JTAM (Jurnal Teori dan Aplikasi Matematika)
http://journal.ummat.ac.id/index.php/jtam
p-ISSN 2597-7512 | e-ISSN 2614-1175
Vol. 9, No. 3, July 2025, pp. 752-768

An Informative Prior of Bayesian Kriging Approach for Monthly Rainfall Interpolation in East Java

Rismania Hartanti Putri Yulianing Damayanti[1], Suci Astutik[1*], Ani Budi Astuti[1]
[1]Department Statistics, Universitas Brawijaya, Indonesia
suci_sp@ub.ac.id

Article History:
Received : 30-04-2025
Revised : 28-05-2025
Accepted : 02-06-2025
Online : 01-07-2025

Keywords:
Bayesian;
Kriging;
Rainfall;
Spatial.

ABSTRACT

In spatial data analysis, interpolation is used to estimate values at unobserved locations, but often faces challenges in capturing complex spatial patterns and estimation uncertainty. One of the main obstacles is the small sample size, which makes the empirical variogram difficult to define well in conventional Kriging methods. The Bayesian Kriging approach overcomes this problem by integrating prior information, so it can still produce stable estimates despite limited data. This study is a quantitative, spatial-based research aimed at interpolating monthly rainfall in East Java Province using the Bayesian Kriging approach. The data consist of monthly rainfall measurements from 11 rain gauge stations distributed across East Java, obtained from the Indonesian Agency for Meteorology, Climatology, and Geophysics (BMKG) for the period of January to April 2024. The entire analysis was conducted using R software. A spherical semivariogram model was selected due to its superior fit to the spatial characteristics of the rainfall data in the study area with the smallest RMSE 37.17. This study demonstrates the effectiveness of Bayesian Kriging for rainfall interpolation in tropical regions with sparse data, providing more stable and accurate estimates compared to conventional methods. The scientific contribution of this research lies in showcasing how the integration of informative priors and Bayesian inference enhances interpolation accuracy in data-limited tropical environments. The resulting interpolated maps can inform land-use planning and flood risk mitigation by identifying areas of high rainfall for improved water infrastructure and lower-rainfall regions for targeted irrigation planning.

Crossref
https://doi.org/10.31764/jtam.v9i3.31027

———————— ◆ ————————

A. INTRODUCTION

Geostatistics refers to a set of quantitative techniques designed to analyze and interpret data that possess spatial or locational attributes (Gao et al., 2024). Emerging in the 1980s, this field developed through the integration of disciplines such as mining, geology, mathematics, and statistics. Its primary strength lies in the ability to model spatial patterns, trends, and correlations within data (Biswas & Biswas, 2024). In parallel, spatial analysis plays a vital role in various field such as environmental planning, risk assessment, and resource management by enabling the evaluation of geographically distributed variables (Breunig et al., 2020; Rai et al., 2022). It enables the detection of spatial patterns and trends that may go unnoticed when relying exclusively on sparse point-based environmental data (Lu et al., 2024).

A challenge in spatial analysis is the limited availability of data, which is often caused by difficulties in collecting comprehensive and representative data across the study area. Factors such as difficult geographical access, limited resources, and high operational costs are barriers to obtaining data in all desired locations. As a result, the spatial data available is usually limited and only covers certain observed points (Zorzetto & Marani, 2020). To overcome this limitation, spatial interpolation techniques are used as an approach to estimate the value of variables in locations that do not have direct observations, by utilizing information from points that have been observed in the vicinity (Uddin & Czajkowski, 2022). Interpolation allows the construction of continuous maps that depict the spatial distribution of an environmental variable more thoroughly. For example in the context of hydrometeorology, rainfall interpolation is used to map spatial distribution patterns of rainfall from data available only at a limited number of observation stations. This technique is also very useful in other fields such as monitoring air quality, soil moisture, surface temperature, and other environmental parameters, especially in regional planning efforts, natural resource management, and disaster mitigation (Nicoletta et al., 2021; Sharma et al., 2021). Interpolation helps policy makers and researchers to understand the condition of the region more accurately and in-depth, even though the available data is spatially limited.

A frequently used interpolation method for estimating values in spatial data is Ordinary Kriging (OK) which is the best unbiased linear estimator and produces reliable predictions at unsampled locations (Erten et al., 2022). OK relies on some basic assumptions that must be met in order for its prediction results to be reliable, namely the existence of spatial dependence and stationarity of spatial data (Gribov & Krivoruchko, 2020). Spatial dependence indicates that values at a point are influenced by values at other points based on distance, which can be modeled with a variogram that describes how much influence the distance between points has on measured values (Mahdi et al., 2020). Several studies have applied this method to analyze rainfall patterns in various locations, such as Peninsular Malaysia (Jamaludin & Suhaimi, 2013), southern Brazil (Charles et al., 2022), and Indonesia (Maulana et al., 2022). OK has been shown to effectively map the spatial distribution of rainfall, showing areas of high intensity and dry conditions (Jamaludin & Suhaimi, 2013; Chutsagulprom et al., 2022). In practice, environmental data such as rainfall, air quality, or soil moisture often exhibit non-linear, heterogeneous, and complex spatial patterns that are difficult to capture using simple linear models like OK. When the underlying data does not meet the assumptions of stationarity or exhibits non-linear relationships, the performance of OK can deteriorate, leading to less accurate spatial predictions. Therefore, OK relies on semivariogram estimation to model the spatial structure of the data. When data is limited, the estimated semivariogram becomes unstable and unrepresentative because a sufficient number of point pairs for each lag distance is not available (Han & Suh, 2024). This puts the OK model at risk of overfitting to local noise or underfitting and thus failing to capture true spatial variation. As a result, the predicted values at unmeasured locations are less accurate and have high variance because the kriging weights obtained are not optimal. One such method is Bayesian Kriging (BK), which integrates prior into the modeling process and allows more robust spatial prediction even when classical assumptions are violated.

Bayesian methods offer a more comprehensive approach where uncertainty is explicitly modeled through probability distributions (Astutik et al., 2023; Schoot et al., 2021). The Bayesian approach provides advantages in terms of producing more realistic and robust estimates, especially in conditions of limited data (Schoot & Miočević, 2020). BK is a more flexible approach in kriging because it considers the covariance structure as unknown (Gelfand & Banerjee, 2017). Unlike traditional methods that estimate parameters using ordinary least squares or maximum likelihood, BK employs prior distributions for both parameters and hyperparameters which are iteratively updated using data through MCMC simulations. This process produces posterior distributions for each parameter, enabling a robust quantification of uncertainty an aspect often limited in conventional approaches. BK has been widely adopted across various disciplines, not only to model spatial structures but also to provide deeper insights into uncertainty within geophysical processes, ecological patterns, environmental pollution, topography, and subsurface characteristics (Lima et al., 2021).

