Mathematical Journal of Modelling and Forecasting
Vol. 3, No. 2, December 2025, pp. 26-35

# Statistical Modelling of Rainfall Data Using Robust Kriging with Gaussian Semivariogram in Bengkulu Province

Cinta Rizki Oktarina[1,*] and Reza Pahlepi[2]

[1]Department of Mathematics, Batam Institute of Technology, Batam, Indonesia
[2]Department of Statistics, University of Bengkulu, Bengkulu, Indonesia
[*]cinta@iteba.ac.id

**Abstract.** This study aims to predict rainfall in Bengkulu Province for January 2024 using the Robust Kriging method, an advanced geostatistical approach designed to handle outliers and non-ideal spatial characteristics. The novelty of this study lies in integrating Robust Kriging with a Gaussian semivariogram for short-term rainfall prediction in Bengkulu Province. This combination has not been explored in previous hydrometeorological studies. Rainfall data were obtained from the Meteorology, Climatology, and Geophysics Agency (BMKG) and analysed to identify spatial dependency and variation. The analysis began with descriptive statistics, assumption testing, and outlier detection, followed by the construction of robust empirical and theoretical semivariogram models. Three semivariogram models, Spherical, Exponential, and Gaussian, were compared to determine the most suitable model based on Mean Squared Error (MSE) and Mean Absolute Percentage Error (MAPE) values. The results indicate that the Gaussian model produced the smallest MSE and MAPE values, showing the best fit to the empirical semivariogram. The Robust Kriging interpolation generated spatial predictions of rainfall intensity across Bengkulu, showing higher rainfall in the north and lower rainfall in the south. The findings demonstrate that Robust Kriging effectively improves prediction accuracy by minimizing the influence of outliers and optimizing spatial weighting. These results provide valuable insights for water resource management, agricultural planning, and hydrometeorological disaster mitigation in Bengkulu Province.

**Keywords:** Robust Kriging, Rainfall Prediction, Gaussian Semivariogram, Spatial Interpolation, Bengkulu Province.

## 1 Introduction

Global warming has resulted in a progressive increase in Earth's average surface temperature due to the continuous accumulation of greenhouse gases, such as carbon dioxide, methane, and nitrous oxide, in the atmosphere. This phenomenon has been recognized as one of the most critical environmental challenges of the 21st century, leading not only to higher global temperatures but also to significant disruptions in climatic and hydrological systems. These disruptions manifest as erratic weather patterns, extreme precipitation events, and prolonged dry spells. Such climatic anomalies, including heavy rainfall during traditionally dry seasons and extended droughts during rainy periods, have become increasingly common across tropical regions. The growing unpredictability of seasonal transitions has had profound implications for agriculture, fisheries, and water resource management, ultimately threatening food security and ecosystem stability. This underscores the urgent need for evidence-based approaches to monitor, model, and mitigate the impacts of global climate change [1].

Bengkulu Province, situated along the western coast of Sumatra Island, Indonesia, provides a relevant case for studying such climatic variability. The province experiences a humid tropical climate, heavily influenced by the Indian Ocean and the prevailing monsoon systems. Based on 2023 rainfall data, the highest precipitation was recorded in July, at approximately 650 mm, whereas September and October were the driest months, with several stations reporting near-zero rainfall. The spatial and temporal variability in precipitation across the province is primarily attributed to its complex topography and its direct exposure to maritime atmospheric circulation [2]. These conditions make Bengkulu highly susceptible to both hydrometeorological disasters, such as floods and landslides, during the wet season, and drought-induced agricultural losses during the dry season.

Given these circumstances, accurate spatial prediction of rainfall intensity is vital for supporting regional planning, water resource allocation, and disaster mitigation strategies. However, rainfall measurements are typically obtained from a limited number of stations, resulting in spatial gaps in data coverage. In this context, *Ordinary Kriging* serves as a powerful geostatistical interpolation technique for estimating rainfall at unsampled locations using spatial autocorrelation and variogram models [3]. Kriging, as a best linear unbiased estimator, accounts for spatial dependence among observed data points to produce optimal predictions. Several variants of this technique have been developed, including *Indicator Kriging*, *Simple Kriging*, *Ordinary Kriging*, and *Co-Kriging*, each offering specific advantages depending on the data structure and modeling objectives [4].

