Selection of variogram model for spatial rainfall mapping using Analytical Hierarchy Procedure (AHP)

M.Z. Muhamad Ali and F. Othman*

Department of Civil Engineering, Faculty of Engineering, University of Malaya, 50603 Kuala Lumpur, Malaysia.

Abstract. In geostatistical analysis, spatial interpolation at any unmeasured point is done using the parameters of a variogram model that matches the experimental data. Several variogram models can be used, and the accuracy of the spatial map produced depends on the selection of the most appropriate variogram model that fits the spatial distribution of the experimental data. Therefore, in this study, a multiple-criteria decision-making method, i.e. Analytical Hierarchy Procedure (AHP), is used to evaluate and select the best variogram model for mapping spatial rainfall in the upper reaches of the Kelang River basin in Malaysia. Using daily rainfall data from 71 rain gauge stations, geostatistical analysis was done by the Ordinary Kriging interpolation method and 5 alternatives to variogram models, namely Spherical, Tetraspherical, Pentaspherical, Exponential, and Gaussian for spatial rainfall mapping. The accuracy of the spatial rainfall map was evaluated using four criteria of spatial interpolation error indicators, which are Root-Mean-Square Error (RMSE), Average Standard Error (ASE), Mean Standardized Error (MSE), and Root-Mean-Square Standardized Error (RMSSE). The results showed that the spherical model was ranked at the top for producing the best spatial rainfall map of the study area.

1. Introduction

In hydrological science, the visualization of spatial rainfall from point-based rainfall for water related analysis is very important. It is generated using a spatial interpolation method via either a deterministic or geostatistical approach.

Recently, the advanced development of the Geographical Information System (GIS) has facilitated the common use of the geostatistical method for producing spatial rainfall distribution. The availability of several interpolation methods in geostatistical analysis, namely built-in GIS applications, is an advantage for analyzing spatial rainfall data. One such method is the Kriging technique variant that requires a variogram model to compute data estimation at the interpolation point.

The Kriging technique is quite similar to Inverse Distance Weight (IDW) method that needs the weight of the observed data to estimate the value at certain location. The IDW derives the weight based on the distance of the interpolated point to the observed data point only. However, the Kriging incorporated the distance with the spatial correlation of the observed data arrangement. The spatial correlation is evaluated using the variogram model before the estimation at unobserved point can be done. Furthermore, several variogram models are available in ArcGIS tools for use in spatial interpolation using the Kriging technique and the most popular models are Spherical, Tetraspherical, Pentaspherical, Exponential, and Gaussian.

The diversity of geostatistical methods has encouraged researchers to study the comparative per-

*. Corresponding author. Tel.: +60379674584;
Fax: +60379675318
E-mail addresses: zaharifudin@siswa.um.edu.my (M.Z. Muhamad Ali); faridahothman@um.edu.my (F. Othman)

formance of interpolation. For instance, Meng et al. [1] conducted a comparative performance study of 7 interpolation techniques in GIS. In their study, the regression Kriging was evaluated as the best interpolation method based on 4 performance indicators, i.e. Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Pearson correlation coefficient, and Bayes factor. Another study on the Kriging interpolation method was conducted by Bargaoui, and Chebbi [2], but the research scope was to suggest the 3-D estimation of the variogram rather than a 2-D perspective.

In terms of the variogram, there are a few models that can be fitted to experimental data and practically only one should be selected for all experimental datasets [3]. Ly, et al. [3] studied seven variogram models to fit the daily rainfall data and evaluated the variogram models using a cross-validation technique and the RMSE factor. It was concluded that the Gaussian model was able to fit most of the daily rainfall data. The same finding was drawn by Othman et al. [4], whose analysis was conducted in a tropical river basin using the spatial correlations Geary's C and Moran's I as performance indicators. The both indicators were the statistical measurement of autocorrelation between observed spatial data. The Geary's C indicator was a measurement of local spatial autocorrelation of the observed data. In contrast, the Moran's I test was for global autocorrelation.