This study uses the BK approach with informative priors for interpolating monthly rainfall in East Java with 11 rain gauges. In Bayesian analysis, informative priors incorporate existing knowledge or findings from previous studies to guide the estimation process, especially when observational data are limited. Informative prior can be obtained through the posterior results of previous studies, thereby ensuring that new estimates are grounded in prior empirical evidence (Wesner & Pomeranz, 2021; Zondervan-Zwijnenburg et al., 2017). In this case, prior information is adapted from previous study (Verdin et al., 2015), who applied BK to model precipitation in Central and South America. The prior distribution used in parameter estimation is a non-informative prior with the distribution used is flat prior for coefficient regression, inverse gamma prior for sill parameters and uniform prior for range parameters, and the nugget parameter is not inccluded. The resulting posterior distributions are: normal for regression coefficients, inverse gamma for the sill, and uniform for the range. However, the analysis reveals that the posterior distribution is imperfect due to difficulties in jointly identifying the sill and range parameters, which is a known issue in geostatistical modeling when using limited data or non-informative priors (Verdin et al., 2015). This challenge can lead to weak identifiability and high uncertainty in spatial prediction.

Based on the background described, this study aims to spatially interpolate monthly rainfall in East Java using the BK approach. By incorporating prior information into the modeling process, this study seeks to overcome challenges related to limited observational data and improve the accuracy of rainfall predictions in unsampled locations. This research also aims to demonstrate the applicability and effectiveness of Bayesian Kriging on monthly rainfall interpolation in East Java with a sparse monitoring network. The analysis results are expected to produce monthly rainfall distribution maps that can support more informed decision-making in areas such as water resources management, agricultural planning, and disaster mitigation.

B. METHODS

1. Data Source

The data used in this study are secondary monthly rainfall data obtained from the East Java Meteorology, Climatology and Geophysics Agency (BMKG) website in November 2023-April 2024 which is the rainy season period. This study uses 11 observation locations as spatial units shown in Figure 1.

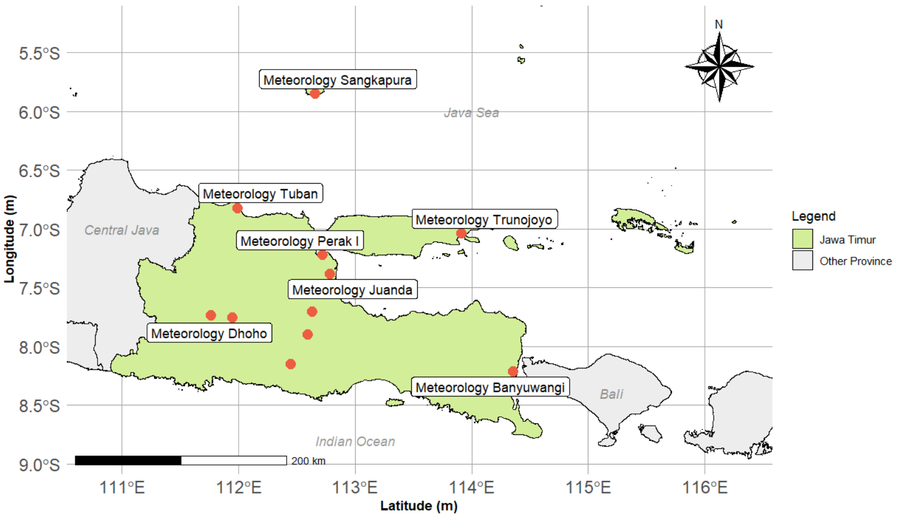

Figure 1. Study Area

The study area in this research is East Java, which is one of the provinces in Indonesia. Astronomically, East Java is located between 111°0' to 114°4' East Longitude and 7°12' to 8°48' South Latitude. East Java Province borders the Java Sea to the north, the Indian Ocean to the south, the Bali Strait to the east, and Central Java Province to the west. This astronomical location gives East Java a tropical climate with weather variations influenced by latitude and altitude from sea level. East Java has 11 weather and climate observation stations. Weather and climate observation stations have a strategic role in providing data that forms the basis for scientific analysis and evidence-based policy making. The resulting long-term data is essential for monitoring climate dynamics, including rainfall patterns, temperature and humidity, to understand trends in environmental change and their impact on ecosystems.

2. Spatial Effect Test

The main concepts in spatial analysis are spatial heterogeneity and spatial dependence. Spatial dependence relates to the correlation between observations based on their spatial proximity (Gao et al., 2024). Spatial dependence or spatial autocorrelation occurs due to the similarity of characteristics that occur in adjacent locations. The test statistic often used to test for spatial autocorrelation is the Moran-I statistic (Efendi et al., 2023). The hypothesis used to test for autocorrelation is as follows.

H_0 : $\forall Cor(Y_i, Y_j) = 0; i \neq j$ (has no spatial autocorrelation) vs

H_1 : $\exists Cor(Y_i, Y_j) \neq 0; i \neq j$ (has spatial autocorrelation)

$$I = \frac{\sum_{j=1}^{n} \sum_{i=1}^{n} w_{ij} (Y_i - \bar{Y})(Y_j - \bar{Y})}{\sum_{i=1}^{n} (Y_i - \bar{Y})^2} \quad (1)$$

I is the Moran's I statistic, Y_i represents the observed value at location-i, $\bar{Y}$ is the mean of all observed values, Y_j is the observed value at location-j, and w_{ij} denotes the spatial weight between locations i and j, typically based on the distance or adjacency between these locations. This test criterion is seen from the value of I which lies between -1 and 1. If the value of I is negative, it states negative autocorrelation, and if I is positive, it can be said that there is positive autocorrelation.

3. Ordinary Kriging

Spatial interpolation is a method to predict values at locations that do not have sample points. One method for spatial interpolation is Kriging. The Kriging method was developed by D.L. Krige to estimate the value of mining ore distribution based on data from several sample locations which was further developed by G. Matheron in 1963 (Cressie, 1993). Kriging is a geostatistical method that uses known values and semivariograms to predict values at other unmeasured locations. The predicted values in the kriging method vary depending on the proximity to the original data location. The Kriging method uses a linear combination of weights to estimate values between data samples.

There are several types of kriging methods such as simple, ordinary, universal, indicator, disjunctive, and probability kriging (Bostan, 2017). Among these methods, OK is the most commonly used spatial interpolation method. OK assumes that the average measured value is constant across the area (spatial stationarity). This method is particularly suitable for environmental data that exhibit local spatial dependence without a strong global trend. It is also preferred when the data show random spatial variability and are limited in number. OK is ideal when there is no clear spatial trend in the data or when observations are sparse for example, rainfall data that are randomly scattered without a strong geographic gradient because it provides a balance between model simplicity and prediction accuracy under minimal assumptions (Han & Suh, 2024).