Nonetheless, conventional Kriging assumes that the underlying data are normally distributed [5], whereas environmental data such as rainfall often contain outliers or exhibit skewed distributions [6]. Such irregularities can distort the variogram and lead to biased predictions. To address these limitations, *Robust Kriging* has been introduced as a modification that reduces the influence of outliers and enhances model stability. This approach refines variogram estimation using robust statistical methods, thereby improving spatial prediction accuracy in heterogeneous datasets. By applying these geostatistical techniques, rainfall prediction in Bengkulu Province can be spatially refined, providing a more precise representation of regional precipitation dynamics. The resulting rainfall distribution maps can serve as valuable inputs for hydrological modeling, agricultural planning, and infrastructure design. Ultimately, this study contributes to strengthening climate resilience and supporting data-driven decision-making for sustainable environmental management in the province.

## 2 Research Methods

### 2.1 Data Source

This study employed historical rainfall data for January from the Meteorology, Climatology, and Geophysics Agency (BMKG) in Bengkulu Province as the primary data source. The research variable was rainfall amount (in millimeters), and each observation point had known geographic coordinates used as spatial references for rainfall distribution. The dataset consisted of rainfall measurements from 97 BMKG observation stations distributed across Bengkulu Province, with denser coverage in the coastal region and fewer stations in inland northern areas. The analysis was performed using R software with the gstat, geoR, and sp packages for semivariogram modeling and Kriging interpolation.

The following steps describe the data analysis procedure using the Robust Kriging method:

1. Descriptive Statistical Analysis.
2. Kriging Assumption Testing.
3. Stationarity Assessment.
4. Identification of Spatial Outliers.
5. Robust Semivariogram Estimation.
6. Kriging Weight Calculation.
7. Robust Estimation with Weighted Median.
8. Outlier Transformation.
9. Data Interpolation.

### 2.2 Spatial Data

Spatial data represent information describing the geographic location of phenomena on the Earth’s surface, such as satellite imagery and location-based services [7]. These data are essential for understanding environmental and urban spatial patterns, enabling the analysis of land use, pollution, and climate variability [8]. In spatial statistics, a spatial outlier is defined as a data point that significantly deviates from the values of its neighboring locations [9]. Such anomalies often indicate localized environmental disturbances, pollutant leaks, or other unique phenomena that are important to detect for effective environmental monitoring and management [10], [11].

## 2.3 Spatial Interpolation

Spatial interpolation is a technique used to estimate the value of a variable at unsampled locations based on measurements from sampled points. This method is fundamental in spatial data analysis, particularly in environmental and climatological studies where observations are often limited in number or irregularly distributed. Spatial interpolation methods are generally classified into deterministic and stochastic approaches. Deterministic methods, such as Inverse Distance Weighting (IDW) and trend surface analysis, rely on mathematical functions to predict values directly from known data. In contrast, stochastic methods incorporate spatial autocorrelation and uncertainty into the estimation process, allowing for statistical assessment of prediction errors [12]. Among stochastic approaches, Kriging is considered the most robust interpolation method because it uses the semivariogram to quantify spatial variation between data pairs. This model describes how similarity decreases with distance and helps determine optimal weights for interpolation, producing predictions with the smallest estimation variance [13]. Kriging not only provides an unbiased estimator but also quantifies uncertainty, making it widely applied in hydrology, meteorology, and environmental monitoring.

## 2.4 Robust Kriging

Kriging is an interpolation method in geostatistics that uses known data points to estimate values at unsampled locations [14]. This method enables spatial prediction based on the dependence between locations by optimizing the variogram, making it effective in addressing spatial correlation and data clustering [5]. However, standard Kriging is sensitive to outliers in the dataset. To overcome this limitation, Robust Kriging was developed as an extension of Ordinary Kriging by transforming the variogram weights to be more resistant to outliers. According to [15], Robust Kriging provides more accurate estimation in datasets containing outliers by reducing the influence of extreme values.

