OPEN ACCESS 159

ISSN 2356-5462
http://socj.telkomuniversity.ac.id/ijoict/

Intl. Journal on ICT
Vol. 11, No. 2, Dec 2025. pp. 159-170
doi: 10.21108/ijoict.v11i2.10021

# Modified Genetic Algorithm in Isotropic Semivariogram Modelling: A Case Study of Groundwater Level in Sumatera Peatland

Kurnia Novita Sari[1,*], Ari Ratna Gumilang[2], Muhammad Rozzaq Hamidi[3], Anwar Efendi Nasution[4], Giraldi Fardiaz Kuswanda[5]

[1,2,3,4]*Department of Mathematics, Bandung Institute of Technology*
*Bandung, 40123, Indonesia*

[5]*Department of Civil Engineering, Trisakti University*
*Jakarta, 11440, Indonesia*

[*]ikur_ain@itb.ac.id

**Abstract**

The purpose of this study is to develop a method for estimating semivariogram parameters using a genetic algorithm (GA). GA is a numerical method that has been extensively applied, which is applicable to estimate semivariogram parameters. In this study, we propose the incorporation of additional parameter constraints into genetic algorithm. By introducing these constraints, we developed a new constrained optimization procedure referred to as the modified-GA (m-GA). The m-GA shows better performance than iterative least square (ILS). The application of spatial analysis to groundwater level (GWL) in peatland areas is still limited, especially semivariogram analysis. Thus, the m-GA is applied to GWL data containing 41 locations in Sumatera peatland and then compared with ILS. In this study, the m-GA is regenerated 5,000 and 100,000 times. The study shows that the spherical semivariogram model estimated using the m-GA provides the best performance, because both the model and kriging have the lowest root mean square error (RMSE) values, at 0.33910 and 0.85880, respectively. The combination of spherical semivariogram model with the m-GA produces optimal and accurate semivariogram parameters to support kriging interpolation on GWL peatland.

**Keywords:** Genetic Algorithm, Groundwater Level, Iterative Least Square, Jackknife Kriging, Peatland, Semivariogram

## I. INTRODUCTION

PEATLAND is a land type created by the accumulation of organic material that consists of dead vegetation that has been submerged in water for a long period. This condition inhibits the decomposition process by microbe [[1], [2], [3], [4]]. Globally, peatland is distributed across more than 180 countries, including Indonesia [5]. Indonesia hosts the largest tropical peatlands in Asia-Pasific [6]. Indonesia is the fourth country with the largest peatland ecosystem in the world, after Canada, Russia, and United States of America, with an area of 24.668 million hectares [7].

Peatland has an important role in climate change mitigation due to its capability to store large amounts of carbon [8]. Based on the latest data about the area and depth of peatland, it is estimated that 38 tropical peatlands store 152 – 288 GtC or about half of total carbon stored in global peatlands [9]. Moreover, peatlands serve as natural water reservoirs that maintain the stability of water flow during both rainy and dry seasons and support the balance of hydrology and ecosystem [10]. One of the main factors influencing peatland drought is the decline

Received on 15 Oct 2025. Revised on 17 Nov 2025. Accepted and Published on 2 Jan 2025.

in groundwater level (GWL). The GWL is affected by various factors, such as rainfall, tidal fluctuations, drainage development, land conversion, and land clearing [11]. Therefore, proper management of GWL is required to maintain peatland moisture and prevent further degradation.

Peatland GWL data are classified as spatial data because they consist of random variables associated with specific location. Previous studies have extensively examined aspects of spatial variability and hydrology in peatlands through various methods. Mukhaiyar et al. [12] applied Generalized Space-Time Autoregressive (GSTAR) model where the spatial weight matrix was constructed using the Minimum Spanning Tree (MST) approach and then using Kriging in predicting risk mapping of peatland GWL. Mukhaiyar et al. [13] also applied GSTAR modelling with three-dimensional spatial weight matrix to predict the peatland GWL. Mukhaiyar et al. [14] used ARIMA model to predict the groundwater level and Ordinary Kriging for estimating unobserved locations. Situngkir et al. [15] applied the combination of logistic regression and Monte Carlo simulation for mapping the risk of peatland fires. These studies provide a strong foundation for further research to understand the spatial patterns of GWL in peatlands, which is important to support the sustainable management and conservation of peatlands.

The estimation of semivariogram parameters is a crucial step in geostatistical analysis, particularly for understanding the spatial variability in data such as peatland GWL. Ordinary Least Squares (OLS) is one of the commonly used methods for estimating semivariogram parameters [[16], [17]], and other techniques, including Iterative Least Squares (ILS), Weighted Least Squares (WLS), and Maximum Likelihood (ML), have also been applied. With the advancement of computational techniques, metaheuristic approaches such as the Genetic Algorithm (GA) have emerged as a promising alternative [[18], [19], [20]]. GA is capable of exploring the solution space globally, reducing the risk of being trapped in local minima [21]. Inspired by the process of natural evolution, GA incorporates key biological concepts such as selection, crossover, and mutation, making it well-suited for optimizing complex functions like semivariogram parameter estimation.

In this study, we apply a genetic algorithm (GA) to estimate the parameters of a semivariogram. To ensure that the algorithm produces solutions consistent with the characteristics of a theoretical semivariogram, certain modifications to the standard GA are required. Accordingly, this study develops a modified genetic algorithm (m-GA), in which specific constraints are imposed based on the boundaries of the theoretical semivariogram. Subsequently, the performance of m-GA will be compared with ILS on peatland GWL data. Subsequently, the performance of m-GA will be compared with ILS on peatland GWL data, and the resulting semivariograms from both methods will be validated using Jackknife Kriging to assess their accuracy in predicting values at unobserved locations.

## II. LITERATURE REVIEW

Observed spatial data represent real-world observations acquired through field measurement. The experimental semivariogram is the first step in modelling spatial data. Matheron proposed an experimental semivariogram formula as variance of the difference between pairs of observations separated by distance $h$. The distances of observations are grouped into several classes. There are three grouping methods according to [22], such as Sturges' rule, Freedman-Diaconis method, and Scott method. Sturges' rule can be applied if the number of distances $h$ is normally distributed. This is due to the normal-reference rule of Sturges' rule [23]. Therefore, there are other methods for skewed data, such as the Freedman-Diaconis method (based on the interquartile range) and the Scott method (based on the standard deviation).

