annisa Nur Falah⁶ and Budi Nurani Ruchjana^{7*}

^{1,2,3,4,5}Department of Statistics, Padjadjaran University, Bandung, Indonesia

^{6,7}Department of Mathematics, Padjadjaran University, Bandung, Indonesia

Email: *budi.nurani@unpad.ac.id

Abstract

Indonesia is a tropical country with high rainfall influenced by its archipelagic geography and phenomena like El Niño and La Niña. According to the Meteorology, Climatology, and Geophysics Agency (BMKG), La Niña can increase Indonesia's monthly rainfall by 20-40% above normal. Despite numerous existing spatial interpolation methods, there remains a significant research gap in accurately predicting rainfall at unsampled locations, specifically when considering both spatial autocorrelation and multiple climate variables simultaneously. This research proposes Spatial Autoregressive Exogenous Kriging (SAR-X Kriging), a novel hybrid approach that integrates the SAR-X model with Ordinary Kriging to enhance rainfall prediction accuracy. Unlike conventional methods, SAR-X Kriging explicitly captures both spatial dependence and the influence of external climate factors, improving predictive performance. SAR-X Kriging first models spatial dependencies between locations and incorporates exogenous climate variables (surface pressure, air temperature, humidity, wind speed, and solar radiation) to enhance prediction accuracy. It also applies kriging for spatial interpolation. This method was chosen for its robustness in capturing spatial dependence and external influences. The analysis revealed significant spatial dependence across districts/cities in Java Island based on the Moran's Index test. The best SAR-X model, utilizing air temperature and wind speed as exogenous variables, achieved a p-value of 6.0352×10^{-9} . Predictions using SAR-X Kriging yielded the lowest Mean Absolute Percentage Error (MAPE) of 3.82%, outperforming the standalone SAR-X method MAPE 4.68% and the Ordinary Kriging method MAPE 3.86%. Practically, these results provide reliable rainfall predictions, enabling better climate-informed decision-making in water resource management, agricultural planning, and flood prevention strategies in Java.

Keywords: climate; rainfall; MAPE; SAR-X; SAR-X Kriging.

Abstrak

Indonesia merupakan negara tropis dengan curah hujan tinggi yang dipengaruhi oleh kondisi geografis kepulauan serta fenomena alam seperti El Niño dan La Niña. Menurut Badan Meteorologi, Klimatologi, dan Geofisika (BMKG), La Niña mampu meningkatkan curah hujan bulanan Indonesia hingga 20-40% di atas normal. Meskipun terdapat berbagai metode interpolasi spasial yang telah dikembangkan, masih terdapat kesenjangan penelitian dalam menghasilkan prediksi curah hujan secara akurat di lokasi yang tidak tersampel, terutama ketika mempertimbangkan secara bersamaan ketergantungan spasial serta pengaruh dari berbagai variabel iklim. Penelitian ini mengusulkan metode bernama Spatial Autoregressive Exogenous Kriging (SAR-X Kriging), sebuah pendekatan hybrid baru yang mengintegrasikan model SAR-X dengan metode Ordinary Kriging untuk meningkatkan akurasi prediksi curah hujan. Tidak seperti metode konvensional, SAR-X Kriging secara eksplisit menangkap ketergantungan spasial serta pengaruh faktor iklim eksternal, sehingga meningkatkan kinerja prediktif. SAR-X Kriging bekerja dengan memodelkan terlebih dahulu ketergantungan spasial antar lokasi, kemudian memasukkan variabel eksogen berupa tekanan permukaan,

*) Corresponding author

Submitted October 24th, 2024, Revised March 8th, 2025,

Accepted for publication March 12th, 2025, Published Online March 31st, 2025

©2025 The Author(s). This is an open-access article under CC-BY-SA license (<https://creativecommons.org/licence/by-sa/4.0/>)

subu udara, kelembaban, kecepatan angin, dan radiasi matahari untuk meningkatkan akurasi prediksi, serta terakhir menerapkan teknik kriging untuk interpolasi spasial. Metode ini dipilih karena mampu menangkap secara lebih baik ketergantungan spasial sekaligus pengaruh variabel eksternal dibandingkan metode konvensional. Hasil analisis menunjukkan adanya ketergantungan spasial yang signifikan antar kabupaten/kota di Pulau Jawa berdasarkan uji Moran's Index. Model SAR-X terbaik diperoleh dengan variabel suhu udara dan kecepatan angin, mencapai nilai p-value sebesar 6.0352×10^{-9} . Prediksi menggunakan SAR-X Kriging menghasilkan Mean Absolute Percentage Error (MAPE) sebesar 3,82%, mengungguli metode SAR-X yaitu MAPE 4,68% dan metode Ordinary Kriging yaitu MAPE 3,86%. Secara praktis, hasil ini dapat meningkatkan kualitas prediksi curah hujan yang bermanfaat dalam pengelolaan sumber daya air, perencanaan pertanian, serta strategi mitigasi banjir di Pulau Jawa.

Kata Kunci: curah hujan; iklim; MAPE; SAR-X; SAR-X Kriging.

2020MSC: 62H11, 86A32

1. INTRODUCTION

Indonesia is a tropical region that receives sunlight throughout the year. This results in consistently high rainfall with complex variations. Observations by the Meteorology, Climatology, and Geophysics Agency (BMKG) show that Indonesia experienced 128 rainy days with an annual rainfall total of 1,562.6 mm in 2023. The high rainfall in tropical areas like Indonesia is generally caused by convection processes and the formation of warm rain clouds [1]. Meanwhile, the complex rainfall variation occurs because Indonesia lies on the equator, where the dry and rainy seasons are heavily influenced by the movement of monsoon winds [2]. In addition, rainfall variations in Indonesia are affected by climate phenomena such as El Niño and La Niña. El Niño is a climate anomaly characterized by reduced rainfall potential and increased air temperature, while La Niña is marked by increased rainfall potential above standard [3]. According to BMKG, La Niña can increase monthly rainfall accumulation in Indonesia by 20 to 40%, while El Niño can reduce rainfall in some regions of Indonesia by more than 40%.

Besides El Niño and La Niña, other climatic factors influence rainfall, such as air temperature, surface pressure, humidity, wind speed, and solar radiation. Fluctuations in air temperature are often closely associated with extreme rainfall events [4]. High temperatures can contribute to increased atmospheric evaporation, raising humidity in the atmosphere [5]. Under the right atmospheric conditions, high humidity can aid in forming rain clouds. Surface pressure also affects rainfall distribution, as pressure differences between land and sea can influence wind direction and speed [6]. Strong winds can carry moisture from the ocean to land, and intense solar radiation can heat the earth's surface, causing convection that ultimately leads to rainfall.

Extreme rainfall can significantly impact human activities in various sectors, particularly in agriculture and transportation [7]. In agriculture, adequate rain is key to successful food production; however, extreme rainfall can damage crops and lead to harvest failure. In the transportation sector, especially sea and air transport, heavy rain can reduce visibility and make transportation routes challenging, ultimately increasing the risk of accidents. Extreme rainfall can also lead to higher water transport costs and increased congestion, especially during peak hours [8].

Besides its impact on economic sectors, rainfall anomalies also increase the risk of natural disasters and public health crises. Excessive rainfall is a major contributor to flooding, damaging infrastructure and elevating human exposure to waterborne diseases [9]. Conversely, prolonged drought reduces river flow, leading to higher concentrations of pollutants and pathogens due to decreased water dilution [10]. Studies indicate that these conditions can heighten the risk of outbreaks

of pneumonia, colds, flu, diarrhea, viral fever, and typhoid [11]. According to the National Disaster Management Agency (BNPB), floods and droughts were the most frequent disasters in Indonesia in 2023, with 5,573 drought incidents and 8,331 flood events. Given these significant environmental and public health risks, developing accurate rainfall prediction models is essential for early warning systems and disaster mitigation efforts.

Predicting rainfall is crucial for water resource planning and disaster risk mitigation, particularly in regions prone to extreme rain [12]. However, rainfall in Indonesia exhibits dynamic spatial and temporal variations [13], posing challenges for accurate modeling. Additionally, limitations in measurement tools often result in missing or unobserved data at several observation points, necessitating the use of spatial interpolation methods to produce more representative and detailed rainfall data.