The steps for constructing a Robust Kriging model are as follows:

1. Calculate the robust semivariogram and fit the model with the smallest Mean Squared Error (MSE).
2. Compute the Kriging weights using the robust semivariogram and predict the value at location $i$ using:
$$\hat{Z}(s_i) = \sum_{j=1}^{n} \lambda_j Z(s_j) \quad (1)$$
where $\lambda_j$ are Kriging weights and $\sigma_k^2$is the Kriging variance.
3. Obtain the robust estimated value using the weighted median approach:
$$\bar{Z}(s_i) = weighted\ median\ \left(\{Z(s_j), i \neq j\}; \{\lambda_{ij}, i \neq j\}\right) \quad (2)$$
4. The observation $Z(s_i)$identified as a spatial outlier is transformed using the winsorized version:
$$Z^{(e)}(s_i) = \begin{cases} \bar{Z}(s_i) + c\sigma_k & if\ Z(s_i) - \bar{Z}(s_i) > c\sigma_k \\ Z(s_i) & if\ |Z(s_i) - \bar{Z}(s_i)| \leq c\sigma_k \\ \bar{Z}(s_i) - c\sigma_k & if\ Z(s_i) - \bar{Z}(s_i) < -c\sigma_k \end{cases} \quad (3)$$
where $1.5 \leq c \leq 2$. The transformed value $Z^{(e)}(s_i)$ represents the robust edited value.
5. Perform interpolation using Ordinary Kriging on the robust semivariogram and transformed data:
$$\hat{Z}^*(s_j) = \sum_{i=1}^{n} \lambda_{ij} Z^{(e)}(s_i) \quad (4)$$

The variogram measures spatial variation between locations and indicates the degree of spatial dependence among observations. It is defined as [15]:
$$2\gamma(h) = E[\{Z(s_i + h) - Z(s_i)\}^2] \quad (5)$$
where:

$s_i$ : spatial point $i$
$h$ : distance between two points
$Z(s_i)$ : observed value at spatial point $s_i$

The semivariogram is half the variogram function and expresses spatial autocorrelation [16]:
$$\gamma(h) = \frac{1}{2} E[\{Z(s_i + h) - Z(s_i)\}^2] \quad (6)$$
where:

$s_i$ : spatial point $i$
$h$ : distance between two points
$Z(s_i)$ : observed value at spatial point $s_i$

The empirical semivariogram is derived from observed data and plotted as a function of distance [6]:

$$\gamma(h) = \frac{1}{2|N(h)|}\sum_{i=1}^{N(h)}[\{Z(s_i + h) - Z(s_i)\}^2] \tag{7}$$

where:
$Z(s_i + h)$ : observation value at point $s_i + h$
$Z(s_i)$ : observation value at spatial point $s_i$
$N(h)$ : set of data pairs $s_i$and $s_i + h$separated by distance $h$
$|N(h)|$ : number of data pairs within the set $N(h)$.

Model fitting is conducted using binning to group distance intervals and identify the spatial trend. Common theoretical semivariogram models include:

1. Spherical Model
   The semivariogram for the Spherical Model is formulated as follows:

$$\gamma(h) = \begin{cases} \sigma^2\left(\frac{3h}{2r}\right) - \frac{h^3}{2r^3} & \text{, for } h \le r \\ \sigma^2 & \text{, other} \end{cases} \tag{8}$$

   where:
   $\sigma^2$: sill, representing the maximum semivariance value,
   $h$: distance between sample locations,
   $r$: range, representing the distance at which the semivariogram reaches the sill.
2. Exponential Model
   The semivariogram for the Exponential model is formulated as follows:

$$\gamma(h) = \sigma^2\left(1 - exp\left(\frac{3h}{r}\right)\right) \tag{9}$$

3. Gaussian Model
   The semivariogram for the Gaussian model is formulated as follows

$$\gamma(h) = \sigma^2\left(1 - exp\left(\frac{3h}{r}\right)^2\right) \tag{10}$$

The semivariogram describes the variability of data as a function of distance. The Spherical Model exhibits a linear increase until reaching the range, the Exponential Model rises rapidly before gradually approaching a plateau, and the Gaussian Model increases smoothly toward the sill. The selection of an appropriate model depends on the data's pattern. According to [15], the robust semivariogram is designed to minimize the negative influence of outliers by reducing the effect of squared differences in semivariance calculations. The following equation expresses the robust semivariogram:

$$\gamma(h) = \frac{\left\{\frac{1}{2|N(h)|}\sum_{i=1}^{N(h)}\left[|Z(s_i + h) - Z(s_i)\}|^{\frac{1}{2}}\right]\right\}^4}{\left(0.457 + \frac{0.494}{|N(h)|}\right)} \tag{11}$$

where:
$Z(s_i + h)$ : observation value at point $s_i + h$
$Z(s_i)$ : observation value at spatial point $s_i$
$N(h)$ : set of paired data points separated by distance $h$
$|N(h)|$ : number of pairs within the set $N(h)$