An experimental semivariogram will be fitted to the theoretical semivariogram by estimating its parameters, such as nugget effect, partial sill, and range. The development of semivariogram parameter estimation has been extensively carried out using various methods. Sari et al. [16] applied the resampling method (bootstrap) for estimating the semivariogram model. Sari et al. [24] applied nonparametric Epanechnikov kernel, which performs better SSE than the ILS for estimating semivariogram parameters. Yu et al. [25] used deep learning to optimize the alignment between experimental and theoretical variogram functions. There are also numerical methods for estimating semivariogram parameters, such as GA. Applying GA to estimate semivariogram parameters has been conducted by [26]. Boroh et al. [26] integrated a genetic algorithm with machine learning to optimize variogram parameters by minimizing root mean square error (RMSE) and maximizing coefficients of determination (R2). In addition, Sequential Gaussian Simulations (SGS) were applied for predictive mapping. Li et al. [27] used GA to optimize the objective function that consists of experimental variogram fit

and interpolation accuracy. Moreover, constraints were applied into the weight of both fitting and interpolating components in objective functions.

In this study, we propose an m-GA, which adds some constraints based on the properties of theoretical semivariogram parameters. This method is a further development of the approach proposed by [28]. Kuswanda [28] determined that the nugget effect and partial sill are constrained to be less than the maximum experimental semivariogram, and the range is constrained to be less than the maximum lag distance. Moreover, the GA is compared with Gauss-Newton method for estimating the semivariogram parameters. The model validation is used through a resampling method, such as Jackknife and will be combined with Ordinary Kriging, named Jackknife Kriging [22].

## III. RESEARCH METHOD

This section contains the methodology used in this study. There are several methods such as calculating experimental semivariogram, fitting theoretical semivariogram, GA representation in semivariogram modelling, and ILS. The validation of the model uses jackknife kriging.

### *A. Semivariogram*

Semivariogram is a statistic that is used to figure out and model the spatial relationship between spatial data points [22]. The experimental semivariogram is the result of calculations obtained from field observations. Experimental semivariogram is formulated by Matheron (1965) as in (1).

$$\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} \left(Z(s_i + h) - Z(s_i)\right)^2 \tag{1}$$

with $\hat{\gamma}(h)$ is the experimental semivariogram at lag-$h$, $N(h)$ is the number of location pairs, and $Z(s_i)$ is the observed value at location $s_i$.

Since the number of location pairs can be large, classification of location pairs is needed using Sturges' rule. Sturges' rule (1926) is defined as (2).

$$K = 1 + 3.3 \log N(h) \text{ and } L = \frac{J}{K} = \frac{h_{\max} - h_{\min}}{K} \tag{2}$$

with $N(h)$ assumed to be normally distributed.

The experimental semivariogram will be fitted to the theoretical semivariogram by estimating its parameters. Theoretical semivariogram has three parameters, such as nugget effect ($C_0$), partial sill ($C$), and range ($a$). [22] explains that nugget effect represents the measurement error of the semivariogram measured at zero lag distance, range is the lag distance at which there is no spatial correlation between locations, and sill ($C_0 + C$) is the threshold value which there is no spatial correlation between locations. There are three theoretical semivariograms that will be used in this study, such as exponential, Gauss, and spherical.

$$\gamma_{\text{exp}}(h) = C_0 + C\left(1 - \exp\left(-\frac{|h|}{a}\right)\right), \; \gamma_{\text{gau}}(h) = C_0 + C\left(1 - \exp\left(-\left(\frac{|h|}{a}\right)^2\right)\right)$$

$$\gamma_{\text{sph}}(h) = \begin{cases} C_0 + C\left(\frac{3|h|}{2a} - \frac{1}{2}\left(\frac{|h|}{a}\right)^3\right), & 0 < |h| < a \\ C_0 + C, & |h| \geq 0 \end{cases}$$

The conditions that must be satisfied by semivariogram parameters are defined as (3). These conditions will be constraints in this study, which represents a novelty in the parameter estimation process.

$$0 \leq C_0 + C \leq \text{Var}\big(Z(s)\big), a \leq \frac{1}{2} h_{\max}, \text{and } C_0 \leq C \tag{3}$$

### *B. Genetic Algorithm*

Genetic algorithm (GA) is an optimization algorithm that is inspired by natural selection, known as evolution [21]. Evolution through various processes, including selection, crossover, and mutation on individual [[20], [29]]. In GA, the individuals that serve as candidate solutions are called chromosomes. These chromosomes consist of genes that represent the variables to be optimized. Chromosomes will go through evolution until the best chromosome is obtained. The best chromosome is chosen based on fitness value, which is a value in GA to evaluate how well a gene solves the problem [20].

GA is applied to estimate the parameters of theoretical semivariogram model. Nugget effect ($C_0$), partial sill ($C$), and range ($a$) serve as gen, thus all chromosomes have a pair of ($\hat{C}_0, \hat{C}, \hat{a}$). Fitness value consists of sum square error (SSE) that formulated as (4).

$$SSE = \sum_{i=1}^{n} \left(\gamma(h_i) - \hat{\gamma}(h_i; \hat{C}_0, \hat{C}, \hat{a})\right)^2 = \sum_{i=1}^{n} \left(\gamma(h_i) - \left(\hat{C}_0 + \hat{C}\, f\left(\frac{|h_i|}{\hat{a}}\right)\right)\right)^2 \quad (4)$$

The m-GA is implemented in this study by adding several constraints to the theoretical model defined in equation (3), in order to ensure that the estimated parameters remain consistent with the theoretical properties of a semivariogram. Without such constraints, GA may explore parameter values outside the theoretically valid solution space, potentially producing results that are not geostatistically meaningful. These constraints are applied throughout all stages of the GA evolution process, including initialization, crossover, and mutation.

During initialization, the estimates $\hat{C}_0$and $\hat{C}$ are sampled from the interval [0,Var($Z(s)$)], while $\hat{a}$ is sampled from $[0, \frac{1}{2}h_{\max}]$. The selection process chooses the two individuals with the highest fitness values as parents for the crossover stage. Subsequently, each child undergoes mutation, where any mutated parameter is also restricted to remain within its corresponding interval.