In the literature, many performance indicators can be considered to justify variogram model selection. However, fulfilling them concurrently remains questionable. This has made researchers only consider the most achievable indicator in their analyses. In GIS application, 5 performance indicators of spatial mapping are available: mean error, root-mean-square error, average standard error, mean standardized error, and root-mean-square standardized error [5]. It is recommended that all indicators be emphasized to evaluate the accuracy of spatial interpolation as well as the smoothness of the produced map.

To conduct such a task, multi-criteria decision-making tool is needed which is able to incorporate the indicators in the decision-making process such as Analytical Hierarchy Process (AHP). The AHP is a simple structured approach of the criterion, and the alternative decisions in hierarchy are formed to analyze the decision-making process. It is able to consider both numerical and non-numerical forms of criterion. Based on this reason, the AHP has been applied in vast application either to technical or non-technical fields.

In this study, the AHP method was applied to analyses of the decision-making process to justify the best variogram model that produced an accurate spatial rainfall map by the geostatistical method. By the Kriging interpolation technique, 5 alternatives to variogram models, namely Spherical, Tetraspherical, Pentaspherical, Exponential, and Gaussian are structured for evaluation based on 4 spatial interpolation indicators (root-mean-square error, average standard error, mean standardized error, and root-mean-square standardized error).

2. Location and datasets

2.1. Location
To present the AHP application to select the variogram model for spatial rainfall mapping, a rain gauge network in the upper Klang River basin was selected for this study as illustrated in Figure 1.

The upper Klang River Basin is mainly located in the Federal Territory of Kuala Lumpur and some parts of Selangor State. The upper part of this study area is virgin forest with an elevation of about 1200 m above sea level. Meanwhile, the downstream area is almost flat with an elevation of more or less 100 m above sea level. This area is fully developed and has high density of residential population.

The climate in the study area is influenced by the monsoon system, which is categorized in 4 seasons: two main monsoons and two transitional monsoons. The main monsoon seasons occur from December to March (also known as the northeast monsoon) and from June to September (also known as the southwest monsoon). The transitional monsoons occur from March to June and September to December. Usually, the study area receives heavy convective rainfall in March, April, October, and November. However, the mean annual and monthly rainfall is relatively uniform with about 2,400 mm per year and 280 mm per month.

2.2. Rainfall datasets
Within the study area, there are 3 rain gauge networks developed by the Drainage and Irrigation Department of Malaysia (DID), which uses the same type of rain gauge instrument but for different hydrological purposes. The main rain gauge network, the National Network, was established for water resource studies in the 1970s. This network consists of 30 rainfall stations that are unevenly distributed throughout the area. As for flood monitoring, a network was developed consisting of 19 stations in the study area. This network is equipped with a telemetry system for hydrological data transmission to the monitoring server. Besides, there is another network specifically constructed for the Stormwater Management and Road Tunnel (SMART) project of Kuala Lumpur city center that has 22 stations. The distribution of station locations is shown in Figure 1.