Conventional interpolation techniques, such as Ordinary Kriging (OK) and Inverse Distance Weighting (IDW), are widely applied for estimating rainfall at unsampled locations [14]. However, these methods primarily rely on spatial relationships and do not explicitly incorporate exogenous climatic variables, which may limit their predictive accuracy [15]. Some advanced kriging variants, such as Regression Kriging (RK) and Universal Kriging (UK), can integrate additional predictors, but their performance depends heavily on the availability and relevance of auxiliary data [16]. In contrast, spatial autoregressive models, such as the Spatial Autoregressive (SAR) and Spatial Durbin models, explicitly account for spatial dependencies by incorporating spatial lag structures [17]. While these models improve the understanding of spatial interactions, they are not inherently designed for interpolation at unsampled locations.

To overcome these challenges, this research proposes Spatial Autoregressive Exogenous Kriging (SAR-X Kriging), a novel hybrid approach that integrates the SAR-X model with kriging. The SAR-X model captures spatial dependencies and external climatic influences, while kriging provides interpolation capabilities for estimating rainfall at unsampled locations [18]. Unlike conventional interpolation techniques that rely solely on spatial proximity or autoregressive models that do not perform interpolation, SAR-X Kriging uniquely combines both spatial dependencies and exogenous climatic influences within a single predictive framework, offering a methodological advancement over existing spatial models. This research focuses on Java Island, Indonesia's most densely populated region, where accurate rainfall prediction is critical for effective water resource management and disaster mitigation. Specifically, this study introduces a novel integration of exogenous climatic variables into SAR-X Kriging, bridging the gap between spatial interpolation and temporal forecasting. By developing an optimal SAR-X model with key exogenous variables and applying it for rainfall forecasting, this research provides a more comprehensive and accurate representation of rainfall variability. The proposed approach is expected to enhance both spatial and temporal prediction accuracy, addressing limitations in existing methods and contributing to more reliable climate-informed decision-making in tropical regions.

2. METHOD

2.1. Research Data and Variables

This study involves factors that influence rainfall in the Java Island region. Data was obtained from the NASA website, with the last update in 2023. Data was obtained from 51 BMKG climate observation station locations spread across various districts/cities throughout Java Island. The observation stations are spread across 37 districts and 14 cities.

Based on several studies, rainfall is influenced by several factors. Those factors are including air temperature, humidity, pressure, and wind speed [19]. In addition, the latitude and longitude of a region also affect rainfall [20]. Therefore, we use the dependent variable under investigation (y) is rainfall in mm, and independent variables consist of surface pressure (x_1) in Kilopascal or kPa, air temperature (x_2) in Celsius, the percentage of humidity (x_3), wind speed (x_4) in km/h, and rate of solar radiation intensity (x_5). Besides that, we also observe the latitude and longitude of each station.

Heat maps were generated to depict variations and illustrate rainfall's spatial distribution and intensity across Java Island (Figure 1). Darker colors indicate higher rainfall intensity, providing an intuitive understanding of regional rainfall patterns. The map demonstrates clear spatial dependence, with distinct high and low rainfall clusters. Higher rainfall is concentrated in the western and central regions, such as Bandung City and Cianjur, while lower rainfall is observed in the eastern areas, including Sumenep and Situbondo. This pattern underscores the spatial dependence of rainfall, highlighting the need to account for spatial variability in effective planning and management strategies. Exploratory data analysis and descriptive statistics shown in Table 1 were conducted for each variable to comprehensively understand its characteristics, including mean, median, minimum, maximum, and variance values.

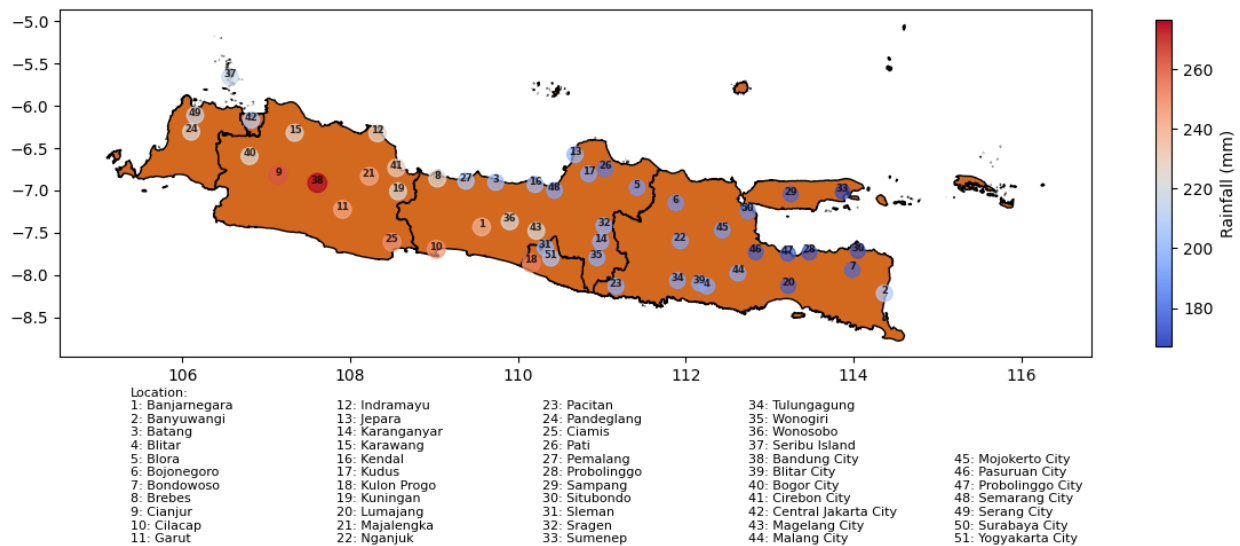


Figure 1. Heat map of rainfall intensity in Java island.

Table 1. Descriptive analysis of research data.

Variable	Minimum	Maximum	Average	Variance
Rainfall (y)	167.20	276.60	206.11	728.08
Surface Pressure (x_1)	90.96	100.97	98.15	5.20
Air Temperature (x_2)	21.54	28.15	25.90	2.60
Humidity (x_3)	76.01	88.72	82.35	13.26
Wind Speed (x_4)	1.42	3.52	2.14	0.37
Solar Radiation (x_5)	17.06	20.20	18.50	0.79

Table 1 presents the descriptive analysis of the research data, showing the minimum, maximum, average, and variance values for key climatic variables on Java Island. Rainfall (y) varies between

167.20 mm and 276.60 mm, with an average of 206.11 mm and a high variance of 728.08, indicating significant spatial variability. Surface pressure (x_1) and air temperature (x_2) show moderate variance, with averages of 98.15 kPa and 25.90°C, respectively. Humidity (x_3) averages at 82.35% with a moderate variance of 13.26. Wind speed (x_4) has the lowest variance (0.37) and an average of 2.14 km/h, reflecting more regional uniformity. Solar radiation (x_5) ranges from 17.06% to 20.20%, with an average of 18.50 % and low variance (0.79), suggesting relatively stable solar exposure. These statistics highlight the diverse climatic conditions across Java Island, emphasizing the need for spatially adaptive climate management strategies.

2.2. Spatial Weight Matrix

The spatial weight matrix is a matrix that describes the relationships between regions and is obtained based on distance or neighborhood information. The weights between locations vary depending on the distance between them; the closer the distance, the greater the weight, and the farther the distance, the smaller the weight. There are two approaches to forming this matrix: the contiguity approach and the distance approach [21]. In this research, the spatial weight matrix is constructed using the distance approach, precisely the inverse distance method with the Euclidean distance metric [22]. The Euclidean distance metric is chosen because it effectively captures the straight-line distance between geographic coordinates, allowing for a precise representation of spatial relationships between regions.

2.3. Spatial Dependency Test

A key assumption must be met with regional or spatial dependence in analyses involving spatial methods. This means heterogeneity between regions affects the value of the dependent variable, in this case, rainfall. Spatial dependency can be tested using Moran's I test and the Lagrange Multiplier test [23].

Moran's I test is the most commonly used method to assess and calculate general spatial autocorrelation. This method is employed to detect the presence of spatial autocorrelation in the data. If spatial autocorrelation exists, it indicates the presence of a pattern or trend forming clusters in space. Moran's I is calculated using the formula proposed by Anselin (1988) as follows [24]:

$$I = \frac{n \sum_{i=1}^n \sum_{j=1}^n W_{ij} (p_i - \bar{p})(p_j - \bar{p})}{\sum_{i=1}^n \sum_{j=1}^n W_{ij} \sum_{i=1}^n (p_i - \bar{p})^2}, \quad (1)$$

where I represent Moran's I index, n refers to the total number of observed locations, p_i and p_j are the observed variable values at locations i and j , $\bar{p}$ represents the mean of the observed variable values across all locations, and W_{ij} denotes the standardized spatial weight between regions i and j .