In using the kriging interpolation method, there are two main steps that must be performed, namely recognizing spatial dependence patterns and making predictions in unmeasured areas (Varga et al., 2023). The first step involves creating variograms and covariance functions to identify and estimate the degree of spatial autocorrelation between the data points. Variograms can help in understanding the extent to which a point at a particular location is affected by points in its vicinity. The information in the variogram can be used to predict values at unmeasured locations through the Kriging equation. The general form of the Kriging equation for estimating data at unsampled locations is presented in equation (2).

$$Z(s) = \mu(s) + \varepsilon(s) \tag{2}$$

where μ is the average of the process and $\varepsilon(s)$ is a random quantity with zero mean and has covariance $C(h)$ where h is the separation in space also known as lag. The covariance ($C(h)$) is shown in equation (3).

$$C(\boldsymbol{h}) = E[\varepsilon(\boldsymbol{s})\varepsilon(\boldsymbol{s} + \boldsymbol{h})] \tag{3}$$

Theoretically, there are semivariogram models such as the exponential model, spherical model, and Gaussian model written as follows.

a. Spherical

$$\gamma(\boldsymbol{h}) = \begin{cases} \sigma^2 \left[\dfrac{3\boldsymbol{h}}{2\phi} - \dfrac{1}{2}\left(\dfrac{\boldsymbol{h}}{\phi}\right)^2\right], & for\ 0 < h < r \\ \sigma^2, & for\ the\ others \end{cases} \tag{4}$$
$$\gamma(0) = 0$$

b. Exponential

$$\gamma(\boldsymbol{h}) = \sigma^2 \left[1 - \exp\left(-\frac{\boldsymbol{h}}{\phi}\right)\right] + \tau^2 \tag{5}$$
$$\gamma(0) = 0$$

with $\gamma(h)$ is defined as semivariogram in h^{th} lag, σ^2 as *sill* parameter, ϕ as *range* parameter, τ^2 as nugget variance parameter. Suppose there is a value of a random variable Z with locations $s_1, s_2, \ldots s_N$ at N observations, $z(s_i)$ with $i = 1,2,\ldots,N$. To estimate the value at an unsampled location or point s_o can use the estimator in equation (6).

$$\hat{Z}(s_0) = \sum_{i=1}^{n} \lambda_i Z(s_i) \tag{6}$$

where $\hat{Z}(s_0)$ is the estimated value and λ_i is the weight of the kriging estimator that satisfies $\sum_{i=1}^{N} \lambda_i = 1$. An illustration of λ_i as the weight of the kriging estimator is shown in Figure 2.

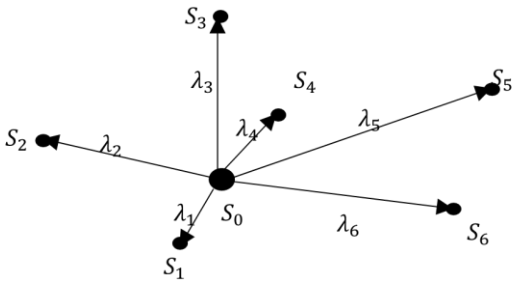

Figure 2. Illustration of Kriging Weights at Location s_0

The unknown value at the target location is estimated as a weighted linear combination of nearby observed data points. Weights ($\lambda_1, \lambda_2, \ldots, \lambda_n$) are determined based on spatial correlation, typically modeled using a semivariogram.

4. Bayesian Kriging

Bayesian Kriging was introduced by Omre (1987) which is a method of combining kriging with parameter estimation using the Bayesian approach (Omre, 1987). The Bayesian approach to Kriging involves prior information about the variogram with its parameters, namely mean (β), sill (σ^2), range (ϕ), nugget (τ^2). The model parameters in Bayesian Kriging are estimated from the posterior distribution shown in equation (7).

$$p(\boldsymbol{\theta}|Z) \propto f(Z|\boldsymbol{\theta})\pi(\boldsymbol{\theta}) \tag{7}$$

Where $f(Z|\theta)$ is the likelihood function and $\pi(\theta)$ is the prior distribution, while Z is the observed data, θ is the vector of model parameters (variogram), $P(\theta|Z)$ is the posterior distribution. The explanation of the likelihood function and posterior distribution of Bayesian Kriging begins with an understanding of the spatial model. Suppose there is a spatial model written hierarchically as follows.

Level 1 : $\boldsymbol{Y}(\boldsymbol{u}) = \boldsymbol{X}(\boldsymbol{u})\boldsymbol{\beta} + S(\boldsymbol{u}) + \varepsilon(\boldsymbol{u})$
$\boldsymbol{Y}(\boldsymbol{u}) = \boldsymbol{X}(\boldsymbol{u})\boldsymbol{\beta} + \sum_{k=1}^{K} \sigma_k T_k(\boldsymbol{u}) + \varepsilon(\boldsymbol{u})$;

Level 2 : $\boldsymbol{T}_k(\boldsymbol{u}) \sim \mathcal{N}\left(0,\ R_k(\phi_k)\right),\ \boldsymbol{T}_1, \dots, \boldsymbol{T}_k$ are independent and $\varepsilon(\boldsymbol{u}) \sim iid\ \mathcal{N}(0, \tau^2 I)$;

Level 3 : $(\boldsymbol{\beta}, \boldsymbol{\sigma^2}, \phi, \tau^2) \sim pr(\,.\,)$, as prior distribution

The likelihood function and the prior distribution of the model parameters can be written in equation (8) and equation (9).

$$L(\beta, \sigma^2, \phi, \tau^2|Y) \propto (2\pi)^{-\frac{n}{2}}. |R_k(\phi_k)|^{-\frac{1}{2}}. \exp\left(-\frac{1}{2}[\boldsymbol{Y} - \boldsymbol{X}\beta]\,' R_k(\phi_k)^{-1}[\boldsymbol{Y} - \boldsymbol{X}\beta]\right) \tag{8}$$

$$p(\beta, \sigma^2, \phi, \tau^2) = p(\phi, \tau^2)p(\beta, \sigma^2|\phi, \tau^2) \tag{9}$$

For the prior distribution each parameter is written in Table 1.

Table 1. Prior Distribution Each Parameter

Paremeter	**Prior Distribution**
β	$\beta \sim \mathcal{N}(\mu_0, \boldsymbol{V}_0)$
σ^2	$\sigma^2 \sim Scl - Inv - \chi^2(v_0, s_0^2)$
ϕ	$\phi \sim U(\phi_{min}, \phi_{maks})$
τ^2	$\tau^2 \sim U(\tau^2_{min}, \tau^2_{maks})$

The posterior of all model parameters can be written as follows.

$$p(\beta, \sigma^2, \phi, \tau^2|Y) = p(\beta, \sigma^2|Y, \phi, \tau^2)p(\phi, \tau^2|Y) \tag{10}$$

$$p(\phi, \tau^2|Y) \propto \frac{p(\beta, \sigma^2, \phi, \tau^2) \times p(y|\beta, \sigma^2, \phi, \tau^2)}{p(\beta, |Y, \sigma^2, \phi, \tau^2)\ p(\sigma^2|Y, \phi, \tau^2)} \tag{11}$$

$$p(\phi, \tau^2|Y) \propto \frac{p(\phi, \tau^2)\, p(\beta, \sigma^2|\phi, \tau^2) \times\, p(y|\beta, \sigma^2, \phi, \tau^2)}{p(\beta, |Y, \sigma^2, \phi, \tau^2)\, p(\sigma^2|Y, \phi, \tau^2)} \tag{12}$$

The determination of posterior distribution for β and σ^2 can be obtained analytically because they belong to the conjugate prior type. However, the determination of posterior distribution for parameters ϕ and τ^2 cannot be done analytically but by MCMC. One MCMC algorithm was introduced by (Tanner, 1996) to obtain the posterior distribution of each parameter in Bayesian Kriging. This algorithm is used in package geoR in RStudio software with the following algorithm steps.

a. Discretize the distribution of $(\phi, \tau^2\ |Y)$, i.e. select a set of values for ϕ and τ^2 within a reasonable interval. In this condition, a uniform prior is used for ϕ and τ^2 on the selected set of values.
b. Calculate the posterior probability with equation (12). The result of this calculation forms the discrete posterior distribution $p(\phi, \tau^2\ |Y)$.
c. Sampling the values of ϕ and τ^2 from the distribution $p(\phi, \tau^2\ |Y)$.
d. Plugging the retrieved values of ϕ and τ^2 into the distribution $p(\beta, \sigma^2 \,\middle|\, Y, \phi, \tau^2\ |Y\ )$ and sampling from this distribution.
e. Repeating step 3 and step 4 for the desired number of iterations until obtaining the desired $(\hat{\beta}, \hat{\sigma}^2, \hat{\phi}, \hat{\tau}^2\ )$ from the posterior distribution.