The total number of rain gauge stations in the study area is 71 stations. Within the study area,
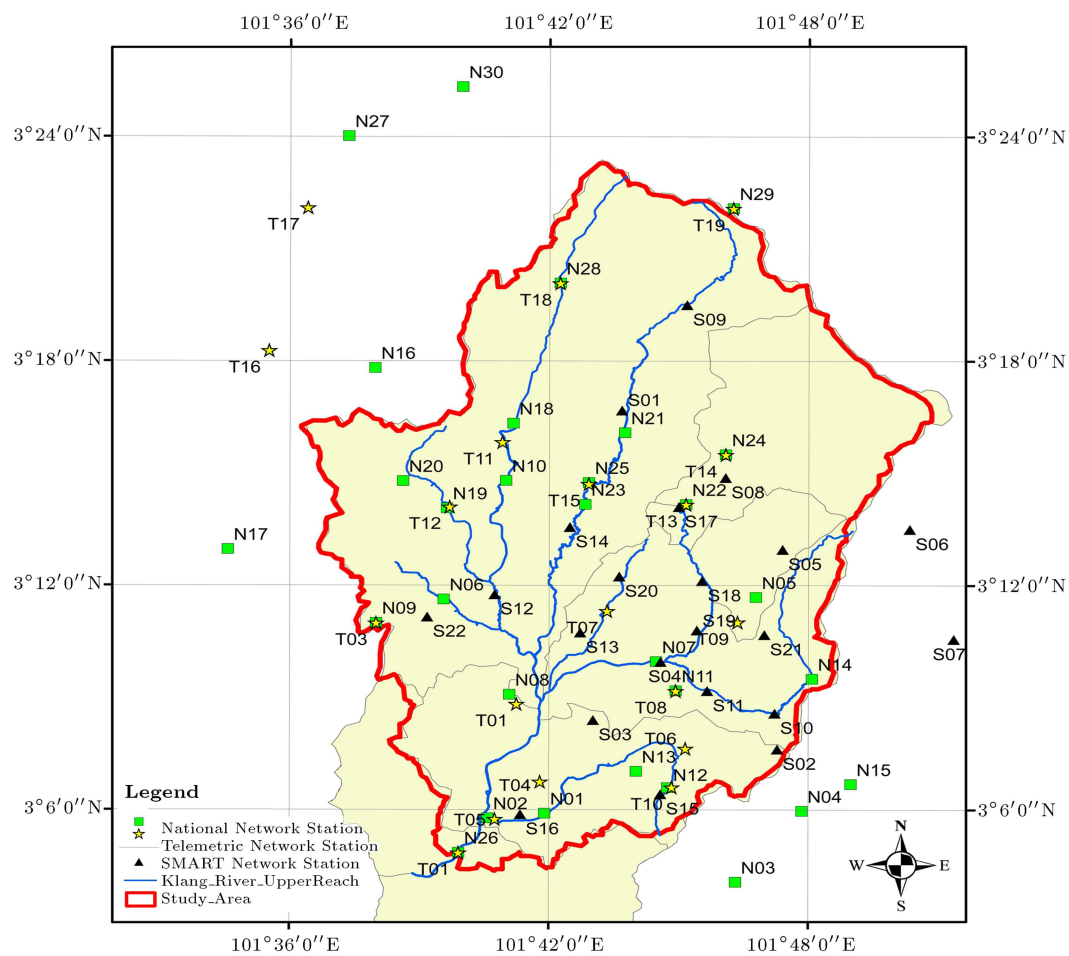
Figure 1. Study area and locations of rainfall stations used in this study.

Table 1. List of daily rainfall data events used in this study.

Year	Date	Minimum rainfall (mm)	Maximum rainfall (mm)	Mean rainfall (mm)	Standard deviation (mm)
2008	24 March 2008	0	109	13.5	23.44
2009	3 February 2009	0	100	19.1	25.15
	3 March 2009	0	134	38	28.01
2011	13 December 2011	0	142	36.1	30.47
	18 September 2011	0	87	31.1	24.08
2012	7 March 2012	0	239.5	78.4	55.09
	18 April 2012	0	111	19.3	25.49
	2 May 2012	0	137.5	65.5	32.14
	21 August 2012	0	130	24.5	27.74

the information of flood events that occurred recently back in the 5 years are determined from the DID. The availability of the rainfall data for each station is assessed for the flood events. Based on these two criteria, there are 9 flood events that are significant to use them in this study. The rainfall data in daily format were extracted from these 71 rainfall stations to perform the spatial rainfall mapping of the study area. The rainfall events selected in this study are tabulated in Table 1.