# 3 Results and Discussion

## 3.1 Spatial Distribution of Rainfall

Fig 1 shows the distribution of rainfall observation points in Bengkulu Province based on geographic coordinates. The points represent rainfall measurement locations, with color intensity indicating rainfall amounts at each site. The data distribution pattern reveals variations in rainfall across different locations, which can be further utilized for spatial analysis, such as interpolation or the identification of areas with the highest and lowest rainfall.

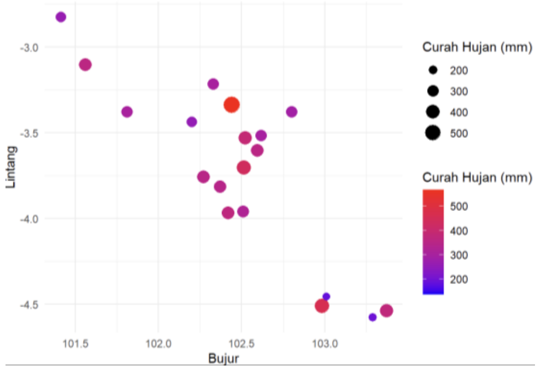


**Fig 1.** Rainfall Distribution Pattern Map

### 3.2 Testing of Kriging Assumption

**Table 1.** Testing of Robust Kriging Assumptions

| Test | Significance | Conclusion |
|---|---|---|
| Normality | 0.5361 | Data are normally distributed |
| Homogeneity | 0.4106 | No serial autocorrelation |
| Spatial Autocorrelation | 0.0891 | Significant spatial autocorrelation |
| Mean Stationarity | 0.5904 | Data are non-stationary in mean |
| Variance Stationarity | 0.7164 | Data are non-stationary in variance |
| Spatial Outlier Detection | 0.0835 | Presence of spatial outliers in the data |

Based on Table 1, the data satisfy the assumptions of normality and the absence of serial autocorrelation, indicating temporal homogeneity. However, the results show significant spatial autocorrelation, and the data are non-stationary in both the mean and the variance, with spatial outliers. To address these issues, the Robust Kriging method is applied, as it can handle outliers and non-ideal spatial characteristics, resulting in more accurate interpolation. Subsequently, spatial outliers are identified using the Z algorithm, as presented in the following table.

### 3.3 Detection of Spatial Outliers

**Table 2.** Identification of Spatial Outliers

| Regency | Station | Latitude | Longitude | Data | Z-value | Outlier Decision |
|---|---|---|---|---|---|---|
| Bengkulu | Muara Bangkahulu | -3.7680 | 102.2900 | 136 | 1.944 | Yes |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| Seluma | Sukaraja | -3.9690 | 102.4183 | 365 | 0.381 | No |

Based on Table 2, it can be concluded that the Robust Kriging method can be applied to interpolate rainfall intensity data. The next step is to calculate the robust semivariogram, which requires grouping data pairs into six classes using Sturges' rule. Subsequently, the robust semivariogram values are computed for each class.

### 3.4 Empirical Semivariogram Analysis

Based on Table 3, robust semivariogram parameters were obtained, with a sill of $5.41 \times 10^{20}$, corresponding to the maximum semivariogram value. The range was determined to be 395.33, corresponding to the point at which the semivariogram reaches the sill. In addition, the nugget value was found to be $8.24 \times 10^{13}$, indicating small-scale variability or noise in the data. These parameters are essential for identifying the most suitable theoretical variogram model.

**Table 3.** Empirical Semivariogram Calculation

| Interval Class (km) | $N_h$ | Mean Distance (km) | Empirical Semivariogram (mm²) |
|---|---|---|---|
| [0,71.8] | 78 | 35.9 | $8.24\times10^{13}$ |
| (71.8,144] | 54 | 103 | $2.97\times10^{16}$ |
| (144,216] | 35 | 178 | $1.46\times10^{18}$ |
| (216,287] | 16 | 247 | $1.71\times10^{19}$ |
| (287,359] | 4 | 310 | $7.57\times10^{19}$ |
| (359,431] | 3 | 395 | $5.41\times10^{20}$ |

### 3.5 Theoretical Semivariogram Model

Several theoretical semivariogram models, namely the spherical, exponential, and Gaussian models, were then applied for comparison with the empirical (robust) semivariogram. Each model was fitted to the corresponding distance class, producing theoretical semivariogram values for each interval.