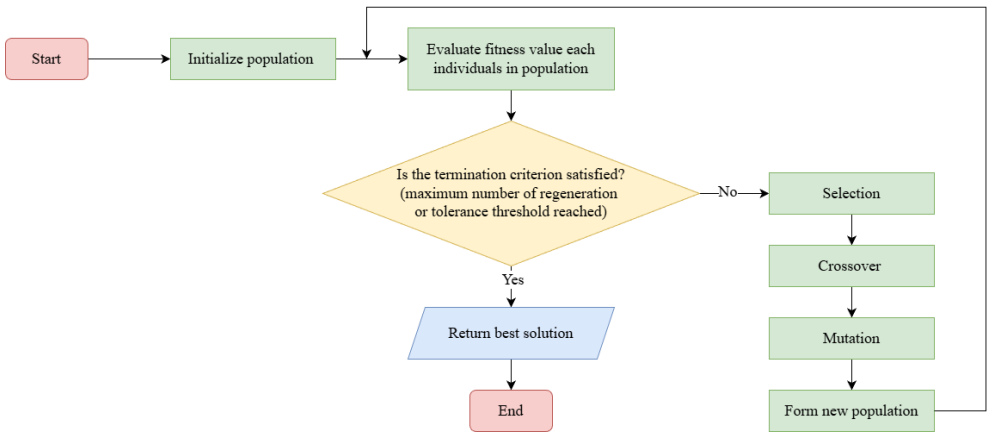


Fig. 1. Flowchart of Modified Genetic Algorithm (m-GA)

The evolutionary process is repeated for 5,000 and 100,000 regenerations according to the regeneration schemes. As a convergence criterion, the GA terminates when the maximum number of regenerations is reached or earlier if the improvement in fitness between generations falls below the tolerance threshold of 0.00001. Through this mechanism, all generated parameters remain theoretically valid while the GA maintains its ability to explore the solution space globally. The flowchart of m-GA is presented in Fig. 1, while the evolution scheme of the m-GA for estimating theoretical semivariogram parameters is illustrated in Fig. 2.

### *C. Iterative Least Square*

Least square is a method that minimizes SSE. Least square is modified in occur to estimate semivariogram parameters, because there is no analytical solution for range ($a$). Therefore, the range will be estimated iteratively, so that least square is modified into ILS. The procedure of ILS is as follows [22]:

1. Take the value of $a \in [h_{\min}, h_{\max}]$. Some literature stated that $a \in [h_{\min}, h_{\max}]$. Partition the value of as small as possible to determine $a$ accurately.
2. Calculate $\hat{C}_0$ and $\hat{C}$ which formulated in [24] for every $a$, then pairs of $(\hat{C}_0, \hat{C}, \hat{a})$ will be obtained.
3. Select a pair of $(\hat{C}_0, \hat{C}, \hat{a})$ that give minimum SSE.

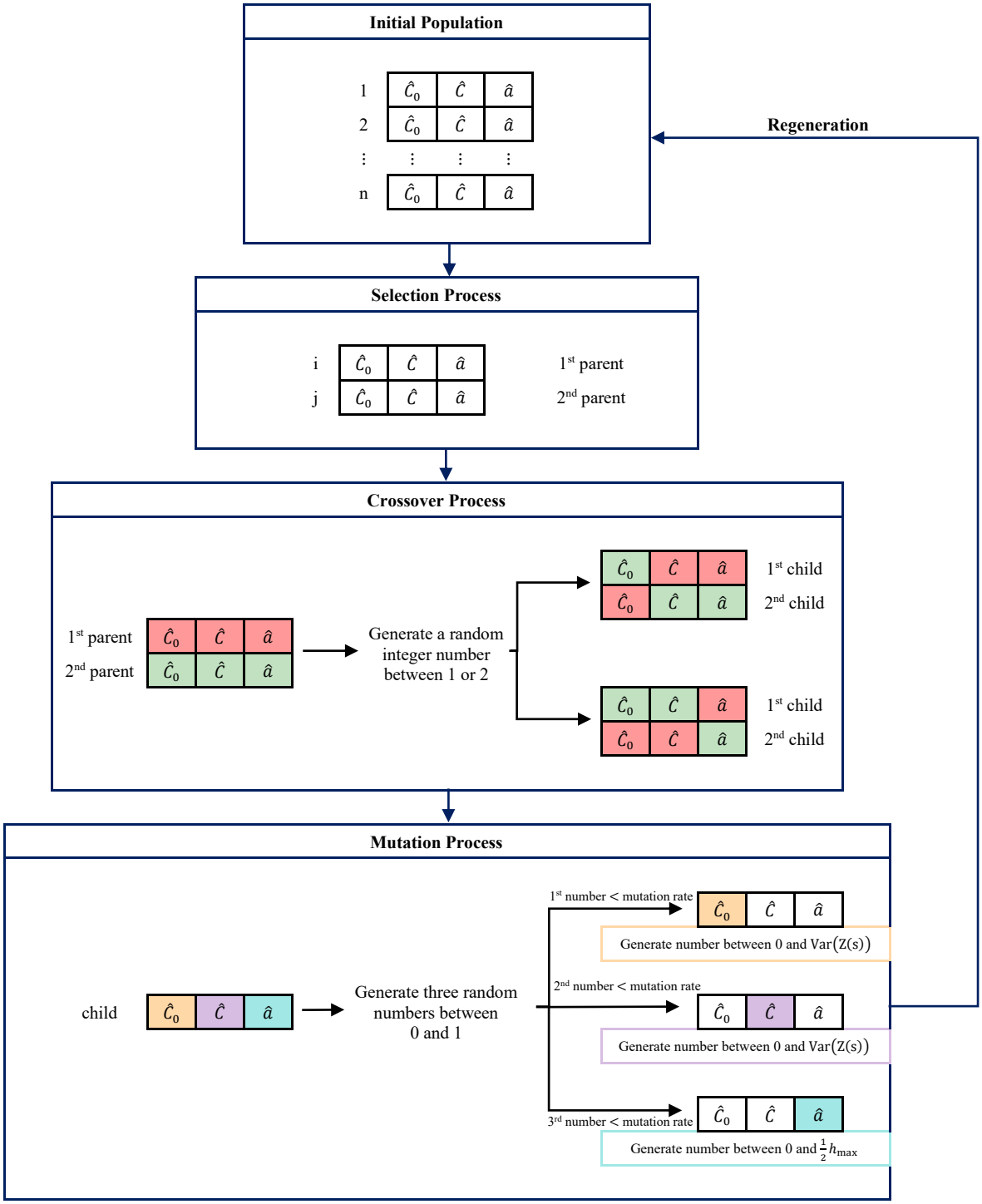


Fig. 2. Illustration of Modified Genetic Algorithm (m-GA) for Estimating Semivariogram Parameters