A positive Moran's I indicates a clustered spatial pattern, while a negative value suggests a dispersed pattern. The null hypothesis (H_0) states no spatial autocorrelation, while the alternative hypothesis (H_1) indicates spatial dependence. The Moran's I test statistic follows a standard normal distribution and can be calculated as follows [24]:

$$Z(I) = \frac{I - E(I)}{\sqrt{Var(I)}} \sim N(0,1), \quad (2)$$

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

$$Var(I) = \frac{n^2 S_1 - n S_2 + 3 S_0^2}{(n^2 - 1) S_0^2} - (E(I))^2, \quad (4)$$

$$S_0 = \sum_{i=1}^n \sum_{j=1}^n W_{ij}, \quad S_1 = \frac{1}{2} \sum_{i \neq j} (W_{ij} + W_{ji})^2, \quad S_2 = \sum_{i \neq j} (\sum_{i=1}^n W_{ij} + \sum_{j=1}^n W_{ji})^2, \quad (5)$$

where $Z(I)$ is the standardized test statistics for Moran's I, I represent Moran's I index, and $E(I)$ is the expected value of Moran's I. The decision for the statistical test in Equation (2) is made by rejecting H_0 at a significance level α if $-Z_{value} \leq -Z_{\frac{\alpha}{2}}$ or $-Z_{value} \geq Z_{\frac{\alpha}{2}}$ or if p -value $\leq \alpha$.

The Lagrange Multiplier test, commonly called the LM test, is used to select the most appropriate spatial regression model [21]. The LM test is also helpful in identifying the spatial effects in the data and determining whether there is lag dependency, error dependency, or both. The H_0 states no spatial dependence exists in the response variable, while the H_1 indicates spatial dependence in the response variable. The LM test statistic follows a Chi-Square distribution with 1 degree of freedom and can be calculated as follows:

$$LM_{Lag} = \frac{\left(\frac{\mathbf{e}^T \mathbf{W} \mathbf{y}}{\delta}\right)^2}{\left[\frac{(\mathbf{W} \mathbf{X} \hat{\boldsymbol{\beta}})^T \mathbf{M} (\mathbf{W} \mathbf{X} \hat{\boldsymbol{\beta}})}{\delta}\right] + [\text{tr}(\mathbf{W}^T \mathbf{W} + \mathbf{W}^2)]} \sim \chi^2(1), \quad (6)$$

$$\delta = \frac{\mathbf{e}^T \mathbf{e}}{N}, \quad (7)$$

$$\mathbf{M} = \mathbf{I} - \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T, \quad (8)$$

where LM is the Lagrange Multiplier test statistics, $\mathbf{e}$ is the residual vector, $\mathbf{W}$ is the spatial weight matrix, $\mathbf{y}$ is the dependent variable vector, δ is the mean square error, $\hat{\boldsymbol{\beta}}$ is the estimated parameter vector from the OLS, and $\mathbf{M}$ is the projection matrix. The decision for the statistical test in Equation (6) is made by rejecting H_0 at a significance level α if $LM_{Lag} > \chi^2_{(1, 1-\alpha)}$ or if p -value $\leq \alpha$.

2.4. Spatial Autoregressive Exogenous Model (SAR-X)

The Spatial Autoregressive model is a model that takes into account the spatial lag effect on its dependent variable. This model is also called the Mixed Regressive Autoregressive model because it combines a standard regression model with a spatial regression model on the dependent variable [25]. The SAR-X model is an extension of the SAR model that adds exogenous variables. In this case, the exogenous variables are those mentioned above: surface pressure, air temperature, humidity, wind speed, and solar radiation intensity. Mathematically, the SAR-X model can be written as follows [26]:

$$\mathbf{y} = \rho \mathbf{W} \mathbf{y} + \mathbf{X} \boldsymbol{\beta} + \boldsymbol{\varepsilon}, \text{ with } \boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2 \mathbf{I}), \quad (9)$$

where $\mathbf{y}$ is an $n \times 1$ vector of the dependent variable, ρ is the spatial lag coefficient of the dependent variable, $\mathbf{W}$ is an $n \times n$ spatial weight matrix, $\mathbf{X}$ is the matrix of independent variables, $\boldsymbol{\beta}$ is the vector of regression coefficient parameters, and $\boldsymbol{\varepsilon}$ is the error term with dimensions $n \times 1$. For the parameter ρ , it can be estimated as $\hat{\rho}$ using the following Equation (10):

$$\begin{aligned} (\mathbf{y} - \mathbf{X} \hat{\boldsymbol{\beta}})^T (\mathbf{W} \mathbf{y} (\mathbf{W} \mathbf{y})^T)^{-1} &= \rho (\mathbf{W} \mathbf{y})^T (\mathbf{W} \mathbf{y} (\mathbf{W} \mathbf{y})^T)^{-1}, \\ \hat{\rho} &= (\mathbf{y} - \mathbf{X} \hat{\boldsymbol{\beta}})^T (\mathbf{W} \mathbf{y} (\mathbf{W} \mathbf{y})^T)^{-1}. \end{aligned} \quad (10)$$

Next, the parameter $\boldsymbol{\beta}$ can be estimated as $\hat{\boldsymbol{\beta}}$ using the following Equation (11):

$$\begin{aligned}\beta(XX^T)(XX^T)^{-1} &= (y - \rho W y) X^T (XX^T)^{-1}, \\ \hat{\beta} &= (y - \rho W y) X^T (XX^T)^{-1}.\end{aligned}\quad (11)$$

2.5. Semivariogram

The semivariogram is half of the variogram, defined as the variance of the difference between the observed values at two locations separated by a distance h . It helps show the spatial correlation between measured data [27]. The experimental semivariogram at distance h can be written as follows:

$$\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(s_i + h) - Z(s_i)]^2, \quad (12)$$

where $\hat{\gamma}(h)$ is the semivariogram value at distance h , $Z(s_i)$ is the observed value at location s_i , $Z(s_i + h)$ is the observed value at location $s_i + h$, $N(h)$ is the number of point pairs separated by distance h , and h is the distance between two locations, namely s_i and $s_i + h$.

All possible pairs at distance h are calculated using the following Equation (13):

$$|h| = \sqrt{(s_i - s_j)^2 + (s_i - s_j)^2}. \quad (13)$$

The three most commonly used theoretical models are the spherical, Gaussian, and exponential models [27]. In the exponential model, the semivariogram increases steeply and reaches the sill value asymptotically. The Gaussian model is the quadratic form of the exponential model, resulting in a parabolic shape at shorter distances. Meanwhile, the Spherical Model assumes that spatial variability reaches a point where the variability becomes constant or stable, with no further increase after that point.

2.6. Ordinary Point Kriging

The Kriging method is a linear interpolation technique used to estimate values at unsampled locations based on the values from sampled surrounding locations. The Ordinary Kriging method provides the Best Linear Unbiased Estimator (BLUE), meaning that the estimate is unbiased and has minimum variance [28]. The kriging estimate for an unsampled location is obtained by summing the weighted values from surrounding locations. The semivariogram is used as an input in the kriging method to determine the optimal weights. The ordinary point kriging estimator is a linear combination of n data values from the regionalized variable around s_0 [29]:

$$\hat{Y}(s_0) = \sum_{i=1}^n w_i Y(s_i), \quad (14)$$

where $\hat{Y}(s_0)$ is the estimated value at the unsampled location, and w_i represents the weight factor at location i (1, 2, ..., n).

The parameter estimates from the SAR-X model are used as inputs for the kriging method in the SAR-X Kriging model. With these results, the SAR-X model's estimates at unsampled locations can be obtained and used as predictions for unsampled locations by incorporating the indicator variables at those locations [30].

2.7 Mean Absolute Percentage Error (MAPE)

MAPE measures the accuracy of a model's predictions for each variable. The principle of calculating MAPE is the percentage of the average absolute value of the residuals in each period divided by the actual value. The MAPE calculation is expressed in the following Equation (15) [31]:

$$MAPE = \left(\left[\frac{1}{n} \sum_{i=1}^n \left| \frac{y(s_i) - \hat{y}(s_i)}{y(s_i)} \right| \right] \right) \times 100\%, \quad (15)$$

where $\mathbf{y}(s_i)$ is the vector of actual data values at location s_i , $\hat{\mathbf{y}}(s_i)$ is the vector of predicted data values at location s_i , and n is the sample size.