**Table 4.** Theoretical Semivariogram Calculation

| Interval Class | Spherical | Exponential | Gaussian |
|---|---|---|---|
| [0,71.8] | $7.35\times10^{19}$ | $4.70\times10^{19}$ | $4.44\times10^{18}$ |
| (71.8,144] | $2.06\times10^{20}$ | $1.24\times10^{20}$ | $3.54\times10^{19}$ |
| (144,216] | $3.41\times10^{20}$ | $1.96\times10^{20}$ | $9.92\times10^{19}$ |
| (216,287] | $4.41\times10^{20}$ | $2.52\times10^{20}$ | $1.75\times10^{20}$ |
| (287,359] | $5.06\times10^{20}$ | $2.94\times10^{20}$ | $2.49\times10^{20}$ |
| (359,431] | $5.41\times10^{20}$ | $3.42\times10^{20}$ | $3.42\times10^{20}$ |
| MSE | $8.81\times10^{40}$ | $3.29\times10^{40}$ | $1.75\times10^{40}$ |
| MAPE | 14998633% | 9580135% | 920601.3% |
| Nugget | $7.35\times10^{19}$ | $4.70\times10^{19}$ | $4.44\times10^{18}$ |
| Sill | $5.41\times10^{20}$ | $3.42\times10^{20}$ | $3.42\times10^{20}$ |
| Range | 395.33 | 395.33 | 395.33 |

1. Spherical Model

$$\gamma(h) = \begin{cases} 7.35 \times 10^{19} + 4.68 \times 10^{20}\left(\frac{3h}{2 \times 395.33} - \frac{h^3}{2 \times 395.33^3}\right), & \text{if } h \leq 395.33 \\ 5.41 \times 10^{20}, & \text{if } h > 395.33 \end{cases}$$

The Spherical model shows a linear increase in semivariance up to a certain distance (range), after which it reaches a constant value (sill). This model is suitable for data with strong spatial autocorrelation up to a certain distance, beyond which the correlation becomes random.

2. Exponential Model

$$\gamma(h) = 4.70 \times 10^{19} + 2.95 \times 10^{20}\left(1 - \exp\left(-\frac{h}{395.33}\right)\right)$$

The Exponential model increases rapidly at short distances and then gradually approaches the sill asymptotically without reaching a plateau. It is more appropriate for data with a gradual, continuous spatial effect that lacks a clear limit.

3. Gaussian Model

$$\gamma(h) = 4.44 \times 10^{18} + 3.38 \times 10^{20}\left(1 - \exp\left(-\left(\frac{h}{395.33}\right)^2\right)\right)$$

The Gaussian model exhibits a smoother increase in semivariance than the exponential model, with a slow initial rise that gradually approaches the sill. This model is suitable for data exhibiting smooth and extended spatial autocorrelation over longer distances.