### *D. Jackknife Kriging*

Jackknife is resampling method purposing to calculate bias and variance from an estimation without any strict assumption [30]. The idea of this method is removing each data sequentially from the dataset, then

estimate the data that has been removed with the model [31]. Jackknife method will be combined with Ordinary Kriging to evaluate semivariogram models. [[22], [31]] shows the procedure of Jackknife Kriging as follows:

1. Remove the data at the $i$-th location, which $i = 1,2,3, \dots, n$.
2. Estimate the parameter of semivariogram model with remaining data.
3. Estimate $Z(s_i)$ at the $i$-th location that has been removed with Ordinary Kriging.
4. Calculate $e_i = \left(\hat{Z}(s_i) - Z(s_i)\right)^2$, which $\hat{Z}(s_i)$ is an estimated data at the $i$-th location and $Z(s_i)$ is an actual data at the $i$-th location.
5. Return the data at the $i$-th location that has been removed.
6. Repeat steps 1-5 for each location.
7. Calculate SSE of every semivariogram model.
8. Select semivariogram model that has minimum SSE.

## IV. Results And Discussion

### A. Data Exploration

Data in this research is groundwater level (GWL) data which is distributed in Sumatera. Data was collected from website Pranata Informasi Restorasi Ekosistem Gambut (PRIMS) which was managed by Badan Restorasi Gambut dan Mangrove (BRGM) [32]. GWL values are expressed in meters. The GWL data are located at latitudes -3.02° to 2.12° and longitudes 100.492° to 105.232°. Table I shows statistical descriptions of the data. Mean of GWL is -0.591 below the ground surface. The standard deviation is 0.849, which reflects a moderate level of data variation. Minimum data is -4.3, indicating the presence of an area with very low GWL. Maximum data is 0.688, showing that there is an area with GWL. Median of the data is -0.382, higher than the first quartile (-0.602) and lower than the third quartile (-0.154), indicating that the distribution of the data is negatively skewed. This is confirmed by the skewness (-2.503), indicating a highly left-skewed distribution. The kurtosis of the data is 7.621, which shows that the distribution of the data is more leptokurtic than normal distribution.

TABLE I
STATISTIC DESCRIPTIONS OF GWL IN SUMATERA

| Number of data | Mean | Standard Deviation | Minimum | Q1 | Median | Q3 | Maximum | Skewness | Kurtosis |
|---|---|---|---|---|---|---|---|---|---|
| 41 | -0.591 | 0.849 | -4.300 | -0.602 | -0.382 | -0.154 | 0.688 | -2.503 | 7.621 |

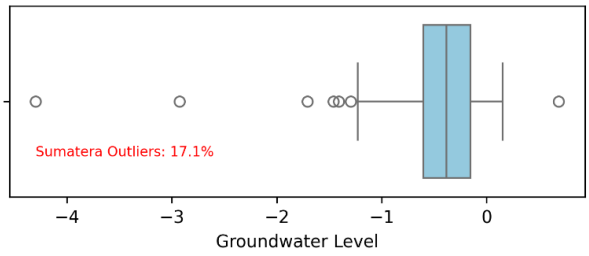


Fig. 3. Boxplot of Groundwater Level in Sumatera

Boxplot of GWL data in Sumatera is presented in Fig. 3. supporting the interpretation from Table I. There are some outliers detected accounting for 17.1%, especially in the lower range, indicating the presence of areas with GWL much lower than the average. This implies that there are extreme local variations at some locations.

Fig. 4. shows the spatial distribution of GWL in Sumatera. Scatterplot on the left shows that nearly all locations have low GWL (showed by red), concentrating in north area. While the higher one (showed by blue), concentrates locally in southeast area. Contour map on the right gives a smooth directional pattern. The directional pattern indicates a decreasing GWL gradient from southeast toward north and central areas, suggesting a possible influence of topography and local hydrological conditions.

### B. Semivariogram Modelling

The first step of semivariogram modelling is calculating experimental semivariogram. Experimental semivariogram strongly depends on the grouping of location pairs based on the distance between them. In this study, the grouping of location pairs use Sturges' rule as defined in (2). From the Sturges' formula, it was found that the number of classes is $K = 11$ and the length of classes is $L = 0.63248$. Experimental semivariogram is calculated using the formula in (1). The results of the distance classification and experimental semivariogram are presented in Table II.

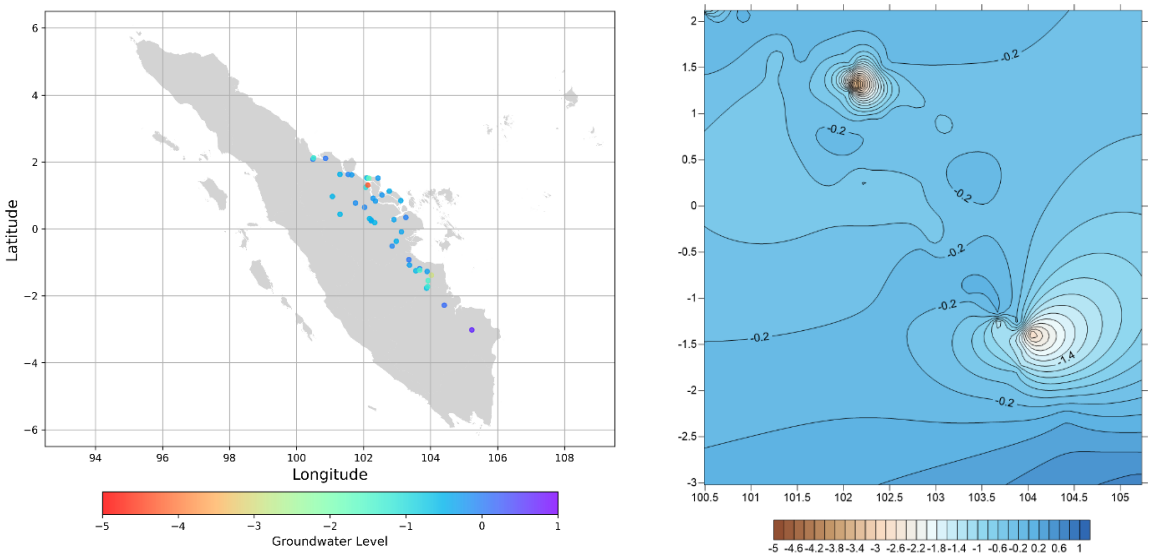


Fig. 4. (left to right) Scatterplot and Contour Map of Groundwater Level in Sumatera.
Data source: PRIMS (https://prims.brgm.go.id/)