The scale for assessing the accuracy of forecast criteria based on MAPE values is as follows [32]:

Table 2. Scale for criteria of forecast accuracy

MAPE	Remarks
<10%	Excellent forecast accuracy
10% - 20%	Good forecast accuracy
20% - 50%	Reasonable forecast accuracy
>50%	Poor forecast accuracy

The research methodology is systematically outlined in the flowchart presented in Figure 2. This flowchart provides a comprehensive visualization of the step-by-step process of data collection, variable selection, model construction, and analysis. Illustrating the methodological framework ensures clarity in understanding the sequence of procedures undertaken, from the initial data acquisition to the final evaluation of results.

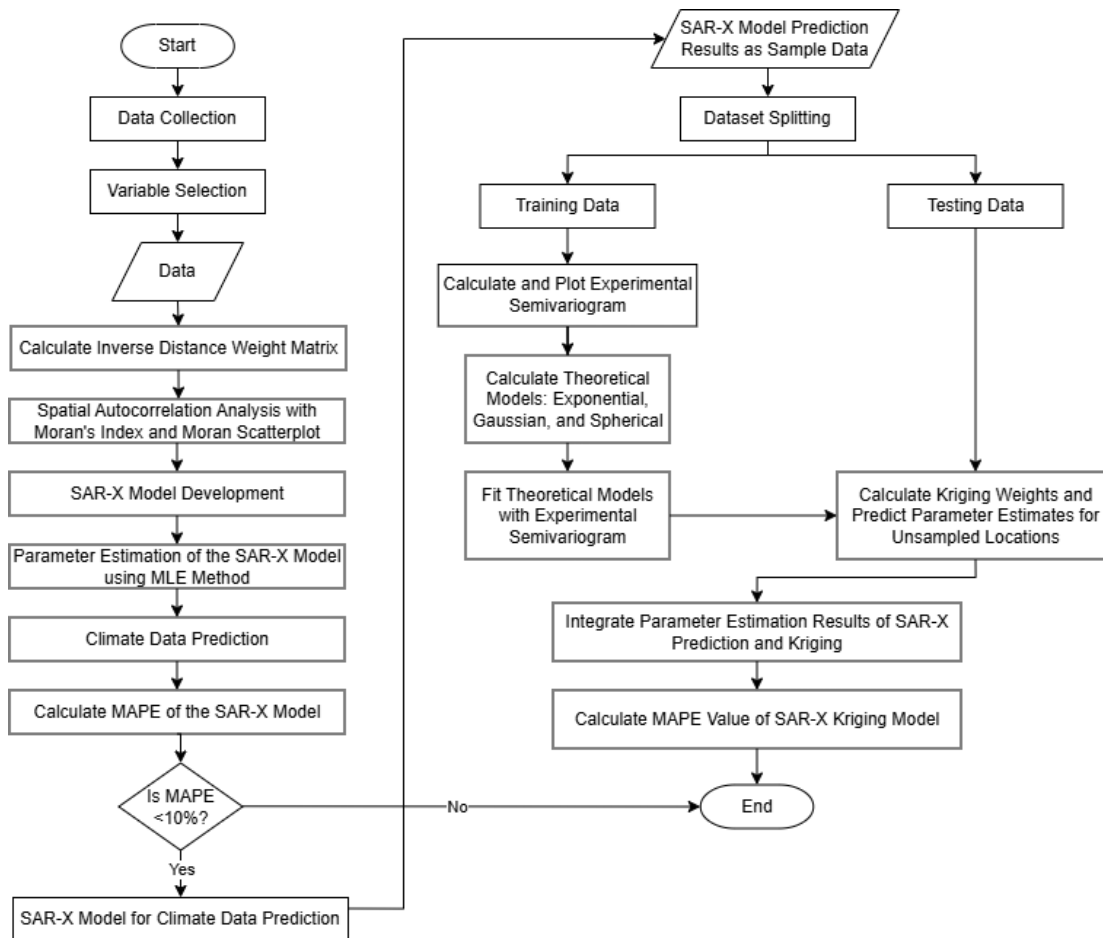


Figure 2. Flowchart of the research process.

3. RESULTS

3.1. Spatial Dependence Test

3.1.1. Moran's I Test

The Moran's I test determines whether observations at one location influence observations at nearby locations. Based on Table 3 of Moran's I test results, rainfall (y) has Moran's I value of 0.35624 with a p-value of 2.2×10^{-16} , indicating significant spatial autocorrelation in rainfall. This means that rainfall patterns tend to cluster in nearby locations, which can be crucial for understanding regional precipitation trends and water resource management. Spatial autocorrelation suggests that traditional statistical models assuming independent observations may not be appropriate. Instead, spatial models such as spatial lag or spatial error models should be considered to account for the spatial dependence on rainfall data. Surface pressure (x_1) has a Moran's I value of 0.03403 with a p-value of 0.03073, indicating spatial autocorrelation, although weaker than other variables. Surface pressure also shows a clustered pattern in nearby areas, suggesting that atmospheric pressure systems influence local climatic conditions. Air temperature (x_2) has a Moran's I value of 0.08647 with a p-value of 0.00013, indicating significant spatial autocorrelation in air temperature. This means that air temperature tends to be similar in nearby locations, which can affect studying heat wave propagation and regional climate modeling. Humidity (x_3) has a Moran's I value of 0.21615 with a p-value of 4.802×10^{-16} , indicating very significant spatial autocorrelation in humidity. This suggests that humidity shows a clustered distribution pattern, reinforcing the role of regional moisture dynamics in climate variability. Wind speed (x_4) has a Moran's I value of 0.00444 with a p-value of 0.2014, indicating no significant spatial autocorrelation in wind speed. This means that wind speed tends to be more randomly distributed without a clustered pattern, reflecting wind dynamics' transient and variable nature across different locations. Meanwhile, solar radiation (x_5) has a Moran's I value of 0.42082 with a p-value of 2.2×10^{-16} , indicating significant spatial autocorrelation in solar radiation. This suggests that solar radiation tends to cluster in nearby locations, which is an essential factor in solar energy potential assessments and climate studies.

Table 3. Result of Moran's I Test

Variable	Moran's I Test	p-value	Remarks
Rainfall (y)	0.35624	2.2×10^{-16}	Significant Spatial Autocorrelation
Surface Pressure (x_1)	0.03403	0.03073	Significant Spatial Autocorrelation
Air Temperature (x_2)	0.08647	0.00013	Significant Spatial Autocorrelation
Humidity (x_3)	0.21615	4.802×10^{-16}	Significant Spatial Autocorrelation
Wind Speed (x_4)	0.00444	0.2014	No Significant Spatial Autocorrelation
Solar Radiation (x_5)	0.42082	2.2×10^{-16}	Significant Spatial Autocorrelation

3.1.2. Lagrange Multiplier (LM) Test

The Lagrange Multiplier test detects spatial dependence, specifically through lag, error, or both (lag and error). The results of the Lagrange Multiplier test are shown in Table 4. Based on Table 4, the results of the Lagrange Multiplier test help identify the appropriate spatial model. According to this test, the Lagrange Multiplier (error) value is 0.1244 with a p-value of 0.7243. This high p-value indicates no substantial evidence to support the spatial error model, making this model statistically insignificant. On the other hand, the Lagrange Multiplier (lag) value is 6.0876 with a p-value of 0.0136. The p-value, being less than 0.05, suggests significant evidence supports the spatial lag model. Therefore, the spatial lag model is more suitable for this analysis than the spatial error model. In the context of rainfall

patterns, the significance of the spatial lag model suggests that rainfall in each location is influenced by rainfall in neighboring locations. This spatial dependence highlights the need for models that explicitly account for such relationships, as ignoring spatial dependence could lead to biased estimates and misleading conclusions. The preference for the spatial lag model indicates that weather systems or regional climatic conditions contribute to clustering rainfall patterns. Consequently, SAR-X modeling is necessary to appropriately capture these spatial dependencies and improve the accuracy of rainfall predictions.