3. Methodology

The methodology of selecting the most appropriate variogram model for spatial rainfall mapping using the AHP is presented. First of all, geostatistical analysis is performed on 9 daily rainfall data to obtain the variography parameters (spatial variance, distance, nugget effect, and sill) for 5 variogram model candidates (Spherical, Tetraspherical, Pentaspherical, Exponential, and Gaussian). These parameter values are used to generate a spatial rainfall map with the prediction error indicators (root-mean-square error, average standard error, mean standardized error, and root-mean-square standardized error). The prediction error indicators are set as criteria in the AHP method and serve to evaluate the variogram model's performance to produce the best spatial rainfall map. To select the best variogram model for spatial rainfall mapping, the AHP method is used to rank the best variogram models for the final decision.

3.1. Geostatistical analysis

Geostatistical analysis is a way to study the environmental datasets related to space and time or the distribution of the datasets. Fundamentally, it consists of three main elements. First, the candidate data are characterized based on the correlation of spatial data distribution using the variogram model. Second, the variogram model is applied to estimate the space-based data through an optimal interpolation method. Then, the distribution is simulated to generate the datasets for the space domain using the variogram model. Further details of geostatistical analysis are explained by Journel and Huijbregts [6], Goovaerts [7], and Chiles and Delfiner [8].

In modeling the space correlation of an experimental dataset (in this study, it is daily rainfall data) using geostatistical analysis, Eq. (1) is used:

$$\gamma(h) = \frac{1}{2n} \sum \left[g(x) - g(x+h) \right]^2, \tag{1}$$

where $\gamma(h)$ is the semivariance, $[g(x) - g(x+h)]$ is the difference in dataset pair value, and n is the size of the dataset. Eq. (1) calculates the semivariance since the product on the right side is divided by 2. The semivariance is plotted against lag to produce the semivariogram.

Then, the experimental dataset is fitted to a curve that represents the spatial correlation of the variography parameters. In this study, the experimental semivariogram is fitted to the 5 variogram model candidates used to calculate the variography parameters for spatial rainfall mapping using the Kriging interpolation method.

The Kriging interpolation variants, namely Ordinary, Simple, Universal, Indicator, Probability, Disjunctive, and Co-Kriging, are associated with the geostatistical method. All of these Kriging variants quantify the spatial structure of the data and the prediction error [5]. The selection of the Kriging method is dependent on the dataset used in this study. Since there is only one variable, that is, the rainfall data used, and the constant mean of the dataset is assumed to be reasonable [5]. Then, the Ordinary Kriging (OK) method is adapted to the spatial interpolation stage. The geostatistical analysis is done using the ArcGIS 9.3 software toolbox.

3.1.1. Ordinary Kriging (OK) and variogram

The OK method is used to estimate the value of spatial interpolation based on assumption that the observed data has constant mean but unknown within the study area. The OK models the observed data using Eq. (2):

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

where $R(s)$ is the value for a location of (s) in coordinate of (x, y), μ is the constant mean, and $\varepsilon(s)$ is a random error.

To estimate the value R_{est} at prediction location, (s_o), OK uses the weighted average λ_i of the observed value, $R(s_i)$, as formed in Eq. (3). The weight λ_i: is calculated based on the distance of observed data to the prediction location and their spatial variation using variogram model. The sum of weight λ_i: must be equal to one to ensure that the predicted value is unbiased:

$$R_{\text{est}}(s_o) = \sum_{i=1}^{m} \lambda_i R(s_i). \tag{3}$$

In this study, 5 variogram models (Spherical, Tetraspherical, Pentaspherical, Exponential, and Gaussian) are evaluated to produce the best spatial rainfall map. Basically, the variogram models have similarity on the variography properties such as sill, range, and nugget as illustrated in Figure 2.