From the Table II, when the number of pairs observations ($N(h)$) decreases with increasing the lag distance ($h$). Meanwhile, the experimental semivariogram $(\hat{\gamma}(h))$ varies at every lag distance. Moreover, experimental semivariogram using Matheron's method for GWL is shown as dot-dash line in Fig. 5. It shows that there is no trend in experimental semivariogram. Experimental semivariogram reaches maximum at lag-9, indicating strongly spatial relationships among all lags. Therefore, all of lag distance are included in semivariogram modelling, because of strong spatial relationship at a large lag distance. The calculated experimental semivariogram will be estimated to be theoretical semivariogram using ILS and GA.

TABLE II
DISTANCE CLASSIFICATION AND EXPERIMENTAL OF GWL IN SUMATERA

| Class Interval | $h$ | $N(h)$ | $\hat{\gamma}(h)$ |
|---|---|---|---|
| (0.02309,0.65557] | 0.33933 | 120 | 0.85533 |
| (0.65557,1.28806] | 0.97181 | 185 | 0.66017 |
| (1.28806,1.92054] | 1.60430 | 139 | 0.66143 |
| (1.92054,2.55302] | 2.23679 | 128 | 0.55771 |
| (2.55302,3.18551] | 2.86927 | 91 | 0.85881 |
| (3.18551,3.81800] | 3.50176 | 71 | 0.63260 |
| (3.81800,4.45048] | 4.13424 | 37 | 0.89825 |
| (4.45048,5.08297] | 4.76673 | 28 | 0.72217 |
| (5.08297,5.71545] | 5.39921 | 12 | 1.70059 |
| (5.71545,6.34799] | 6.03170 | 6 | 0.53614 |
| (6.34799,6.98042] | 6.66418 | 3 | 0.80190 |

TABLE III
PARAMETER ESTIMATION AND SSE OF THEORETICAL SEMIVARIOGRAM WITH ILS

| Model | Parameter Estimation | | | SSE |
|---|---|---|---|---|
| | $\hat{C}_0$ | $\hat{C}$ | $\hat{a}$ | |
| Exponential | 0.63692 | 0.20354 | 1.42721 | 1.00258* |
| Gauss | 0.68621 | 0.13511 | 2.66688 | 1.00947 |
| Spherical | 0.62756 | 0.15258 | 6.65786 | 1.96368 |

* showing the model with the best fit

ILS is applied to estimate the parameters of theoretical semivariogram, such as exponential, Gauss, and spherical. The result of the theoretical semivariogram parameters is shown in Table III. The best model is exponential model with smaller SSE than the other models. Besides experimental semivariogram, Fig. 5.

presents theoretical semivariograms that have been fitted using ILS. Theoretical semivariograms are marked as colored continuous lines.

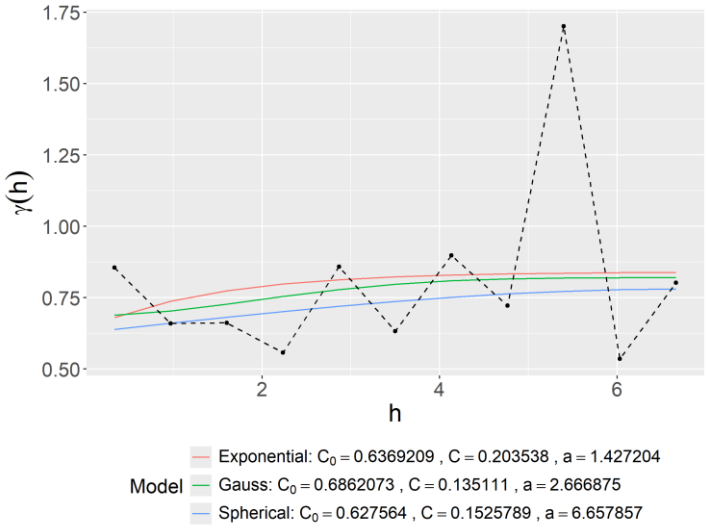


Fig. 5. Experimental and Theoretical Semivariogram using ILS

Fig. 5. shows that spherical model position under exponential and Gauss model. This pattern indicates an underestimated potential from spherical models in the occurrence of spatial relationships representation in this observed area. The underestimation can be caused by the characteristics of spherical models which have longer range, thus spatial relationships still exist at larger lag distances. Despite the value of SSE spherical model (1.96368) is still competitive compared to exponential model (1.00258) and Gauss model (1.00947), the underestimate indicates that spherical model does not ideally represent spatial relationship. Therefore, exponential model with the smallest SSE is the perfect choice to represent spatial relationships of GWL in Sumatera. Exponential model that has been estimated by ILS is formulated as (5).

$$\gamma_{\exp}(h) = 0.63692 + 0.20354\left(1 - \exp\left(-\left(\frac{h}{1.42720}\right)\right)\right) \tag{5}$$

The other method that is used in this study is m-GA. The parameter estimation $(\hat{C}_0, \hat{C}, \hat{a})$ is formulated as an optimization problem and GA is adopted to search for the optimal solution. GA requires a constraint or interval value to obtain the estimation of semivariogram parameters. Table IV presents the interval of parameter estimation in this study. The parameter constraints are proposed by [22], which are defined in Equation (3).

TABLE IV
INTERVAL VALUE OF ESTIMATION PARAMETER IN MODIFIED GENETIC ALGORITHM

| Parameter | Interval |
|---|---|
| $C_0$ (Nugget Effect)<br>$C$ (Partial Sill) | $[0, \mathrm{Var}(Z(s_i))] = [0,0.850295]$ |
| $a$ (Range) | $\left[0, \frac{1}{2}h_{\max}\right] = [0,3.33209]$ |

Subsequently, the population size and mutation rate are defined. In this study, mutation rate is considered as a binomial distribution. Let $X \sim Binomial(3, p)$ state the number of genes that mutate, thus the probability of the child is calculated with binomial distribution. Fig. 6. shows the simulation of the probability of the mutate genes to an individual with some scenario of mutation rate.