Table 4. Lagrange Multiplier Test Results

Lagrange Multiplier Test	Test Value	p-value
Lagrange Multiplier (error)	0.1244	0.7243
Lagrange Multiplier (lag)	6.0876	0.0136

3.2. Spatial Autoregressive Exogenous (SAR-X) Model

After conducting the spatial dependency test, the SAR-X model is then implemented. Below are the output results from the R software for SAR-X modeling, including the respective parameter values for each. **Table 5 shows the** results of SAR-X modeling, and **Table 6 shows the** results of the overall SAR-X models. Based on Table 5, the equation of the SAR-X model is

$$\hat{y} = -137.796 + 0.6703Wy - 0.3530x_1 - 0.4684x_2 + 4.2816x_3 + 9.2701x_4 - 6.4662x_5.$$

Table 5. SAR-X modeling results.

	Estimate	Std. Error	z-value	p-value	Remarks
Intercept	-137.796	203.7655	-0.6762	0.4988	No Significant
ρ	0.6703	0.1825	-0.0513	0.9590	No Significant
Surface Pressure (x_1)	-0.3530	6.8811	-0.0513	0.9590	No Significant
Air Temperature (x_2)	-0.4684	15.3785	-0.0305	0.9757	No Significant
Humidity (x_3)	4.2816	3.0472	1.4051	0.1599	No Significant
Wind Speed (x_4)	9.2701	4.3928	2.1103	0.0348	Significant
Solar Radiation (x_5)	-6.4662	3.1659	-2.0424	0.0411	Significant

Table 6 presents the results of the overall SAR-X models generated, displaying key statistical parameters for each model. The table includes the model structure, intercept, parameter estimates for different explanatory variables x_1, x_2, x_3, x_4, x_5 , and the autoregressive coefficient ρ (rho). Additionally, the p-value for each model is provided to indicate statistical significance, with all models marked as "Significant." The intercept values vary across models, showing different baseline predictions. Each parameter estimate represents the effect of the corresponding explanatory variable on the dependent variable. The ρ values are consistently high, indicating strong spatial dependence. The p-values are all very low (mostly below 10^{-2}), supporting the statistical significance of the models. The remarks confirm that all models are considered significant based on these p-values. The table comprehensively compares different SAR-X models, helping to identify the best-fitting model for the given data.

3.3. Selection of the Best Model

After estimating the parameters for each model, the next step is to determine the best model. In some applied settings or stepwise procedures, a smaller p-value is often taken to indicate stronger evidence against the null hypothesis (i.e., that a given parameter is zero), thus suggesting that the corresponding

predictor is statistically significant in explaining variations in the dependent variable [33]. Table 7 presents the best model identified.

Table 6. Results of the Overall SAR-X Models

No	Model	Intercept	Parameter x_1	Parameter x_2	Parameter x_3	Parameter x_4	Parameter x_5	Rho	p-value	Remarks
1	$y \sim x_1$	379.9396	-3.7536					0.9464	9.51×10^{-9}	Significant
2	$y \sim x_2$	193.4474		-6.9874				0.9417	1.15×10^{-8}	Significant
3	$y \sim x_3$	-306.0706			3.99537			0.8908	8.91×10^{-7}	Significant
4	$y \sim x_4$	32.2666				-9.239		0.9417	4.15×10^{-8}	Significant
5	$y \sim x_5$	236.0153					-11.2536	0.8669	3.69×10^{-4}	Significant
6	$y \sim x_1 + x_2$	-497.9444	11.052	-21.9814				0.9173	1.67×10^{-7}	Significant
7	$y \sim x_1 + x_3$	-493.0477	1.3966		4.66074			0.8669	6.37×10^{-6}	Significant
8	$y \sim x_1 + x_4$	377.86062	-3.7292			-0.14221		0.9463	1.10×10^{-8}	Significant
9	$y \sim x_1 + x_5$	582.73313	-3.5547				-10.8394	0.8403	7.15×10^{-4}	Significant
10	$y \sim x_2 + x_3$	-505.5053		3.4728	5.4152			0.8546	1.46×10^{-5}	Significant
11	$y \sim x_2 + x_4$	227.7383		-8.9497		7.462		0.9443	6.04×10^{-9}	Significant
12	$y \sim x_2 + x_5$	369.8823		-6.1999			-9.2874	0.8199	1.06×10^{-3}	Significant
13	$y \sim x_3 + x_4$	-388.7415			4.82232	7.9540		0.8787	1.17×10^{-6}	Significant
14	$y \sim x_3 + x_5$	-150.1347			3.66271		-5.33813	0.7459	3.40×10^{-4}	Significant
15	$y \sim x_4 + x_5$	228.258				-5.4557	-10.2059	0.8673	3.40×10^{-4}	Significant
16	$y \sim x_1 + x_2 + x_3$	-384.6021	-7.9759	19.9619	8.3601			0.8176	1.08×10^{-4}	Significant
17	$y \sim x_1 + x_2 + x_4$	-493.84	11.672	-25.2974		9.4233		0.9206	7.18×10^{-8}	Significant
18	$y \sim x_1 + x_2 + x_5$	-175.521	8.0881	-17.5125			-6.6217	0.7967	1.51×10^{-3}	Significant
19	$y \sim x_1 + x_3 + x_4$	-454.3911	0.5434		5.0125	7.2749		0.8695	3.89×10^{-6}	Significant
20	$y \sim x_1 + x_3 + x_5$	-259.8476	0.6983		4.0485		-4.77122	0.7406	4.55×10^{-3}	Significant
21	$y \sim x_1 + x_4 + x_5$	696.8622	-4.6188			6.2958	-11.9617	0.8286	1.03×10^{-3}	Significant
22	$y \sim x_2 + x_3 + x_4$	-465.0124		1.5171	5.3363	6.9133		0.8636	7.42×10^{-6}	Significant
23	$y \sim x_2 + x_3 + x_5$	-286.8735		1.9658	4.5465		-4.5039	0.7343	5.16×10^{-3}	Significant
24	$y \sim x_2 + x_4 + x_5$	458.4678		-9.1338		11.323	-10.7784	0.7748	2.75×10^{-3}	Significant
25	$y \sim x_3 + x_4 + x_5$	-219.2123			4.6012	8.4847	-5.9333	0.6711	1.21×10^{-2}	Significant
26	$y \sim x_1 + x_2 + x_3 + x_4$	-397.3393	-4.8092	11.7285	7.1144	5.922		0.8421	4.73×10^{-5}	Significant
27	$y \sim x_1 + x_2 + x_3 + x_5$	-219.7454	-6.3516	15.3584	7.0608		-4.0003	0.7006	1.00×10^{-2}	Significant
28	$y \sim x_1 + x_2 + x_4 + x_5$	-98.4947	8.4574	-21.1302		11.844	-8.2813	0.7273	4.95×10^{-3}	Significant
29	$y \sim x_1 + x_3 + x_4 + x_5$	-137.3609	-0.5593		4.37159	9.2068	-6.43334	0.6694	1.23×10^{-2}	Significant
30	$y \sim x_2 + x_3 + x_4 + x_5$	-140.1878		-1.2441	4.1375	9.3603	-6.5107	0.6719	1.19×10^{-2}	Significant
31	$y \sim x_1 + x_2 + x_3 + x_4 + x_5$	-137.7962	-0.3530	-0.46842	4.28169	9.2701	-6.46629	0.6704	1.33×10^{-2}	Significant

Table 7. Best SAR-X Model

	Estimate	Std. Error	z-value	p-value	Remarks
Intercept	227.7383	41.7128	5.4597	4.77×10^{-8}	Significant
ρ	0.9442	0.0391	24.129	6.035×10^{-9}	Significant
Air Temperature (x_2)	-8.9497	1.8182	-4.9222	8.556×10^{-7}	Significant
Wind Speed (x_4)	7.4620	4.8306	1.5447	0.1224	No Significant

Thus, the equation of the SAR-X model is

$$\hat{y} = 227,7383 + 0,9442Wy - 8,9497x_2 + 7,4620x_4.$$

Based on the parameter estimation results for the SAR-X model, the smallest p-value was found in the SAR-X model (11), which is 6.0351×10^{-9} . This model indicates that the intercept, rho, and the variable that significantly affects rainfall is air temperature (x_2), with an estimated parameter value of -8.9497. Meanwhile, the variable that is not significant is wind speed (x_4), with an estimated parameter value of 7.4620.

3.4. Rainfall Prediction Using SAR-X

Based on the SAR-X prediction model at 51 observation locations, the predicted rainfall results influenced by exogenous factors such as air temperature, humidity, solar radiation, wind speed, and surface pressure are presented in Table 8 as follows:

Table 8. Rainfall prediction (mm) using the SAR-X model.