5. Performance Evaluation

Root Mean Square Error (RMSE) is a statistical metric used to measure how well a model's predictions approximate the actual observed values. RMSE measures the root of the mean square of the difference between the predicted value and the true value. The formula of RMSE is shown in equation (13) (Hodson, 2022).

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(z_0 - \hat{z}_0)^2} \tag{13}$$

where n is the number of observations, z_0 is the actual value, and $\hat{z}_0$ is the predicted value. If the RMSE is smaller, it indicates that the model predictions are more accurate, so the model performance can be said to be good. The advantage of RMSE over MSE is that the criterion is in the same units as the target variable, making it easier to interpretation (Althoff & Rodrigues, 2021).

C. RESULT AND DISCUSSION

1. Descriptive Statistics

Rainfall is the amount of rain over a specific area and time measured in millimeters (mm) and is vital for ecosystems, agriculture, and human needs. In Indonesia, rainfall varies spatially and temporally, influenced by geography, wind, and climate phenomena like El Niño and La Niña. In East Java, during the study period, rainfall ranged from 139.6 mm (Banyuwangi) to

450.6 mm (Nganjuk). Monthly rainfall is classified as medium (100–300 mm) and high (300–500 mm), with high rainfall observed at Nganjuk and Dhoho stations.

2. Spatial Autocorrelation

Before conducting Bayesian Kriging analysis, it is essential to perform a spatial autocorrelation test to ensure the presence of spatial dependence in the data. One of the most commonly used methods is the Moran's I statistic, which measures the degree of correlation between the values of a variable at one location and those at neighboring locations, based on a predefined spatial weights matrix. The results of the spatial autocorrelation test in this study are presented in Table 2.

Table 2. Moran's I Test

E(I)	Var(I)	Moran Statistics	P-value
-0.1	0.11	0.463	0.046

The p-value of 0.046 indicates that the Moran's I statistic is significant at the 5% level, suggesting a significant spatial autocorrelation in the monthly rainfall data across East Java. This implies that rainfall values at nearby locations are spatially related.

3. Empirical an Theoritical Variogram

The empirical variogram in Figure 3 derived from the monthly rainfall data reveals a characteristic spatial structure. At short distances, the semivariance values are relatively low, indicating a strong similarity in rainfall amounts between nearby locations. As the distance increases, the semivariance rises, reaching a peak around 100–120 km, reflecting a decline in spatial correlation. Beyond this range, the semivariance drops significantly, likely due to the limited number of data pairs at larger distances, leading to less stable estimates. In the context of Bayesian Kriging, this variogram serves as a basis for constructing the spatial prior. The relatively small nugget effect suggests minimal measurement error or small-scale variability in the data. The sill, estimated between 12,000 and 13,000, represents the total data variance, while the range, approximately 100–120 km, defines the maximum distance over which spatial correlation persists.

In order to accurately model the spatial structure of the monthly rainfall data, several theoretical variogram models were considered. Two models spherical and exponential were selected due to their widespread use and suitability for various types of spatial data. Each model was fitted to the empirical variogram derived from the observed data, using consistent initial parameters for nugget, sill, and range. The fitting results were then compared both visually and statistically to determine the model that best represents the underlying spatial correlation structure. This selection is crucial for ensuring the robustness of subsequent Bayesian Kriging analysis.

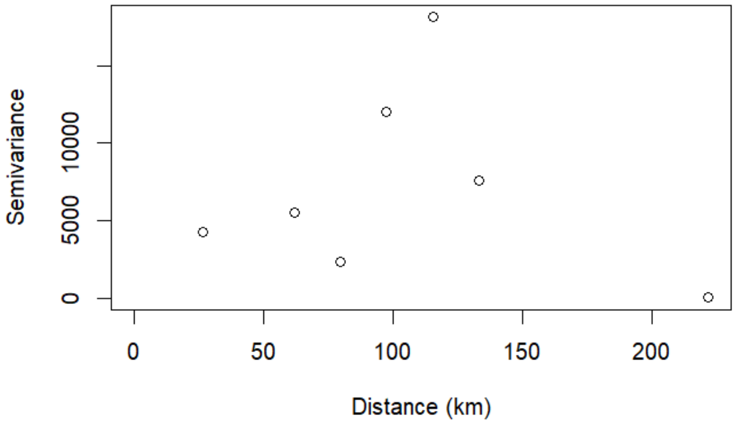

Figure 3. Empirical Variogram

4. Model Fitting

In this section, the results of Bayesian Kriging using four different theoretical variogram model as spherical and exponential are compared. Each model was fitted with identical prior information to isolate the influence of variogram structure on the prediction outcomes. The analysis includes a comparison of the trace plot, autocorrelation plot, and Monte Carlo Error, aiming to identify the variogram model that provides the most accurate and reliable spatial interpolation of monthly rainfall data. Trace plot of of each posterior parameter for spherical model are shown in Figure 4 and the trace plot for exponential model are shown in Figure 5 below.

Trace plots are used to check the convergence of MCMC chains in the context of Bayesian Kriging. Trace plots that show good convergence are characterized by random fluctuations around the mean value without any consistent upward or downward trend, indicating that the sample has reached a stationary state. In this analysis, the trace plots for the spherical and exponential models show that all parameters (β, σ^2, ϕ, and τ^2) have reached convergence, which is reflected by the stable and dense fluctuation patterns throughout the 10,000 iterations. Overall, there is no significant difference between the trace plots for the spherical and exponential models. Both models show almost identical patterns, with the most striking difference being the parameter σ^2. In the exponential model, this parameter shows slightly more extreme spikes compared to the spherical model, although the difference is relatively small.