Based on Fig. 6, $p = 0.5$ shows the optimal balance with the mutation probability for $X = 1$ and $X = 2$ is 0.375, while the mutation probability for $X = 0$ and $X = 3$ is lower than the other one, that is 0.125. This indicates a moderate distribution, with almost all genes mutated without affecting the stability of the solution.

This is important to maintain a balance between exploration (introducing new variation) and exploitation (preserving the best genes). Therefore, $p = 0.5$ is selected as the optimal value to ensure that the m-GA can search for solutions effectively without losing stability or causing overly drastic changes in the population.

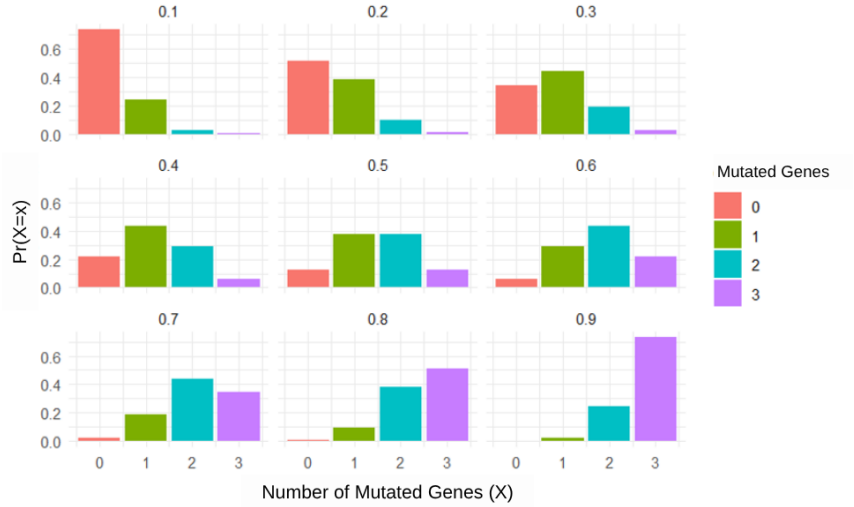


Fig. 6. Visualization of The Probability of Mutated Genes Based on Number of Mutated Genes

Thereafter, the population size is set at 10 individuals, which are randomly generated from the interval values defined in Tabel IV. This approach provides enough initial population, which is important to ensure wider solution exploration in the initial step of optimization. Fitness value is calculated based on SSE, which is formulated in (4). Fitness value is used to evaluate the quality of each individual in population. An individual with smaller SSE value is considered to have a good fitness value.

The search for solutions is performed randomly through selection, crossover, and mutation process, with a stopping criterion based on regeneration number. In this study, there are two regeneration schemes, such as 20 iterations with 5,000 regenerations per iteration and one iteration with 100,000 regenerations. The best result of these schemes is presented in Table V.

TABLE V
THE BEST SOLUTION OF MODIFIED GENETIC ALGORITHM WITH SOME REGENERATION SCHEMES

| Model | Number of Regeneration | $\hat{C}_0$ | $\hat{C}$ | $\hat{a}$ | Fitness Value |
|---|---|---|---|---|---|
| Exponential | 100,000 | 0.60022 | $1.276 \times 10^{-5}$ | 3.33203 | 0.78881 |
| | 5,000 | 0.66046 | $2.228 \times 10^{-4}$ | 3.32937 | 0.79000 |
| Gauss | 100,000 | 0.66000 | $4.869 \times 10^{-6}$ | 3.33175 | 0.78876* |
| | 5,000 | 0.66036 | $1.892 \times 10^{-6}$ | 3.33022 | 0.78980 |
| Spherical | 100,000 | 0.66003 | $1.203 \times 10^{-5}$ | 3.33194 | 0.78883 |
| | 5,000 | 0.66059 | $3.602 \times 10^{-4}$ | 3.33017 | 0.79060 |

* showing the model with the best fit

Gauss model with 100.000 regenerations, showing the lowest fitness value than the other models. This indicates that the greater the number of regenerations, the better the fitness value. Beside of it, there is no significant difference between the fitness value of 100.000 regenerations and 5.000 regenerations with 20 iterations. For optimizing computation, 5.000 regenerations with 20 iterations are enough to estimate the parameters. Moreover, the optimal result does not depend on number of regenerations, but also frequency of estimating and the efficiency of model structure. Based on Tabel V, Gauss model is expressed as (6).

$$\gamma(h) = 0.66000 + 4.869 \times 10^{-6}\left(1 - \exp\left(-\left(\frac{h}{3.33175}\right)^2\right)\right) \quad (6)$$

*C. Model Validation with Jackknife Kriging*

The main idea of semivariogram modelling is supporting estimation process of observations through Kriging method. Thus, the advantage of semivariogram model is determined not only by its fit to the experimental semivariogram but also by its accuracy in interpolating unobserved values. The validation model process is through Jackknife Kriging. Jackknife Kriging is designed to evaluate model performance based on its prediction ability to estimate the observed value.

TABLE VI
PARAMETER AND RMSE COMPARISON BASED ON ESTIMATION METHOD

| Model | Method | Estimation | | | RMSE | |
|---|---|---|---|---|---|---|
| | | $\widehat{C_0}$ | $\widehat{C}$ | $\widehat{a}$ | Model | $\widehat{C_0}$ |
| Exponential | ILS | 0.63692 | 0.20354 | 1.42721 | $0.30190^{(1)}$ | 1.42840 |
| | m-GA | 0.66046 | $2.228 \times 10^{-4}$ | 3.32937 | 0.33923 | $0.85883^{(2)}$ |
| Gauss | ILS | 0.68621 | 0.13511 | 2.66688 | $0.30278^{(2)}$ | 1.08846 |
| | m-GA | 0.66036 | $1.892 \times 10^{-6}$ | 3.33022 | 0.33934 | $0.85886^{(3)}$ |
| Spherical | ILS | 0.62756 | 0.15258 | 1.96368 | $0.30970^{(3)}$ | 2.23363 |
| | m-GA | 0.66059 | $3.602 \times 10^{-4}$ | 3.33017 | 0.33910 | $0.85880^{(1)}$ |

Table VI presents a comparison between ILS and m-GA for estimating semivariogram parameters in each theoretical model. The table reports both the RMSE Model, which measures the goodness of fit to the theoretical semivariogram, and the RMSE Kriging, which evaluates spatial prediction accuracy. The results clearly show that ILS achieves the lowest RMSE Model across all models—for example, the exponential model under ILS yields the best fit with an RMSE Model of 0.30190. In contrast, m-GA consistently produces lower RMSE Kriging values than ILS for every semivariogram model, indicating superior prediction performance. The spherical model optimized by m-GA provides the best predictive accuracy, with an RMSE Kriging of 0.85880.