No	Observation Location	$\hat{y}$	No	Observation Location	$\hat{y}$
1	Banjarnegara Regency	215.851	27	Pemalang Regency	211.729
2	Banyuwangi Regency	199.862	28	Probolinggo Regency	175.172
3	Batang Regency	209.635	29	Sampang Regency	179.795
4	Blitar Regency	205.384	30	Situbondo Regency	176.345
5	Blora Regency	188.658	31	Sleman Regency	214.900
6	Bojonegoro Regency	184.927	32	Sragen Regency	196.052
7	Bondowoso Regency	204.713	33	Sumenep Regency	183.056
8	Brebes Regency	220.392	34	Tulungagung Regency	200.817
9	Cianjur Regency	234.752	35	Wonogiri Regency	212.381
10	Cilacap Regency	217.664	36	Wonosobo Regency	233.106
11	Garut Regency	245.839	37	Seribu Islands	207.123
12	Indramayu Regency	214.233	38	Bandung City	255.336
13	Jepara Regency	193.994	39	Blitar City	199.127
14	Karanganyar Regency	196.266	40	Bogor City	225.928
15	Karawang Regency	222.242	41	Cirebon City	211.640
16	Kendal Regency	204.964	42	Central Jakarta City	208.200
17	Kudus Regency	190.422	43	Magelang City	230.122
18	Kulonprogo Regency	208.959	44	Malang City	202.874
19	Kuningan Regency	225.883	45	Mojokerto City	184.547
20	Lumajang Regency	196.680	46	Pasuruan City	178.720
21	Majalengka Regency	243.951	47	Probolinggo City	175.784
22	Nganjuk Regency	190.765	48	Semarang City	195.100
23	Pacitan Regency	211.797	49	Serang City	205.897
24	Pandeglang Regency	215.088	50	Surabaya City	178.237
25	Pangandaran Regency	220.489	51	Yogyakarta City	212.173
26	Pati Regency	194.020			

MAPE is calculated using Equation (15) and the error calculations at each location. The error calculation shows that the MAPE value for the SAR-X model prediction is 5.076%, which indicates that the prediction accuracy is very high since the MAPE is less than 10%. According to Table 8, the highest monthly rainfall predictions, with values exceeding 210 mm, are found in the following regions: Banjarnegara Regency, Brebes Regency, Cianjur Regency, Cilacap Regency, Garut Regency, Indramayu Regency, Karawang Regency, Kuningan Regency, Majalengka Regency, Pacitan Regency, Pandeglang Regency, Pangandaran Regency, Pemalang Regency, Sleman Regency, Wonogiri Regency, Wonosobo Regency, Bandung City, Bogor City, Cirebon City, Magelang City, and Yogyakarta City.

3.5. Experimental Semivariogram and Fitting of Theoretical Model in Java Island

The predicted data using the SAR-X model are used as input for the kriging model, where 10 regencies/cities in Java Island with the closest distances to the data center are used as testing data, and the remaining 41 locations are used as training data. Prediction calculations using ordinary point kriging require the input of experimental semivariogram calculations. The experimental semivariogram values are calculated based on all possible distance pairs, where the distance function used is the Euclidean distance, and an experimental semivariogram plot is formed, as shown in Figure 3.

Based on Figure 3, the plot of the experimental semivariogram calculation is obtained, which is used for the theoretical model fitting process. The fitting process of the experimental semivariogram

with the theoretical model approach is determined by the minimum Sum of Squared Errors (SSE) among the three models. After conducting the theoretical calculations, the fitting between the experimental semivariogram and the theoretical model is performed.

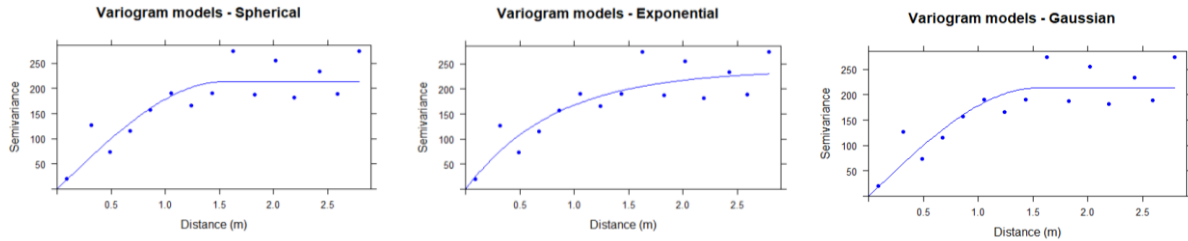


Figure 3. Experimental Semivariogram plot in Java island: Spherical model (left), Exponential model (center), and Gaussian model (right).

Table 9 shows the theoretical models in Java Island. The exponential model is the best of the three because it has the smallest SSE. This best theoretical model is then used as input in the calculation process of the ordinary point kriging method.

Table 9. Theoretical model in Java island

Model	SSE
Exponential	646427.6
Gaussian	1267426.0
Spherical	741826.4

3.6. Rainfall Prediction at Unsamped Locations in Java Island Using SAR-X Kriging

After obtaining the theoretical model from the semivariogram fitting process, this model is used as input for the ordinary point kriging method. The prediction is performed by splitting the sampled location data, which consists of 51 coordinate-based prediction samples, into 41 training data points and 10 testing data points. The testing data are unsampled observation locations that will be predicted using the ordinary point kriging method. The prediction at 10 unsampled observation locations using SAR-X Kriging can be seen in Table 10. Based on the error calculation results, the MAPE for the prediction using the SAR-X Kriging model is 3.82%, indicating that the prediction accuracy is very high, as the MAPE is less than 10%.

Table 10. Prediction at 10 unsampled observation locations using SAR-X Kriging

No.	Location	Latitude	Longitude	$\hat{y}$
1	Semarang City	-7.0022329	110.434226	202.475000
2	Magelang City	-7.4805408	110.217695	209.316900
3	Kendal Regency	-6.9326769	110.203074	204.544238
4	Sleman Regency	-7.6816700	110.323330	210.184609
5	Wonosobo Regency	-7.3679310	109.900846	211.342835
6	Yogyakarta City	-7.8004570	110.391280	210.733905
7	Kudus Regency	-6.8073946	110.840369	195.755013
8	Sragen Regency	-7.4202758	111.023247	202.787127
9	Kulonprogo Regency	-7.8596000	110.157900	211.304991
10	Karanganyar Regency	-7.6069193	110.984515	207.327185
MAPE				3.82%

3.7. Comparison of Prediction Results Kriging, SAR-X, and SAR-X Kriging

To ensure a fair comparison between the SAR-X, Ordinary Kriging, and SAR-X Kriging models, we focus on predictions at 10 unsampled observation locations. This approach provides a consistent basis for evaluating the predictive performance of each technique under identical conditions.

3.7.1. Rainfall Prediction Using SAR-X Method

The SAR-X method was used to predict rainfall at all 51 observation locations in Java Island. However, only the expected rainfall values at the selected 10 observation locations are considered for MAPE comparison. The prediction results using the SAR-X method are shown in Table 11. The MAPE value for the SAR-X method is 4.68%, indicating high prediction accuracy.

Table 11. Prediction at 10 unsampled observation locations using SAR-X

No.	Location	Latitude	Longitude	$\hat{y}$
1	Semarang City	-7.0022329	110.434226	195.100
2	Magelang City	-7.4805408	110.217695	230.122
3	Kendal Regency	-6.9326769	110.203074	204.964
4	Sleman Regency	-7.6816700	110.323330	214.900
5	Wonosobo Regency	-7.3679310	109.900846	233.106
6	Yogyakarta City	-7.8004570	110.391280	212.173
7	Kudus Regency	-6.8073946	110.840369	190.422
8	Sragen Regency	-7.4202758	111.023247	196.052
9	Kulonprogo Regency	-7.8596000	110.157900	208.959
10	Karanganyar Regency	-7.6069193	110.984515	196.266
MAPE				4.68%

3.7.2. Rainfall Prediction Using Ordinary Kriging Method

The Ordinary Kriging method was used to predict rainfall at 10 unsampled locations on Java Island. Kriging is a geostatistical interpolation technique that uses the spatial autocorrelation of data points to estimate unknown values. The rainfall data was first transformed using a logarithmic transformation to meet the normality assumption required by the Kriging method. This transformation ensures the data distribution is more normal, allowing the Kriging model to provide more accurate predictions. The technique relies on the semivariogram model, with the exponential model being selected as the best fit with the smallest SSE of 0.0002686902. The prediction results using the Ordinary Kriging method are shown in Table 12. The MAPE for the Kriging method is 3.86%, indicating a high level of prediction accuracy, which is slightly higher than the MAPE of the SAR-X Kriging method (3.82%).