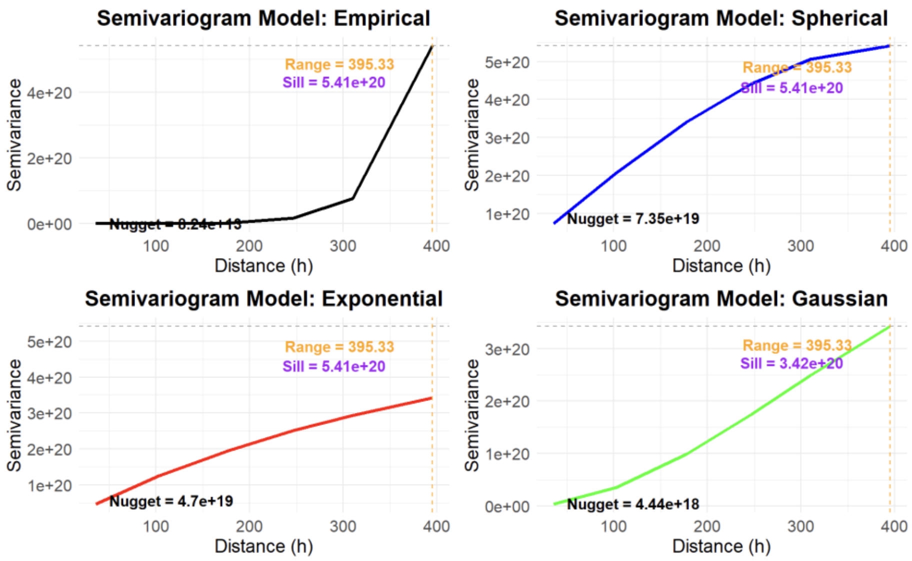


**Fig 2.** Comparison of Semivariograms

### 3.6 Robust Kriging

The Robust Kriging prediction aims to produce optimal estimates for unobserved locations by utilizing the spatial structure identified through the variogram model. During the prediction process, weights are determined for each prediction location to represent the contribution of each observation point to the estimation of values at new locations. These weights are calculated based on the spatial relationship between observation and prediction points, as defined by the Gaussian semivariogram model. The calculated weights are then used to estimate rainfall intensity at the specified locations, as shown in Table 5.

**Table 5.** Weight Values for Spatial Outliers

| Location | Observation Point | Weight | Location | Observation Point | Weight |
|---|---|---|---|---|---|
| Bengkulu (Muara Bangkahulu) | 1 | 0.084 | Lebong (Rimbo Pengadang) | 1 | 0.019 |
| | 2 | 0.086 | | 2 | 0.018 |
| | 3 | 0.084 | | 3 | -0.060 |
| | 4 | 0.071 | | 4 | -0.045 |
| | 5 | 0.078 | | 5 | 0.018 |
| | 6 | 0.041 | | 6 | 0.059 |
| | 7 | 0.051 | | 7 | 0.080 |
| | 8 | 0.102 | | 8 | 0.046 |
| | 9 | 0.046 | | 9 | -0.039 |
| | 10 | 0.028 | | 10 | -0.020 |
| | 11 | 0.014 | | 11 | 0.090 |
| | 12 | 0.015 | | 12 | 0.097 |
| | 13 | -0.001 | | 13 | 0.129 |
| | 14 | -0.001 | | 14 | 0.142 |
| | 15 | 0.101 | | 15 | 0.076 |
| | 16 | 0.084 | | 16 | 0.119 |
| | 17 | -0.002 | | 17 | 0.112 |
| | 18 | -0.050 | | 18 | 0.163 |
| | 19 | 0.078 | | 19 | 0.003 |
| | 20 | 0.093 | | 20 | -0.009 |

Rainfall estimation can be performed by calculating the weights of all sampled points using the product of the inverse covariance matrix ($\boldsymbol{C}^{-1}$) and the distance matrix ($\mathbf{D}$), as formulated in the corresponding equation. These weights are then applied to estimate rainfall values at the specified prediction locations. The same procedure can be extended to other observation points to obtain rainfall estimates across multiple locations. As a result, rainfall prediction values were generated for 20 observation stations in Bengkulu Province.

**Table 6.** Rainfall Prediction Results

| Location | Station | Latitude | Longitude | Prediction |
|---|---|---|---|---|
| Bengkulu | Sawah Lebar (Bk) | -3.797 | 102.284 | 314.5944 |
| Bengkulu | Stasiun Klimatologi Pulau Baai Bengkulu | -3.881 | 102.308 | 309.1993 |
| Bengkulu Selatan | Karang Cayo Pino (Masat) | -4.378 | 102.914 | 311.9895 |
| Bengkulu Selatan | Kedurang Ulu | -4.461 | 103.1 | 317.2942 |
| Bengkulu Tengah | Merigi Sakti | -3.6528 | 102.4022 | 328.5469 |
| Bengkulu Tengah | Pagar Jati | -3.621 | 102.36 | 327.0504 |
| Bengkulu Utara | Kemumu | -3.4321 | 102.2400 | 332.5912 |
| Bengkulu Utara | Tanjung Agung Palik | -3.5605 | 102.137 | 319.2539 |
| Kaur | Padang Guci Hulu | -4.4875 | 103.2622 | 321.5083 |
| Kaur | Tanjung Ganti | -4.5358 | 103.22 | 319.3246 |
| Kepahiang | Merigi | -3.51 | 102.55 | 346.7597 |
| Kepahiang | Muara Kemumu | -3.6678 | 102.73 | 344.4882 |
| Lebong | Gunung Alam | -3.152 | 102.186 | 347.9659 |
| Lebong | Lemeu | -3.076 | 102.226 | 355.695 |
| Mukomuko | Karang Jaya/Teras Terunjam | -2.52 | 101.24 | 334.4595 |
| Mukomuko | Lubuk Pinang | -2.4313 | 101.1772 | 336.6231 |
| Rejang Lebong | Pal 8 | -3.3787 | 102.47 | 350.4376 |
| Rejang Lebong | PT Agro Teh Bukit Daun | -3.438 | 102.465 | 343.4581 |
| Seluma | Sandabi B | -4.031 | 102.46 | 306.3453 |
| Seluma | Sandabi C | -4.028 | 102.51 | 310.151 |