Theoretically, the semivariogram started at zero value. The measurement will be raised up to the sill where the line will be off or almost flat. However, due
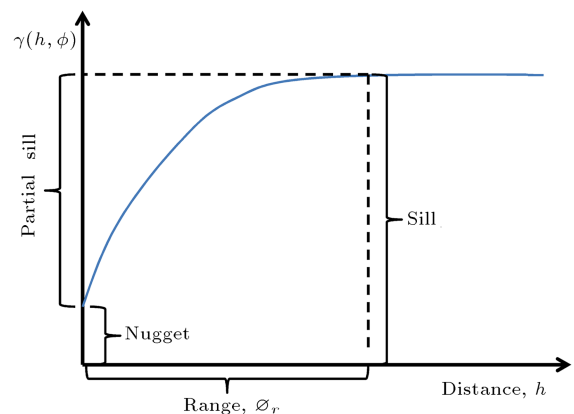
Figure 2. A typical structure of semivariogram.

Table 2. Mathematical equations for variogram models.

Variogram model	Mathematical model
Spherical	$\gamma(h,\phi) = \begin{cases} \phi_s \left[\frac{3}{2} \frac{\|h\|}{\varnothing_r} - \frac{1}{2} \left(\frac{\|h\|}{\varnothing_r} \right)^3 \right] & \text{for } 0 \leq \|h\| \leq \varnothing_r \\ \varnothing_s & \text{for } h \leq \|h\| \end{cases}$
Tetraspherical	$\gamma(h,\phi) = \begin{cases} \frac{2\varnothing_s}{\pi} \left[\arcsin \left(\frac{\|h\|}{\varnothing_s} \right) + \frac{\|h\|}{\varnothing_r} \sqrt{1 - \left(\frac{\|h\|}{\varnothing_r} \right)^2} + \frac{2}{3} \frac{\|h\|}{\varnothing_r} \left(1 - \left(\frac{\|h\|}{\varnothing_r} \right)^2 \right)^{\frac{3}{2}} \right] & \text{for } 0 \leq \|h\| \leq \varnothing \\ \varnothing_s & \text{for } \varnothing_r < \|h\| \end{cases}$
Pentaspherical	$\gamma(h,\phi) = \begin{cases} \phi_s \left[\frac{15}{8} \frac{\|h\|}{\varnothing_r} - \frac{5}{4} \left(\frac{\|h\|}{\varnothing_r} \right)^3 + \frac{3}{8} \left(\frac{\|h\|}{\varnothing_r} \right)^5 \right] & \text{for } 0 \leq \|h\| \varnothing_r \\ \varnothing_s & \text{for } h \leq \|h\| \end{cases}$
Exponential	$\gamma(h,\phi) = \phi_s \left[1 - \exp \left(-\frac{3\|h\|}{\varnothing_r} \right) \right]$
Gaussian	$\gamma(h,\phi) = \phi_s \left[1 - \exp \left(-3 \left(\frac{\|h\|}{\varnothing_r} \right)^2 \right) \right]$

to the measurement error, the semivariogram has offset value at the origin called nugget effect. This Nugget is a value of initial variability in the smallest group distance (lag). The sill value can be read at where the line off and partial sill value can be calculated by the residual of the sill and the nugget value. Range is the lag value extracted from the diagram at the sill location on the diagram which beyond this range, the autocorrelation measure is zero.

The dissimilarity between the variogram models is that how their mathematical equation presents the experimental data. The mathematical equation of variogram models are tabulated in Table 2. The spherical model presents the semivariogram curve linearly increasing at the early distance or lag and gradually change before it reaches the range. In contrast, the Gaussian model presents the parabolic form within the range value up to the sill. The Exponential model's curve is quite similar to Spherical model at the early distance, but exponentially increases the sill value as the distance increases the range. The Tetraspherical and Pentaspherical are differentiated by the mathematical formula from others variogram model. The Gaussian can be unstable numerically in the kriging interpolation if the nugget effect is not considered.