Table 12. Prediction at 10 unsampled observation locations using Ordinary Kriging.

No.	Location	Latitude	Longitude	$\hat{y}$
1	Semarang City	-7.0022329	110.434226	198.7248
2	Magelang City	-7.4805408	110.217695	211.5173
3	Kendal Regency	-6.9326769	110.203074	200.7293
4	Sleman Regency	-7.6816700	110.323330	211.0061
5	Wonosobo Regency	-7.3679310	109.900846	220.0400
6	Yogyakarta City	-7.8004570	110.391280	210.0692
7	Kudus Regency	-6.8073946	110.840369	188.0655
8	Sragen Regency	-7.4202758	111.023247	191.4540
9	Kulonprogo Regency	-7.8596000	110.157900	218.8666
10	Karanganyar Regency	-7.6069193	110.984515	193.1760
MAPE				3.86%

3.7.3. Comparison of MAPE for Kriging, SAR-X, and SAR-X Kriging Methods

To assess the performance of each prediction method, we give a comprehensive comparison of the MAPE. This comparison aims to determine which method provides the most accurate rainfall predictions by evaluating the error prediction generated by the Kriging, SAR-X, and SAR-X Kriging methods. A lower MAPE indicates higher accuracy and reliability in predictive modeling, critical for practical applications in climate analysis and environmental management. The MAPE for each prediction method is compared in Table 13. The SAR-X Kriging method shows the lowest MAPE, proving its superior accuracy in rainfall prediction. This method effectively combines the strength of the spatial autoregressive model (SAR-X) with the precision of the Kriging interpolation technique. By incorporating exogenous factors such as air temperature and wind speed and accounting for spatial dependencies, the SAR-X Kriging method offers a more accurate prediction model for regional climate analysis.

Table 13. Comparison of MAPE for prediction methods

Method	MAPE (%)
Kriging	3.86
SAR-X	4.68
SAR-X Kriging	3.82

4. DISCUSSION

This research demonstrates that SAR-X Kriging serves as a robust alternative to conventional kriging methods, effectively addressing spatial dependencies in rainfall prediction. By incorporating several exogenous variables, SAR-X Kriging utilises spatial information and captures atmospheric dynamics that influence rainfall distribution more comprehensively. The influence of air temperature and solar radiation on rainfall patterns has been widely recognized in previous research. Studies utilizing the Generalized Space Time Autoregressive (GSTAR) model have shown that incorporating spatial and temporal dependencies significantly enhances temperature forecasting accuracy. Since air temperature is a critical determinant of precipitation, integrating spatial autoregressive models with exogenous variables further improves predictive performance in rainfall estimation [34].

Additionally, prior research has explored the prediction of wind speed using the Autoregressive Fractionally Integrated Moving Average (ARFIMA) model, highlighting the significance of advanced time-series techniques in meteorological forecasting. Wind speed, one of the key exogenous variables employed in SAR-X Kriging, plays a crucial role in shaping rainfall patterns by influencing cloud movement and precipitation intensity. Incorporating autoregressive models in climate studies reinforces the necessity of spatial regression techniques in improving prediction accuracy [35]. These findings align with previous studies employing the Expanded Spatial Durbin Model with Ordinary Kriging (ESDMOK), which confirmed that air temperature and wind speed are critical factors in identifying spatial rainfall patterns and enhancing prediction accuracy. However, unlike earlier studies that primarily focused on spatial interpolation, the present research integrates exogenous climatic factors, improving prediction reliability and offering more profound insights into rainfall variability. This suggests that conventional kriging methods, which do not account for spatial dependencies in climatic variables, may underestimate the complexity of rainfall distribution [36].

The methodological strengths of this research lie in its ability to handle missing or sparse observational data while maintaining high predictive accuracy. This makes SAR-X Kriging particularly

valuable for water resource management, disaster mitigation, and agricultural planning applications, where rainfall data is critical for decision-making [37]. However, the research also has certain limitations. The limited spatial coverage of observation points (51 locations on Java Island) may reduce the model's accuracy in predicting rainfall in less sampled regions. Additionally, excluding factors such as soil moisture and land cover changes may limit the model's adaptability to long-term climate shifts [38]. Future research could improve the model by incorporating additional exogenous variables, such as soil moisture, atmospheric circulation indices, or topographic features, which may refine rainfall predictions. Applying higher-resolution satellite data and machine learning-based spatial models could also enhance the model's adaptability to evolving climatic conditions.

Furthermore, testing the SAR-X Kriging approach in other tropical regions would help assess its generalizability beyond Java Island. This research demonstrates that SAR-X Kriging is a promising approach for improving rainfall prediction by integrating spatial dependence and exogenous climatic factors. Its application can be extended to regions with complex rainfall patterns, providing more accurate and reliable forecasts for climate-sensitive sectors.

5. CONCLUSIONS

This research confirms that rainfall in Java Island exhibits spatial dependence, which is effectively captured using the Spatial Autoregressive Exogenous (SAR-X) model. The best SAR-X model, incorporating air temperature and wind speed, achieved a high predictive accuracy with a p-value of 6.0351×10^{-9} . Among the three methods tested, the SAR-X Kriging method provided the highest prediction accuracy with a MAPE of 3.82%, followed by the Ordinary Kriging method with a MAPE of 3.86%, and the SAR-X model alone with a MAPE of 4.68%.

The superior performance of the SAR-X Kriging method demonstrates its effectiveness in capturing spatial dependence and integrating exogenous climatic variables. This accuracy is particularly beneficial for practical agriculture, transportation, and disaster management applications. Accurate rainfall predictions can help optimize irrigation schedules, enhance infrastructure resilience to extreme weather, and improve flood early warning systems, supporting better decision-making and risk management in weather-sensitive sectors.

Despite its promising results, this study has several limitations, including the limited spatial coverage of observation points (51 locations in Java Island) and the exclusion of potential influential variables such as soil moisture, atmospheric circulation indices or topographic features. Future research could integrate more diverse exogenous variables, such as soil moisture, atmospheric circulation indices, or topographic features, and explore advanced machine learning techniques to improve prediction accuracy and model robustness.

ACKNOWLEDGEMENTS

The authors would like to thank the Rector of Padjadjaran University, the Director of DRPM Unpad, and the Center for Modeling and Computation Studies FMIPA Unpad for providing research funds for the dissemination of lecturers' and students' research results through the Academic Leadership Grant with contract number 1425/UN6.3.1/PT.00/2024.

REFERENCES

- [1] I. Dainty, S. H. Abdullah, and A. Priyati, "Analisis Peluang Curah Hujan untuk Penetapan Pola