The results show that the focus on only RMSE model is not enough to evaluate the overall performance of semivariogram parameters. Note that the RMSE model of ILS and m-GA is not significantly different, but the difference in their RMSE Kriging is quite significant. Thus, the main advantage of m-GA is the ability to optimize parameters in supporting spatial predictions more accurately, even though it produces a slightly higher RMSE model. The combination of spherical models with m-GA proves that providing best balance between RMSE model and RMSE Kriging, which is the most effective approach in this study.

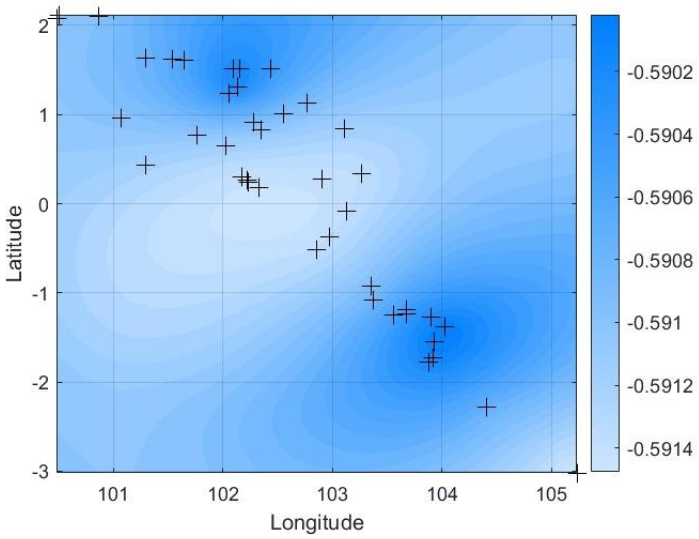


Fig.7. Interpolation of GWL using Jackknife Kriging with m-GA Spherical Model

Contour map that is shown in Fig. 7 illustrates the result of GWL interpolation using Ordinary Kriging from spherical model that is estimated by m-GA. The light blue presents a dry area, but still in the good condition. The observation locations are marked with plus symbol (+). This map is important for policy-making in

planning to restore the dry area and protect the wet area to keep the stability of peat ecosystem. The accuracy of m-GA algorithms makes an effective approach to supporting management and the sustainability of peatland conservation.

## V. CONCLUSION

This study compares the performance of m-GA and ILS for estimating semivariogram parameters of peatland GWL in Sumatera. The results show that m-GA consistently produces lower RMSE Kriging than ILS. Spherical model that is estimated with m-GA presents the best results with RMSE model is about 0.33910 and RMSE Kriging is about 0.85880. Validation using Jackknife Kriging supports the discovery of m-GA to obtain an optimal semivariogram parameter. This approach supports the spatial interpretation more accurately than ILS method. The m-GA methods can be applied to other datasets, although several limitations remain. The algorithm requires higher computational time as the dataset becomes larger, and its performance may still depend on the choice of the theoretical semivariogram model.

## REFERENCES

[1] C. Chen, L. Loft, and B. Matzdorf, "Lost in action: Climate friendly use of European peatlands needs coherence and incentive-based policies," *Environ Sci Policy*, vol. 145, pp. 104–115, 2023, doi: https://doi.org/10.1016/j.envsci.2023.04.010.

[2] J. Shi *et al.*, "Research Progress in the Field of Peatlands in 1990–2022: A Systematic Analysis Based on Bibliometrics," *Land (Basel)*, vol. 13, p. 549, Apr. 2024, doi: 10.3390/land13040549.

[3] J. Loisel *et al.*, "Expert assessment of future vulnerability of the global peatland carbon sink," *Nat Clim Chang*, vol. 11, no. 1, pp. 70–77, 2021, doi: 10.1038/s41558-020-00944-0.

[4] R. J. M. Temmink *et al.*, "Recovering wetland biogeomorphic feedbacks to restore the world's biotic carbon hotspots," *Science (1979)*, vol. 376, no. 6593, p. eabn1479, Oct. 2025, doi: 10.1126/science.abn1479.

[5] UNEP Global Peatlads Assesment, "The state of the world's peatlands: Evidence for Action toward the Conservation, Restoration, and Sustainable Management of Peatlands," Nairobi, 2022.

[6] Y. Choy and A. Onuma, "The Tropical Peatlands in Indonesia and Global Environmental Change: A Multi-Dimensional System-Based Analysis and Policy Implications," *Regional Science and Environmental Economics*, vol. 2, p. 17, Jul. 2025, doi: 10.3390/rsee2030017.

[7] P. Winarto, W. Y. Utomo, R. M. Subekti, W. Utami, A. Suwendar, and Sudarmono, *Corrective Action on Peatland Protection and Management in Indonesia Toward Sustainable Peatland Management 2019-2020*, Cetakan Pertama. Central Jakarta : Directorate of Peatland Degradation Control, Directorate General of Environmental Pollution and Degradation Control, Ministry of Environment and Forestry Republic of Indonesia, 2021.

[8] L. I. Harris *et al.*, "The essential carbon service provided by northern peatlands," *Front Ecol Environ*, vol. 20, no. 4, pp. 222–230, May 2022, doi: https://doi.org/10.1002/fee.2437.

[9] K. Ribeiro *et al.*, "Tropical peatlands and their contribution to the global carbon cycle and climate change," *Glob Chang Biol*, vol. 27, no. 3, pp. 489–505, Feb. 2021, doi: https://doi.org/10.1111/gcb.15408.

[10] M. E. Armanto and E. Wildayana, *Keunikan dan Kemampuan Lahan Gambut*, First Edition. Palembang: UPT Penerbit dan Percetakan Universitas Sriwijaya 2023, 2023.