3.1.2. Spatial interpolation indicator

The accuracy of the spatial rainfall map produced by the geostatistical method is evaluated by spatial interpolation error using cross-validation approach. For this purpose, hold-one-out cross-validation technique is applied, where a station is removed one by one and the rainfall magnitude is estimated using the variogram parameters. Then, the spatial interpolation errors

are computed by Root-Mean-Square Error (RMSE), Average Standard Error (ASE), Mean Standardized Error (MSE), and Root-Mean-Square Standardized Error (RMSSE). The respective equations for each indicator are listed as below:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} (R_{\text{est}} - R_{\text{obs}})^2}{n}}, \tag{4}$$

$$\text{ASE} = \sqrt{\frac{\sum_{i=1}^{n} \sigma_i}{n}}, \tag{5}$$

$$\text{MSE} = \frac{\sum_{i=1}^{n} \left(\frac{R_{est} - R_{obs}}{\sigma_i} \right)^2}{n}, \tag{6}$$

$$\text{RMSSE} = \sqrt{\frac{\sum_{i=1}^{n} \left(\frac{(R_{\text{est}} - R_{\text{obs}})}{\sigma_i} \right)^2}{n}}, \tag{7}$$

where:

R_{obs} :   Observed rainfall value at the rain gauge;

R_{est} :   Interpolation estimate rainfall value at the rain gauge;

σ :   Standard error of estimated value at the rain gauge;

n :   Number of rain gauge station.

The best spatial rainfall map produced by the geostatistical method should have appropriate value of each indicator. The MSE should be close to zero, but the RMSSE is near to 1. However, the RMSE should be small and the ASE nears the RMSE value.

Table 3. Comparative weighting score for pairwise matrix comparison.

Criteria	Comparative weighting score																		Criteria
	More important than								Equal		Less important than								
C1	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	C2
C1	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	C3
C1	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	C4
C2	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	C3
C2	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	C4
C3	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	C4

Most researchers considered one or 2 indicators in their study, because these criteria are rarely achieved. In this study, this problem is solved using the multi-criteria decision-making tool.

3.2. Analytical Hierarchy Procedure (AHP)

The Analytical Hierarchy Process (AHP) was proposed by Saaty in the 1970s. The AHP is a methodology of multi-criteria decision-making for qualitative or quantitative study through an evaluation of a set of variables in the hierarchical structure [9]. It consists of 4 evaluation stages in the hierarchical structures, i.e. modeling, assessment, ranking, and conclusion. In the modeling stage of the hierarchical structure, the study objective is placed at the top of the structure. Meanwhile, the criteria, sub-criteria, and alternatives are structured at the bottom. To rank the alternatives for a decision, the criteria and sub-criteria are evaluated to set the priority weights used to assess the alternatives' attributes.

#### 3.2.1. Priority weight of criteria

The goal of this study is to select the best variogram model from 5 candidate models based on 4 prediction error criteria of spatial rainfall mapping. The AHP model used in this study is shown in Figure 3. Prior to executing AHP, the preference for each criterion is compared using the pair-wise matrix comparison [10]. A comparative scale with numerical values of 1 to 9 represents the influence of one criterion on another [11]. The criteria are labeled as C1 for RMSE, C2 for ASE, C3 for MSE, and C4 for RMSSE. These criteria are

compared and tabulated in Table 3, and the assignment of numerical values is based on the best output of spatial rainfall mapping. To achieve the best spatial rainfall mapping [5], criterion C2 must be the same as C1, which is the smallest value. Criterion C4 must be nearest to 1, but criterion C3 must approach 0.

In a real case, spatial rainfall mapping that fulfills all these criteria is difficult to achieve. Furthermore, the Consistency Ratio (CR) of the pairwise matrix must be less than 0.1 to ensure that the Priority Weights (PW) of the criteria are reliable for analysis, and it is the measurement of the consistency of the decision as structured by the AHP. The PW are calculated by first assigning the comparative numerical value of the criterion. For instance, in Set 1, a comparative numerical value 4 is assigned for C1 over C3. This means that C1 is more important than C3. In contrast, a comparative numerical value 0.11 (or 1/9) is assigned to C1 over C4. This means that the C4 is greatly more important than C1. After the pairwise matrix is created, each comparative value is normalized using the sum of each column according to Eq. (8), where X_{ij} is the normalized comparative numerical value, C_{ij} is the comparative numerical value, $\sum_{i=1}^{n} C_{ij}$ is the sum of columns for the comparative value, and notations of i and j are the row and column, respectively:

$$X_{ij} = \frac{C_{ij}}{\sum_{i=1}^{n} C_{ij}}. \tag{8}$$

Then, the PW are calculated by averaging the normalized value of comparative value for each row using
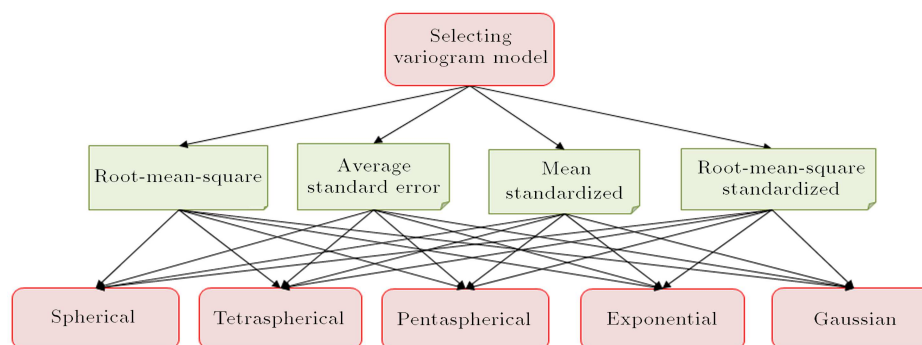
Figure 3. AHP model for variogram model selection.

Table 4. Random index for number of criteria [10].

Number of criteria, n	1	2	3	4	5	6	7
RI	0	0	0.58	0.9	1.12	1.24	1.32

Eq. (9):

$$PW_{ij} = \frac{\sum_{j=1}^{n} X_{ij}}{n},$$ (9)

where the $\sum_{j=1}^{n} X_{i,j}$ is the sum of normalized comparative value of row, and n is number of criterion.

The CR value is computed using Eq. (10), where the CI is the Consistency Index which is calculated using Eq. (11) and the RI is the Random Index. The RI value represents the random consistency of the developed pairwise matrix, and the value is based on the number of criteria in AHP structure as tabulated in Table 4. Since there are 4 criteria in this study, the RI value is 0.9. The $\lambda_{\max}$ in Eq. (11) is the largest eigenvalue of PW of each criterion:

$$CR = \frac{CI}{RI},$$ (10)

$$CI = \frac{\lambda_{\max} - n}{n - 1}.$$ (11)

The numerical value selection of a pairwise matrix is crucial, because it determines the consistency and robustness of the decision made. To achieve this, the pairwise matrix comparison using comparative numerical is simulated until the CR value is less than 0.1. As a result, there are 3 sets of pairwise matrices with different comparative numerical values, and the associated priority weights are used in the analysis. A detailed calculation of PW and the CR is given by Coulter et al. [12], Ishizaka and Lusti [13], and Saaty [14].

3.2.2. Alternative evaluation

The alternatives (variogram models) are evaluated using the criterion value and the PW value. To rank the alternatives for decision making, Eq. (12) is used to calculate the Alternatives' Weighted Priority (AWP) value, where I_i is the normalized criterion value, and n is the number of criterion. The I_i value for each criterion is calculated by dividing the criterion value with the respective mean, thus transforming the values to dimensionless. The alternatives' rank position is produced by sorting the AWP values from the highest to the lowest:

$$AWP = \sum_{i=1}^{n} I_i.PW_i.$$ (12)

The AHP method produced a rank for decision making for a single data set. However, in this study, 9 rainfall

Table 5. S_r value for each rank position.