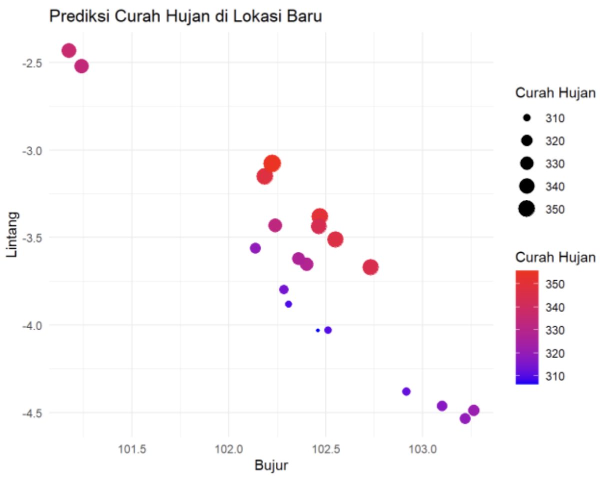


**Fig 3.** Spatial Distribution Map of Predicted Rainfall

The spatial pattern of predicted rainfall in Bengkulu Province is closely linked to the region's physical and atmospheric characteristics. Higher rainfall intensities observed in the northern and central areas can be attributed to orographic effects, as these zones lie near hilly and mountainous terrain where moist air masses from the Indian Ocean are forced to rise, cool, and condense. The province is also strongly influenced by the west-southwest monsoon, which transports moisture-laden winds toward the western coast, enhancing precipitation along elevated landscapes. Conversely, the southern and southeastern regions receive lower rainfall due to relatively flat topography and reduced exposure to monsoonal moisture. These combined topographic and atmospheric factors explain the spatial variability captured by the Robust Kriging predictions.

## 4 Conclusion

This study successfully applied the Robust Kriging method to predict rainfall in Bengkulu Province for January 2024. The process was chosen for its ability to handle outliers and non-ideal spatial characteristics in the data. Among the tested semivariogram models, the Gaussian model performed best, with the lowest MSE and MAPE. The prediction results revealed an apparent spatial variation in rainfall distribution, with higher intensities in the northern region and lower intensities in the southern region, providing valuable insights for water resource management, agricultural planning, and hydrometeorological disaster mitigation. The method proved effective in producing accurate estimations by incorporating data transformation, weighted median estimation, and spatial relationships among observation points.

However, this study has several limitations. The analysis relied on rainfall data from a limited temporal window (one month), which may not capture seasonal or inter-annual variability in Bengkulu's climate. In addition, the spatial prediction accuracy is influenced by the uneven distribution of the 97 BMKG stations, particularly in sparsely monitored inland regions. Future research could integrate multi-month or multi-year datasets, incorporate additional environmental covariates such as elevation or wind patterns, and explore advanced geostatistical approaches, such as co-Kriging, spatio-temporal Kriging, or machine-learning-based spatial models, to further enhance prediction reliability.

## References


[1] I. Muhandhis, A. S. Ritonga, and M. H. Murdani, "Implementasi Metode Inferensi Fuzzy Tsukamoto Untuk Memprediksi Curah Hujan Dasarian Di Sumenep," *J. Ilm. Edutic Pendidik. dan Inform.*, vol. 8, no. 1, pp. 01 10, 2021, doi: 10.21107/edutic.v8i1.8907.