[11] R. Nusantara, G. Anshari, and W. Ramadhan, "Fluktuasi Tinggi Muka Air Tanah Gambut Di Lahan Perkebunan Kelapa Sawit Desa Kubu Kecamatan Kubu Kabupaten Kubu Raya," *Jurnal Ilmu Lingkungan*, vol. 21, pp. 781–788, Jul. 2023, doi: 10.14710/jil.21.4.781-788.

[12] U. Mukhaiyar *et al.*, *Risk Mapping of Groundwater Level in Peatland Area Utilizing a Spatio- Temporal Model with Weight Constructed Based on Minimum Spanning Tree*. 2024. doi: 10.21203/rs.3.rs-4119220/v1.

[13] U. Mukhaiyar *et al.*, "The generalized STAR modelling with three-dimensional of spatial weight matrix in predicting the Indonesia peatland's water level," *Environ Sci Eur*, vol. 36, no. 1, p. 180, 2024, doi: 10.1186/s12302-024-00979-6.

[14] U. Mukhaiyar, A. W. Mahdiyasa, T. Prastoro, B. C. Suherlan, U. S. Pasaribu, and S. W. Indratno, "Spatial and Time Series Modelling for the Groundwater Level of Peatlands in Riau and Central Kalimantan, Indonesia," in *Decision Mathematics, Statistical Learning and Data Mining*, W. F. Wan Yaacob, Y. B. Wah, and O. U. Mehmood, Eds., Singapore: Springer Nature Singapore, 2024, pp. 89–104.

[15] A. P. O. Situngkir, A. W. Mahdiyasa, K. N. Sari, and U. S. Pasaribu, "Modelling Peatland Fire Risk and Economic Losses in Indonesia," in *2025 5th International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET)*, 2025, pp. 1–6. doi: 10.1109/IRASET64571.2025.11008022.

[16] K. N. Sari, U. S. Pasaribu, O. Neswan, and A. K. Permadi, "Estimation of the parameters of isotropic semivariogram model through bootstrap," *Applied Mathematical Sciences*, vol. 9, no. 103, pp. 5123–5137, 2015.

[17] F. Usman, G. M. Tinungki, and E. T. Herdiani, "Efficiency of Ni Content in Laterite Nickel Deposits through The Least Square Method Approach on Semivariogram," *J Phys Conf Ser*, vol. 2123, no. 1, p. 012015, 2021, doi: 10.1088/1742-6596/2123/1/012015.

[18] J.-S. Ryu, Y. Park, and K.-J. Cha, "The Estimation of Theoretical Semivariogram Adapting Genetic Algorithm for Kriging," *Commun Stat Appl Methods*, vol. 11, Jan. 2004, doi: 10.5351/CKSS.2004.11.2.355.

[19] S. S. Rocha, C. S. Pitombo, L. H. M. Costa, and S. de F. Marques, "Applying optimization algorithms for spatial estimation of travel demand variables," *Transp Res Interdiscip Perspect*, vol. 10, p. 100369, 2021, doi: https://doi.org/10.1016/j.trip.2021.100369.

[20] A. Salhi and R. Abo Alsabeh, "Genetic Algorithm: A study survey," *Iraqi Journal of Science*, vol. 63, pp. 1231–1251, Mar. 2022, doi: 10.24996/ijs.2022.63.3.27.

[21] S. Katoch, S. S. Chauhan, and V. Kumar, "A review on genetic algorithm: past, present, and future," *Multimed Tools Appl*, vol. 80, no. 5, pp. 8091–8126, 2021, doi: 10.1007/s11042-020-10139-6.

[22] R. K. N. Sari, *Bahan Ajar Geostatistika dan Aplikasinya*. Bandung: ITB Press, 2020.

[23] D. W. Scott, "Sturges' rule," *WIREs Computational Statistics*, vol. 1, no. 3, pp. 303–306, Nov. 2009, doi: https://doi.org/10.1002/wics.35.

[24] K. N. Sari, Y. Y. Budiman, U. S. Pasaribu, and A. Sonhaji, "Comparison Between Iterative Least Square and Nonparametric Epanechnikov Kernel in Semivariogram Modeling, Case study: Urban Land Cover in East Java Province," *ITM Web Conf.*, vol. 58, 2024, [Online]. Available: https://doi.org/10.1051/itmconf/20245804007

[25] S. Yu, L. Zhao, and S. Li, "Leveraging Deep Learning for Automated Experimental Semivariogram Fitting," *Atmosphere (Basel)*, vol. 16, p. 191, Feb. 2025, doi: 10.3390/atmos16020191.

[26] A. W. Boroh, A. B. Kenfack Fokem, M. L. Mfenjou, F. D. Hamat, and F. Mbounja Besseme, "Variogram modelling optimisation using genetic algorithm and machine learning linear regression: application for Sequential Gaussian Simulations mapping," *Artificial Intelligence in Geosciences*, vol. 6, no. 1, p. 100124, 2025, doi: https://doi.org/10.1016/j.aiig.2025.100124.

[27] Z. Li, X. Zhang, K. C. Clarke, G. Liu, and R. Zhu, "An automatic variogram modeling method with high reliability fitness and estimates," *Comput Geosci*, vol. 120, pp. 48–59, 2018, doi: https://doi.org/10.1016/j.cageo.2018.07.011.

[28] G. F. Kuswanda, "Penaksiran Parameter Variogram Robust pada Asuransi Harta Benda di Kota Bandung Menggunakan Metode Gauss-Newton dan Algoritma Genetik," Bandung Institute of Technology, Bandung, 2020.

[29] F. M. Isa, W. N. M. Ariffin, M. S. Jusoh, and E. P. Putri, "A Review of Genetic Algorithm: Operations and Applications," *Journal of Advanced Research in Applied Sciences and Engineering Technology*, vol. 40, no. 1, pp. 1–34, Feb. 2024, doi: 10.37934/araset.40.1.134.

[30] H. Friedl and E. Stampfer, "Jackknife Resampling," Jun. 2001, doi: 10.1002/9780470057339.vaj001.

[31] N. Rohma, "PERBANDINGAN PENDUGAAN METODE ORDINARY KRIGING DAN METODE ORDINARY KRIGING DENGAN TEKNIK JACKKNIFE," *MAp (Mathematics and Applications) Journal*, vol. 4, pp. 101–111, Feb. 2023, doi: 10.15548/map.v4i2.4736.

[32] PRIMS (Pranata Informasi Restorasi Ekosistem Gambut), "Groundwater Level Data in Sumatera," prims.brgm.go.id.