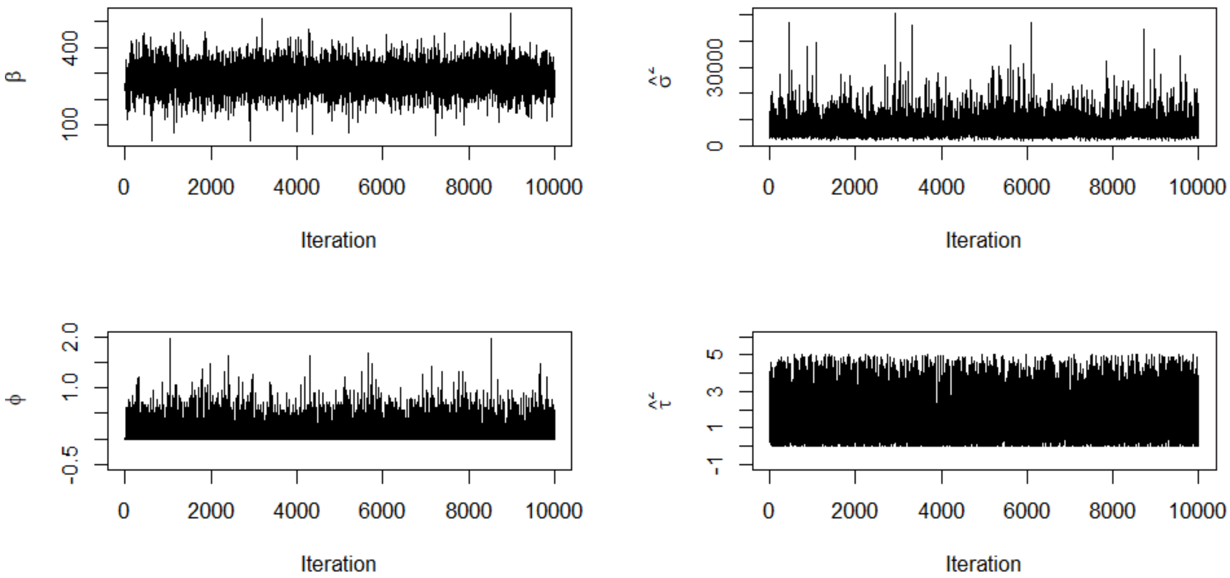

Figure 4. Trace Plot of Each Parameter for Spherical Variogram Model

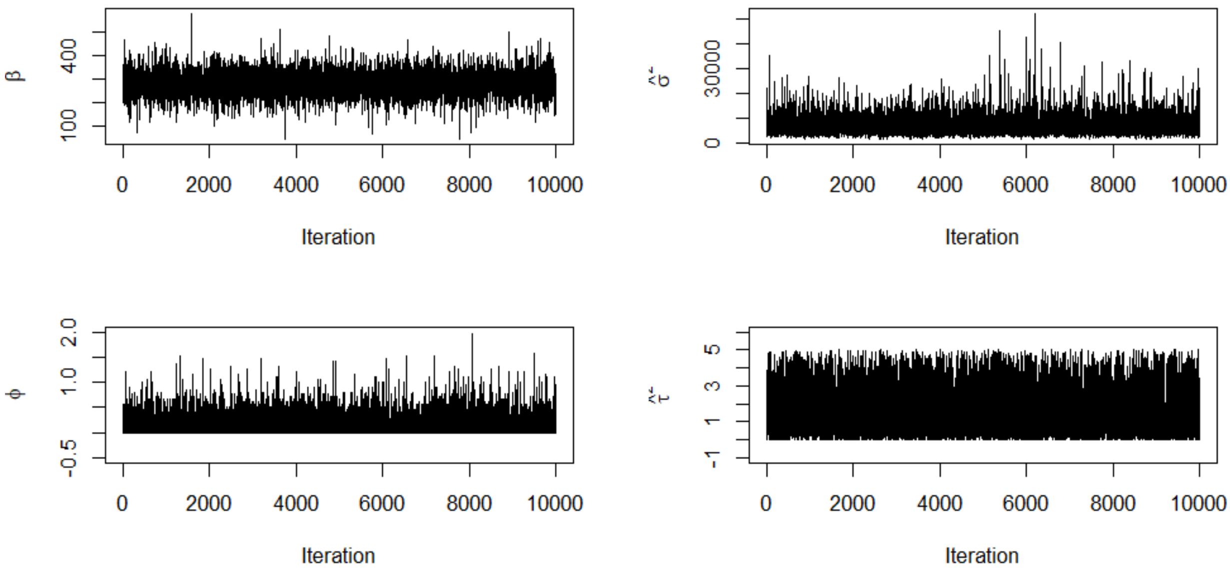

Figure 5. Trace Plot of Each Parameter for Exponential Variogram Model

Upon observing the ACF plots for both the spherical and exponential variogram models, it can be seen that the autocorrelation values for all parameters (β, σ^2, ϕ, and τ^2) largely remain within the blue dashed confidence intervals across the first 20 lags. This suggests that there is no strong correlation between consecutive samples after a few steps, which is a strong indicator of efficient mixing in the MCMC sampling. Efficient mixing is crucial because it means the generated samples are representative of the true posterior distribution, rather than being heavily dependent on previous samples.

Furthermore, the minimal autocorrelation observed complements the previous results from the trace plots where the chains fluctuated randomly around a stable mean without obvious trends and the density plots where smooth and unimodal distributions were obtained. These three types of diagnostics (trace plot and ACF plot) collectively suggest that the MCMC chains have likely converged and that the samples can be considered reliable for posterior inference. Autocorrelation plot of of each posterior parameter for spherical model are shown in Figure 6 and the autocorrelation plot for exponential model are shown in Figure 7 below.