[2] B. M. K. dan G. (BMKG, "Prakiraan Cuaca dan Informasi Iklim untuk Provinsi Bengkulu," BMKG. Accessed: Nov. 25, 2024. [Online]. Available: https://bengkulu.bmkg.go.id/

[3] N. Nur Rohma, "Pendugaan Metode Ordinary Kriging," *J. Penelit. Ilmu Sos. dan Eksakta*, vol. 2, no. 1, pp. 21 29, 2022, doi: 10.47134/trilogi.v2i1.33.

[4] L. Belkhiri, A. Tiri, and L. Mouni, "Spatial Distribution of The Groundwater Quality Using Kriging and Co-Kriging Interpolations," *Groundw. Sustain. Dev.*, vol. 11, no. February, p. 100473, 2020, doi: 10.1016/j.gsd.2020.100473.

[5] G. Bohling, *Introduction to Geostatistics and Variogram Analysis*. Kansas Geological Survey, 2005. doi: 10.1163/156852901753498098.

[6] E. Mahdi, A. H. Abuzaid, and A. M. A. Atta, "Empirical Variogram for Achieving The Best Valid Variogram," *Commun. Stat. Appl. Methods*, vol. 27, no. 5, pp. 547 568, 2020, doi: 10.29220/CSAM.2020.27.5.547.

[7] C. R. Oktarina, I. Sriliana, E. N. F. Sidik, and M. A. Firmansyah, "Pemodelan Data Geospasial Balita Kurang Gizi Dengan Pendekatan Geographically Weighted Regression Principal Component Analysis," *J. Gaussian*, vol. 13, no. 2, pp. 339 350, 2024, doi: 10.14710/j.gauss.13.2.339-350.

[8] Cinta Rizki Oktarina, Sri Syuhada Putri, Reza Pahlepi, Avrillia Permata Hati4, and Dyah Setyo Rini, "Pemetaan Spasial Keterkaitan Faktor Risiko Kematian Neonatal dengan Mixed Geographically Weighted Regression," *J. Ilmu Kesehat. dan Gizi*, vol. 2, no. 2, pp. 15 26, 2024, doi: 10.55606/jikg.v2i2.2818.

[9] C. R. Oktarina, J. Rizal, F. Faisal, Q. L. Tasyah, and S. C. Pratiwi, "Pemodelan IPM di Provinsi Bengkulu dengan Pendekatan Metode Geographically Weighted Regression (GWR) dan Geographically Temporally Weighted Regression (GTWR)," *J. EurekaMatika*, vol. 12, no. 1, pp. 23 34, 2024.

[10] B. Huang and J. Wang, "Big Spatial Data for Urban and Environmental Sustainability," *Geo-Spatial Inf. Sci.*, vol. 23, no. 2, pp. 125 140, 2020, doi: 10.1080/10095020.2020.1754138.

[11] M. Qu, J. Chen, B. Huang, and Y. Zhao, "Enhancing Apportionment of The Point and Diffuse Sources of Soil Heavy Metals Using Robust Geostatistics and Robust Spatial Receptor Model With Categorical Soil-Type Data.," *Environ. Pollut.*, vol. 265, p. 114964, 2020, doi: 10.1016/j.envpol.2020.114964.

[12] S. K. Al-Mamoori, L. A. Al-Maliki, A. H. Al-Sulttani, K. El-Tawil, and N. Al-Ansari, "Statistical analysis of the best GIS interpolation method for bearing capacity estimation in An-Najaf City, Iraq," *Environ. Earth Sci.*, vol. 80, no. 20, pp. 1 14, 2021, doi: 10.1007/s12665-021-09971-2.

[13] M. A. Oliver and R. Webster, *Basic Steps in Geostatistics: The Variogram and Kriging*, 1st ed. New York: Springer Cham, 2015.

[14] Z. Tatalovich, “A comparison of thiessen-polygon, kriging, and spline models of UV exposure,” *Univ. Consort. Geogr. Inf. Sci. Summer Assem.*, no. January 2005, 2005.

[15] N. A. C. Cressie, *Statistics for Spatial Data*, 1st ed. New York: John Wiley & Sons, 2013. doi: 10.1016/B0-12-348530-4/00169-7.

[16] H. Jo and M. J. Pyrcz, “Automatic Semivariogram Modeling by Convolutional Neural Network,” *Math. Geosci.*, vol. 54, no. 1, pp. 177 205, 2022, doi: 10.1007/s11004-021-09962-w.