- dan Waktu Tanam serta Pemilihan Jenis Komoditi yang Sesuai di Desa Masbagik Kecamatan Masbagik Kabupaten Lombok Timur,” *J. Ilm. Rekayasa Pertan. dan Biosist.*, vol. 4, no. 1, pp. 207–216, 2016.
- [2] L. Y. Worku, A. Mekonnen, and C. J. Schreck, “Diurnal cycle of rainfall and convection over the Maritime Continent using TRMM and ISCCP,” *Int. J. Climatol.*, vol. 39, no. 13, pp. 5191–5200, 2019, doi: 10.1002/joc.6121.
 - [3] B. Yuniasih, W. N. Harahap, and D. A. S. Wardana, “Anomali Iklim El Nino dan La Nina di Indonesia pada 2013-2022,” *AGROISTA J. Agroteknologi*, vol. 6, no. 2, pp. 136–143, 2022, doi: 10.55180/agi.v6i2.332.
 - [4] S. Sharma and P. P. Mujumdar, “On the relationship of daily rainfall extremes and local mean temperature,” *J. Hydrol.*, vol. 572, no. 2019, pp. 179–191, 2019, doi: 10.1016/j.jhydrol.2019.02.048.
 - [5] T. P. Roderick, C. Wasko, and A. Sharma, “Atmospheric Moisture Measurements Explain Increases in Tropical Rainfall Extremes,” *Geophys. Res. Lett.*, vol. 46, no. 3, pp. 1375–1382, 2019, doi: 10.1029/2018GL080833.
 - [6] C. Zhang, S. Wu, T. Li, Z. Yu, and J. Bian, “Interpreting the Trends of Extreme Precipitation in Florida through Pressure Change,” *Remote Sens.*, vol. 14, no. 6, pp. 1–12, 2022, doi: 10.3390/rs14061410.
 - [7] S. Mostamandi *et al.*, “Sea Breeze Geoengineering to Increase Rainfall over the Arabian Red Sea Coastal Plains,” *J. Hydrometeorol.*, vol. 23, no. 1, pp. 3–24, 2022, doi: 10.1175/JHM-D-20-0266.1.
 - [8] M. J. Koetse and P. Rietveld, “The impact of climate change and weather on transport: An overview of empirical findings,” *Transp. Res. Part D Transp. Environ.*, vol. 14, no. 3, pp. 205–221, 2009, doi: 10.1016/j.trd.2008.12.004.
 - [9] A. F. Prein and A. J. Heymsfield, “Increased melting level height impacts surface precipitation phase and intensity,” *Nat. Clim. Chang.*, vol. 10, no. 8, pp. 771–776, 2020, doi: 10.1038/s41558-020-0825-x.
 - [10] M. Kim, Y. Kienast, J. K. Hatt, A. E. Kirby, and K. T. Konstantinidis, “Metagenomics indicate that public health risk may be higher from flooding following dry versus rainy periods,” *Environmental Microbiol.*, vol. 14, no. 2, pp. 265–273, 2022, doi: <https://doi.org/10.1111/1758-2229.13047>.
 - [11] A. J. De Roos *et al.*, “Heavy precipitation, drinking water source, and acute gastrointestinal illness in Philadelphia, 2015-2017,” *PLoS One*, vol. 15, no. 2, pp. 2015–2017, 2020, doi: 10.1371/journal.pone.0229258.
 - [12] J. C. Semenza, “Cascading risks of waterborne diseases from climate change,” *Nat. Immunol.*, vol. 21, no. 5, pp. 484–487, 2020, doi: 10.1038/s41590-020-0631-7.
 - [13] H. Kuswanto, D. Setiawan, and A. Sopaheluwakan, “Clustering of Precipitation Pattern in Indonesia Using TRMM Satellite Data,” *Eng. Technol. Appl. Sci. Res.*, vol. 9, no. 4, pp. 4484–4489, 2019, doi: 10.48084/etasr.2950.
 - [14] K. Pakoksung, “Impact of Spatial Rainfall Scenarios on River Basin Runoff Simulation a Nan River Basin Study Using the Rainfall-Runoff-Inundation Model,” *Eng.*, vol. 5, no. 1, pp. 51–69, 2024, doi: 10.3390/eng5010004.
 - [15] H. Chen *et al.*, “A spatiotemporal estimation method for hourly rainfall based on F-SVD in the recommender system,” *Environ. Model. Softw.*, vol. 144, p. 105148, 2021, doi: 10.1016/j.envsoft.2021.105148.
 - [16] S. Shadeed, A. Jayyousi, A. Khader, C. Chwala, and H. Kunstmann, “Comparative analysis of

- interpolation methods for rainfall mapping in the Faria catchment, Palestine,” *An-Najah Univ. J. Res.*, vol. 36, no. 1, pp. 1–20, 2022.
- [17] J. Lee, P. C. B. Phillips, and F. Rossi, “Consistent Misspecification Testing in Spatial Autoregressive Models,” *ERN Cross-Sectional Model*, 2020, doi: <https://doi.org/10.2139/ssrn.3683830>.
- [18] A. N. Falah, B. N. Ruchjana, A. S. Abdullah, and J. Rejito, “Clustering spatial autoregressive kriging model for climate: A bibliometric analysis approach,” *Int. J. Data Netw. Sci.*, vol. 7, no. 2, pp. 637–646, 2023, doi: 10.5267/j.ijdns.2023.3.008.
- [19] A. Luthfiarta, A. Febriyanto, H. Lestiawan, and W. Wicaksono, “Analisa Prakiraan Cuaca dengan Parameter Suhu, Kelembaban, Tekanan Udara, dan Kecepatan Angin Menggunakan Regresi Linear Berganda,” *JOINS (Journal Inf. Syst.*, vol. 5, no. 1, pp. 10–17, 2020, doi: 10.33633/joins.v5i1.2760.
- [20] O. C. Satya, M. Arsali, A. K. Affandi, and P. M. Mandailing, “Spatial distribution and diurnal characteristics of rainfall in South Sumatra and surrounding areas based on Tropical Rainfall Measuring Mission (TRMM) data,” *J. Phys. Conf. Ser.*, vol. 1568, no. 1, 2020, doi: 10.1088/1742-6596/1568/1/012024.
- [21] J. P. LeSage, *The Theory and Practice of Spatial Econometrics*. 1999.
- [22] K. Kopczewska, *Applied Spatial Statistics and Econometrics*, 1st ed. London: Routledge, 2020.
- [23] A. Anwar, “Spatial Analysis of Regional Poverty in Central Java Indonesia,” *J. Din. Ekon. Pembangunan*, vol. 5, no. 1, pp. 36–55, 2022, doi: 10.14710/jdep.5.1.36-55.
- [24] L. Anselin, “Spatial Econometrics: Methods and Models,” *J. Am. Stat. Assoc.*, 1988.
- [25] L. Anselin, *Spatial Econometrics: Methods and Models*, 1st ed. Springer Dordrecht, 1988.
- [26] L. Anselin, “Spatial Econometrics,” in *A Companion to Theoretical Econometrics*, B. H. Baltagi, Ed. Blackwell Publishing, 2003.
- [27] Y. Chen, Y. Yang, and W. Wu, “Coal Seam Thickness Prediction based on Least Squares Support Vector Machines and Kriging Method,” *Electron. J. Geotech. Eng.*, vol. 20, no. 1, pp. 167–176, 2015.
- [28] M. Armstrong, *Basic Linear Geostatistics*, 1st ed. Springer Berlin, Heidelberg, 1998.
- [29] A. Setiawan and R. Rosadi, “Spasial Data Mining Menggunakan Model SAR-Kriging,” *IJCCS (Indonesian J. Comput. Cybern. Syst.*, vol. 5, no. 3, p. 52, 2011, doi: 10.22146/ijccs.5213.
- [30] A. S. Abdullah, “Implementasi Spasial Data Mining menggunakan Model Spatial Autoregressive-Kriging (SAR-Kriging),” Universitas Gadjah Mada, 2009.
- [31] W. W. S. Wei, *Time Series Analysis and Multivariate Methods*. Canada : Addison Wesley Publishing Company, 2006.
- [32] C. D. Lewis, *Industrial and Business Forecasting Methods*. London: Butterworths, 1982.
- [33] D. N. Gujarati and D. C. Porter, *Basic Econometrics*, 5th ed. New York: McGraw-Hill, 2009.
- [34] A. Aprianti, N. Faulina, and M. Usman, “Generalized Space Time Autoregressive (GSTAR) Model for Air Temperature Forecasting in the South Sumatera, Riau, and Jambi Provinces,” *Inpr. Indones. J. Pure Appl. Math.*, vol. 6, no. 1, pp. 1–13, 2024, doi: 10.15408/inprime.v6i1.36049.
- [35] D. I. Octaviyani, M. Y. Wijaya, and N. Fitriyati, “Estimation Parameter d in Autoregressive Fractionally Integrated Moving Average Model in Predicting Wind Speed,” *Inpr. Indones. J. Pure Appl. Math.*, vol. 1, no. 2, 2019, doi: 10.15408/inprime.v1i2.13676.
- [36] J. A. Acero, P. Kestel, H. T. Dang, and L. K. Norford, “Impact of rainfall on air temperature, humidity and thermal comfort in tropical urban parks,” *Urban Clim.*, vol. 56, no. June, p. 102051, 2024, doi: 10.1016/j.uclim.2024.102051.
- [37] A. N. Falah, B. N. Ruchjana, A. S. Abdullah, and J. Rejito, “The Hybrid Modeling of Spatial

- Autoregressive Exogenous Using Casetti's Model Approach for the Prediction of Rainfall," *Mathematics*, vol. 11, no. 17, pp. 1–21, 2023, doi: 10.3390/math11173783.
- [38] M. Abedini, M. A. Md Said, and F. Ahmad, "Integration of statistical and spatial methods for distributing precipitation in tropical areas," *Hydrol. Res.*, vol. 44, no. 6, pp. 982–994, 2013, doi: 10.2166/nh.2012.159.