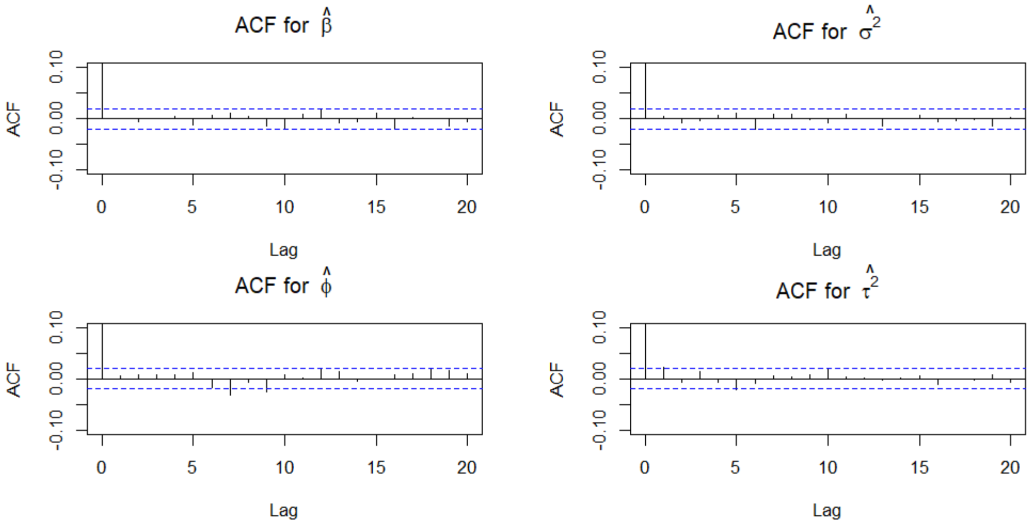

Figure 6. Autocorrelation Plot of Each Parameter for Spherical Variogram Model

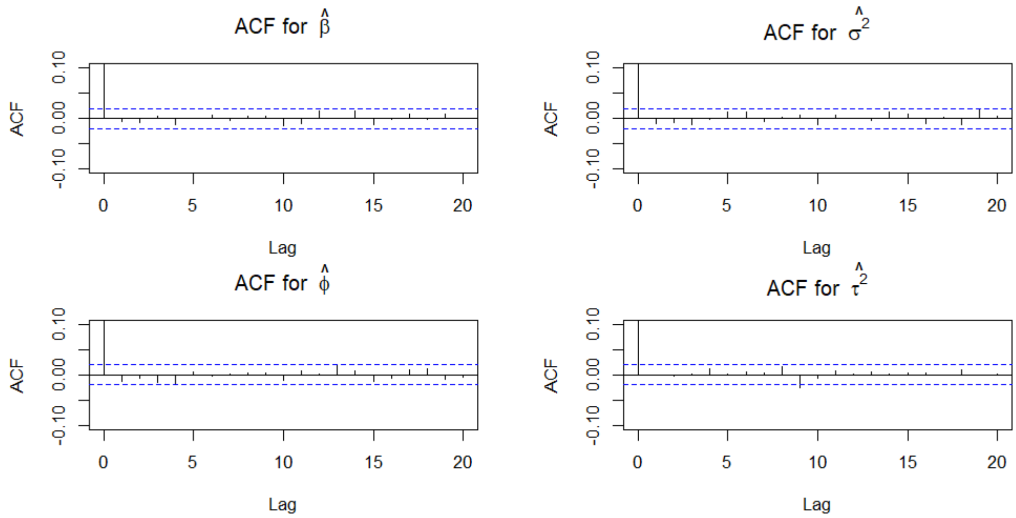

Figure 7. Autocorrelation Plot of Each Parameter for Exponential Variogram Model

When comparing the ACF patterns between the spherical and exponential models, the plots are very similar in terms of autocorrelation structure: both models show quick decay of autocorrelation and maintain values close to zero after the first few lags. This indicates that changing the variogram models has minimal impact on the sampling performance and convergence behavior of the Bayesian Kriging approach used. Thus, in terms of MCMC diagnostics, both models perform equally well, and the choice between them may be guided more by the theoretical or practical suitability for spatial data rather than concerns about convergence. The following are the parameter estimation results of two variogram models, namely the Spherical and Exponential models shown in Table 3.

Table 3. Summary of Posterior Distribution Each Model

Variogram Model	Parameter Estimator	*Mean*	*Standard Deviation*	*Credible Interval* 2,5%	*Credible Interval* 97,5%	MC *Error*	5% SD
Spherical	$\hat{\beta}$	246.66	39.53	166.82	323.66	0.39	1.98
	$\hat{\sigma}^2$	7345.19	2933.93	3449.37	14509.57	29.34	146.69
	$\hat{\phi}$	0.56	0.49	0	1.85	0.005	0.02
	$\hat{\tau}^2$	1.09	1.04	0	3.99	0.015	0.05
Exponential	$\hat{\beta}$	282.03	45.62	189.04	373.28	0.46	2.28
	$\hat{\sigma}^2$	7003.62	3616.43	2346.33	16157.38	36.16	180.82
	$\hat{\phi}$	0.18	0.19	0	0.65	0.002	0.01
	$\hat{\tau}^2$	1.51	1.25	0	4.54	0.001	0.06

The estimation results of Bayesian Kriging based on the trace plots, ACF plots, and summary statistics indicate that both models using spherical and exponential semivariograms have reached convergence, as evidenced by the stable and random fluctuations of parameters across iterations. However, notable differences emerge in terms of estimation stability and precision. The exponential model yields a more concentrated estimate for the spatial range parameter $\hat{\phi}$ compared to the spherical model, indicating a shorter spatial correlation and greater precision, supported by a smaller Monte Carlo error. Meanwhile, the spherical model exhibits higher variability in the estimation of $\hat{\sigma}^2$ (SD = 2933.93) than the exponential model (SD = 3616.43), though it features a slightly narrower credible interval. The parameter $\hat{\beta}$ is estimated with a higher mean under the exponential model (282.03), albeit with greater uncertainty (SD = 45.62 vs. 39.53). These differences highlight the importance of selecting an appropriate variogram model; the exponential model appears more stable in estimating spatial parameters, while the spherical model is more sensitive to local variation.

5. Validation Model

The performance of Bayesian Kriging in interpolating monthly rainfall in East Java can be seen through model validation on testing data. Validation on the testing data is seen based on the RMSE value on the interpolation results shown in Table 4.

Table 4. RMSE Value

Variogram Model	RMSE
Spherical	37.17476
Exponential	40.22298

The validation results of the Bayesian Kriging model on the test data show that the model with the spherical variogram produces a RMSE lower than the exponential variogram. The lower RMSE value of the spherical variogram indicates that this model has better accuracy in predicting monthly rainfall in East Java. This indicates that the spatial structure of rainfall data in the region is more in line with the assumptions built by the spherical variogram, where the level of interrelationship between locations tends to decrease gradually until it reaches a certain threshold (range). This difference in performance between variograms also emphasizes the importance of selecting the right variogram model in the spatial interpolation process using the Bayesian Kriging approach. The selection of an appropriate variogram can improve the accuracy of predictions in unobserved areas, so that interpolation results are more reliable for decision making, especially in the context of planning and mitigating hydrometeorological disasters such as floods or droughts. The spherical variogram can be recommended as a more optimal model for rainfall characteristics in East Java in this study. The interpolation results of East Java monthly rainfall based on Bayesian Kriging can be seen through the contour map in Figure 8.

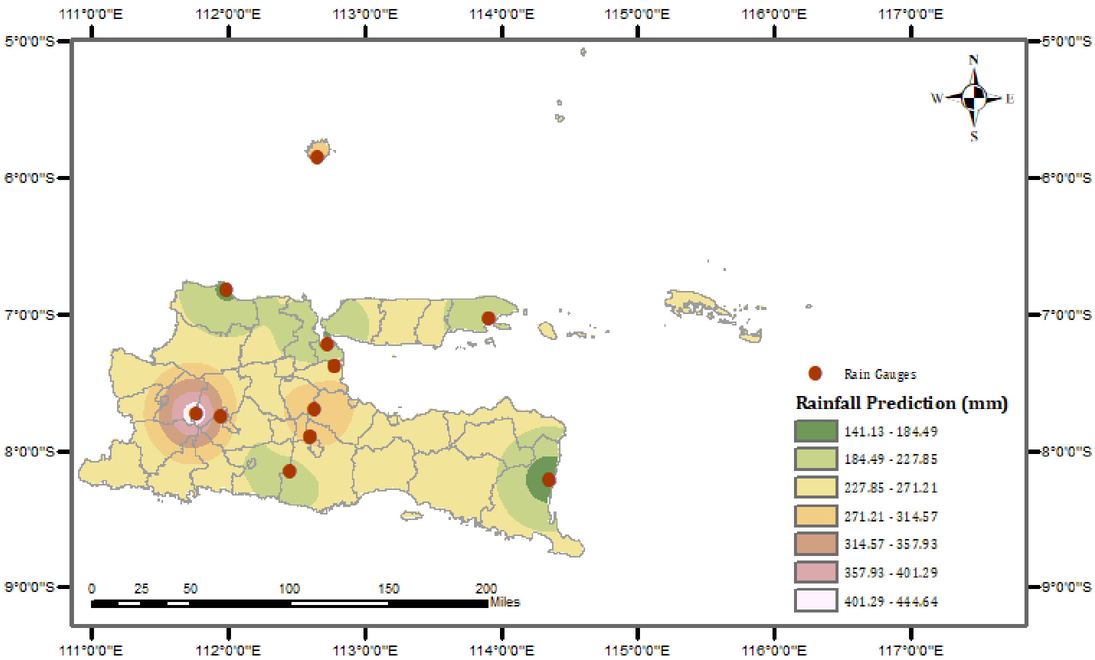

Figure 8. Monthly Rainfall Interpolation Map in East Java Based on Bayesian Kriging Model

The interpolated map (Figure 8) of monthly rainfall in East Java illustrates the spatial distribution of rainfall predicted using the Bayesian Kriging method. The interpolated results show that the highest rainfall is concentrated in the southwestern part of East Java, marked in purple and pink. This indicates the possible influence of topography such as mountains or upstream watershed areas that tend to receive more precipitation. In contrast, most other areas show moderate to low rainfall, especially in the north and east. The spatial distribution shown is quite smooth and consistent with the geographical pattern of East Java, which demonstrates the effectiveness of the Bayesian Kriging method in capturing spatial structure and accommodating prediction uncertainty. These results can serve as the basis for policy planning for water management, agriculture, and disaster mitigation at the local level.

When compared with previous studies that used OK with Gaussian semivariogram to map rainfall in Sulawesi, the general spatial patterns particularly the high rainfall in mountainous regions appear to be consistent (Sanusi et al., 2024). However, the Bayesian Kriging approach employed in this study offers the added advantage of incorporating prior knowledge and explicitly modelling uncertainty, resulting in smoother and more reliable spatial predictions (Kanooni & Amogein, 2025; Lima et al., 2021). This supports the notion that Bayesian Kriging can provide improved interpolation performance, especially in regions with sparse data, as also noted in (Verdin et al., 2015). Thus, the findings of this study not only align with earlier research but also highlight the potential of Bayesian approaches in enhancing spatial analysis in tropical regions.

D. CONCLUSION AND SUGGESTIONS

Bayesian Kriging is a parameter estimation approach in spatial interpolation method designed to overcome the limitations of sample size and the complexity of variogram structure. This study applies Bayesian Kriging to interpolate monthly rainfall in East Java. Semivariogram parameters including intercept, sill, range, and nugget are estimated using a Bayesian approach with informative priors. The reliability of parameter estimation is proven through trace plots, autocorrelation plots, and Monte Carlo error analysis, which demonstrate convergence and stability of the posterior samples. Two theoretical semivariogram models were used, namely spherical and exponential. The analysis shows that the spherical semivariogram model provides more accurate interpolation results than the exponential model, as indicated by lower RMSE values (spherical: RMSE = 37.17 mm vs. exponential: RMSE = 40.22 mm). This highlights the spherical model's better capacity to capture medium-range spatial dependencies commonly found in rainfall patterns across East Java.

The novelty of this study lies in the use of informative priors to improve estimation accuracy under limited data conditions, which has rarely been implemented in previous rainfall interpolation studies in Indonesia. Thus, it can be concluded that the Bayesian Kriging approach with the spherical model is better able to represent the spatial structure of monthly rainfall data in East Java. The use of informative priors also makes this method effective despite limited data. The interpolated maps (Figure 8) clearly show the highest rainfall concentrated in the southwestern part of East Java, particularly in districts such as Nganjuk, Madiun, dan Ponorogo, with estimated values more than 300 mm/month. These findings indicate that the research objectives namely to evaluate the performance of Bayesian Kriging with informative priors in interpolating monthly rainfall have been successfully addressed.

Based on the results obtained, future research is recommended to explore the Bayesian Kriging approach by considering other semivariogram models such as Gaussian or Matérn, which offer greater flexibility in representing smooth or complex spatial correlations. Additionally, the inclusion of environmental covariates such as elevation, distance from the coast, or land cover is expected to improve model performance by accounting for key physical factors influencing rainfall distribution. In addition, testing for longer time periods or different seasons can also enrich the understanding of the spatial dynamics of rainfall in East Java. the interpolated map shows that some areas, especially in the central and western parts of East Java, receive high monthly rainfall. Therefore, local governments and stakeholders are advised to increase vigilance against potential flooding and optimize water management infrastructure in these areas. For instance, regions like Ponorogo and Nganjuk consistently showed interpolated rainfall above 300 mm/month in this period, suggesting priority zones for flood mitigation planning. Conversely, areas with lower rainfall can be prioritized in planning irrigation systems and water conservation, especially to support the agricultural sector. Dry zones identified in the east such as Banyuwangi may benefit from targeted irrigation investments. Thus, the results of this study can be used as a basis for spatially-based land use planning and hydrometeorological risk mitigation.

ACKNOWLEDGEMENT
The researcher expressed gratitude for support from many parties.

REFERENCES

Althoff, D., & Rodrigues, L. N. (2021). Goodness-of-fit criteria for hydrological models: Model calibration and performance assessment. *Journal of Hydrology*, *600*, 126674. https://doi.org/10.1016/j.jhydrol.2021.126674

Astutik, S., Astuti, A. B., Efendi, A., Darmanto, D., Irsandy, D., & Saniyawati, F. Y. D. A. S. (2023). *Analisis Bayesian: Teori dan Aplikasi dengan R*. Universitas Brawijaya Press.

Biswas, A., & Biswas, A. (2024). Geostatistics in Soil Science: A Comprehensive Review on Past, Present and Future Perspectives. *Journal of the Indian Society of Soil Science*, *72*(1), 1–22. https://doi.org/10.5958/0974-0228.2024.00017.1

Bostan, P. (2017). Basic Kriging Methods in Geostatistics. *Yüzüncü Yıl Üniversitesi Tarım Bilimleri Dergisi*, *27*(1), 10–20. https://doi.org/10.29133/yyutbd.305093

Breunig, M., Bradley, P. E., Jahn, M., Kuper, P., Mazroob, N., Rösch, N., Al-Doori, M., Stefanakis, E., & Jadidi, M. (2020). Geospatial Data Management Research: Progress and Future Directions. *ISPRS International Journal of Geo-Information*, *9*(2), 95. https://doi.org/10.3390/ijgi9020095

Charles, T. D. S., Lopes, T. R., Duarte, S. N., Nascimento, J. G., Ricardo, H. D. C., & Pacheco, A. B. (2022). Estimating average annual rainfall by ordinary kriging and TRMM precipitation products in midwestern Brazil. *Journal of South American Earth Sciences*, *118*, 103937. https://doi.org/10.1016/j.jsames.2022.103937

Chutsagulprom, N., Chaisee, K., Wongsaijai, B., Inkeaw, P., & Oonariya, C. (2022). Spatial interpolation methods for estimating monthly rainfall distribution in Thailand. *Theoretical and Applied Climatology*, *148*(1–2), 317–328. https://doi.org/10.1007/s00704-022-03927-7

Cressie, N. A. C. (1993). *Statistics for spatial data* (2nd ed.). New York: John Wiley & Sons.

Efendi, A., Irsandy, D., Fitriani, R., & Munawar, M. (2023). Bayesian Spatial Analysis for Minimum Wage Schemes in East Java, Indonesia. *International Journal of Agricultural and Statistical Sciences*, *19*(1), 33. https://doi.org/10.59467/IJASS.2023.19.33

Erten, G. E., Yavuz, M., & Deutsch, C. V. (2022). Combination of Machine Learning and Kriging for Spatial Estimation of Geological Attributes. *Natural Resources Research*, *31*(1), 191–213. https://doi.org/10.1007/s11053-021-10003-w

Gao, Y., Meng, H., Pei, T., & Liu, Y. (2024). Understanding Spatial Dependency Among Spatial Interactions. In X. Meng, X. Zhang, D. Guo, D. Hu, B. Zheng, & C. Zhang (Eds.), *Spatial Data and Intelligence* (Vol. 14619, pp. 28–43). Singapore: Springer Nature Singapore. https://doi.org/10.1007/978-981-97-2966-1_3

Gelfand, A. E., & Banerjee, S. (2017). Bayesian Modeling and Analysis of Geostatistical Data. *Annual Review of Statistics and Its Application*, *4*(1), 245–266. https://doi.org/10.1146/annurev-statistics-060116-054155

Gribov, A., & Krivoruchko, K. (2020). Empirical Bayesian kriging implementation and usage. *Science of The Total Environment*, *722*, 137290. https://doi.org/10.1016/j.scitotenv.2020.137290

Han, H., & Suh, J. (2024). Spatial Prediction of Soil Contaminants Using a Hybrid Random Forest–Ordinary Kriging Model. *Applied Sciences*, *14*(4), 1666. https://doi.org/10.3390/app14041666

Hodson, T. O. (2022). Root-mean-square error (RMSE) or mean absolute error (MAE): When to use them or not. *Geoscientific Model Development*, *15*(14), 5481–5487. https://doi.org/10.5194/gmd-15-5481-2022

Jamaludin, S., & Suhaimi, H. (2013). Spatial Interpolation on Rainfall Data over Peninsular Malaysia Using Ordinary Kriging. *Jurnal Teknologi*, *63*(2), 51–58. https://doi.org/10.11113/jt.v63.1912

Kanooni, A., & Amogein, E. F. (2025). Comparison of sensitivity of interpolation methods by rain gauge network density. *Journal of Irrigation Sciences and Engineering*, *47*(4), 105–122. https://doi.org/10.22055/jise.2025.48259.2140

Lima, C. H. R., Kwon, H.-H., & Kim, Y.-T. (2021). A Bayesian Kriging model applied for spatial downscaling of daily rainfall from GCMs. *Journal of Hydrology*, *597*, 126095. https://doi.org/10.1016/j.jhydrol.2021.126095

Lu, X., Fan, Y., Hu, Y., Zhang, H., Wei, Y., & Yan, Z. (2024). Spatial distribution characteristics and source analysis of shallow groundwater pollution in typical areas of Yangtze River Delta. *Science of The Total Environment*, *906*, 167369. https://doi.org/10.1016/j.scitotenv.2023.167369

Maulana, I. A., Prasetiyowati, S. S., & Sibaroni, Y. (2022). Prediction of Rainfall Classification of Java Island with ANN-Feature Expansion and Ordinary Kriging. *Jurnal Media Informatika Budidarma*, *6*(4), 1988. https://doi.org/10.30865/mib.v6i4.4556

Nicoletta, L., Daniele, S., Aristide, G., Diego, B., & Massimo, P. (2021). Spatial Interpolation Techniques For innovative Air Quality Monitoring Systems. *Chemical Engineering Transactions*, *86*, 391–396. https://doi.org/10.3303/CET2186066

Omre, H. (1987). Bayesian Kriging—Merging Observations and Qualified Guesses in Kriging. *Mathematical Geology*, *19*, 25–39. https://doi.org/10.1007/BF01275432

Rai, P. K., Mishra, V. N., & Singh, P. (2022). *Geospatial technology for landscape and environmental management: Sustainable assessment and planning* (1st ed.). Singapore: Springer. https://doi.org/10.1007/978-981-16-7373-3

Sanusi, W., Sidjara, S., Patahuddin, S., & Danial, M. (2024). A Comparison of Spatial Interpolation Methods for Regionalizing Maximum Daily Rainfall Data in South Sulawesi, Indonesia. *ITM Web of Conferences*, *58*, 04003. https://doi.org/10.1051/itmconf/20245804003

Schoot, R. V. D., Depaoli, S., King, R., Kramer, B., Märtens, K., Tadesse, M. G., Vannucci, M., Gelman, A., Veen, D., Willemsen, J., & Yau, C. (2021). Bayesian statistics and modelling. *Nature Reviews Methods Primers*, *1*(1), 1. https://doi.org/10.1038/s43586-020-00001-2

Schoot, R. V. D., & Miočević, M. (2020). *Small Sample Size Solutions: A Guide for Applied Researchers and Practitioners* (1st ed.). London: Routledge. https://www.taylorfrancis.com/books/9780429273872

Sharma, S., Aggarwal, V., & Yadav, M. P. (2021). Comparison of linear and non-linear GARCH models for forecasting volatility of select emerging countries. *Journal of Advances in Management Research*, *18*(4), 526–547. https://doi.org/10.1108/JAMR-07-2020-0152

Tanner, M. A. (1996). *Tools for Statistical Inference: Methods for the Exploration of Posterior Distributions and Likelihood Functions*. New York: Springer New York. https://doi.org/10.1007/978-1-4612-4024-2

Uddin, M. S., & Czajkowski, K. P. (2022). Performance Assessment of Spatial Interpolation Methods for the Estimation of Atmospheric Carbon Dioxide in the Wider Geographic Extent. *Journal of Geovisualization and Spatial Analysis*, *6*(1), 10. https://doi.org/10.1007/s41651-022-00105-1

Varga, B., Pereira, M., Kulcsár, B., Pariota, L., & Péni, T. (2023). Data-Driven Distance Metrics for Kriging-Short-Term Urban Traffic State Prediction. *IEEE Transactions on Intelligent Transportation Systems*, *24*(6), 6268–6279. https://doi.org/10.1109/TITS.2023.3251022

Verdin, A., Rajagopalan, B., Kleiber, W., & Funk, C. (2015). A Bayesian kriging approach for blending satellite and ground precipitation observations. *Water Resources Research*, *51*(2), 908–921. https://doi.org/10.1002/2014WR015963

Wesner, J. S., & Pomeranz, J. P. F. (2021). Choosing priors in Bayesian ecological models by simulating from the prior predictive distribution. *Ecosphere*, *12*(9), e03739. https://doi.org/10.1002/ecs2.3739

Zondervan-Zwijnenburg, M., Peeters, M., Depaoli, S., & Van De Schoot, R. (2017). Where Do Priors Come From? Applying Guidelines to Construct Informative Priors in Small Sample Research. *Research in Human Development*, *14*(4), 305–320. https://doi.org/10.1080/15427609.2017.1370966

Zorzetto, E., & Marani, M. (2020). Extreme value metastatistical analysis of remotely sensed rainfall in ungauged areas: Spatial downscaling and error modelling. *Advances in Water Resources*, *135*, 103483. https://doi.org/10.1016/j.advwatres.2019.103483