Rank Position, r	1	2	3	4	5
S_r	5	4	3	2	1

data sets are used to select the best variogram model. The AHP results in a different rank set for each data set, and it is unable to make a decision from multi-data sets. To overcome this problem, a new set of scores is introduced by multiplying the probability (f_r) of the alternatives placed at a rank position by S_r value. The value of f_r is computed by counting the frequency of the alternatives placed at a rank position divided by the total number of datasets (9 in this study). The S_r value is a factor that consists of a set of numbers in a descending order from 5 (number of alternatives) to 1. The S_r value is assigned according to rank position as tabulated in Table 5. The Final Score is then calculated by summing up the scores as in Eq. (13), where r is the rank position. The Final Score value is sorted in a descending order to produce the final rank for decision-making. The alternative ranked in the top place is the best model to use for spatial rainfall mapping:

$$\text{Final score} = \sum_{r=1,2,3...} f_r.S_r.$$ (13)

4. Results

4.1. Priority weight of criteria

The results of priority weights in pairwise matrix analysis for criteria versus goal in the study are presented in Table 6. As mentioned earlier, there are 3 sets of priority weights employed in this study. All sets indicate satisfactory Consistency Ratio values of less than 0.1 to ensure the consistency of calculated priority weight. These 3 weight sets were used to observe the results of different levels of influence among the criteria. The CR value in Set 1 is 0.095, i.e. 9.5% inconsistency in decision, which is the highest; in Set 2, the CR value is the lowest (0.012 or 1.2% inconsistency). Whereas CR value in Set 3 has moderate CR value that is 0.039 (3.9% inconsistency).

In Set 1, criterion C4 had 71.3% influence on the decision, but criterion C3 had the least influence with only 4.6%. Criteria C1 and C2 had the equal influence of 12.1%. However, in Sets 2 and 3, criterion C4 still had the greatest weight, but criterion C3 increased up to 26.7% in Set 2 and 20.6% in Set 3. Meanwhile, criteria C1 and C2 had influences of only 6% and 8.5% on the decision, respectively. In other words, in Set 1, the dominant criterion was C4 (Root Mean Square Standardized Error), and there was no emphasis on criterion C3 (Mean Standardized Error). On the other hand, in Sets 2 and 3, criterion C4 remained important with the same contribution as criterion C3 on the evaluation process. These 3 scenarios were

Table 6. Results of priority weight and its associated comparative numerical value and consistency ratio of pairwise matrix comparison of the criteria versus objective.

Set no.	Priority weight	Factor	C1	C2	C3	C4	Consistency ratio
1	0.121	C1	1.00	1.00	4.00	0.11	
	0.121	C2	1.00	1.00	4.00	0.11	0.095
	0.046	C3	0.25	0.25	1.00	0.11	
	0.713	C4	9.00	9.00	9.00	1.00	
2	0.060	C1	1.00	1.00	0.20	0.11	
	0.060	C2	1.00	1.00	0.20	0.11	0.012
	0.267	C3	5.00	5.00	1.00	0.33	
	0.612	C4	9.00	9.00	3.00	1.00	
3	0.085	C1	1.00	1.00	0.33	0.17	
	0.085	C2	1.00	1.00	0.33	0.17	0.039
	0.206	C3	3.00	3.00	1.00	0.20	
	0.623	C4	6.00	6.00	5.00	1.00	

employed to ensure the robustness of the decision made by AHP.

4.2. Candidate model performance

The alternatives are evaluated using the final score value. Based on the final score value, the variogram models are ranked for decision-making. The results of the variogram model rank based on the Final Score value for each set are tabulated in Table 7. Based on the table, each set unanimously resulted in the same model ranke at the top, that is, the Spherical model. The minimum Spherical model's score value is 3.56 (in Set 2) compared with the maximum score in this study that is 5. This means that the Spherical model demonstrates a good performance to fulfill all criteria with at least 71.2%, and the maximum performance is 75.6% (score of 3.78 in Set 1).

The Gaussian and Tetraspherical models are ranked second and third in Set 2 and Set 3, respectively. In contrast, for Set 1, the Pentaspherical ranked second followed by the Gaussian model in the third place. The Gaussian model's score recorded is equal to 3.56 in Set 2 and Set 3 . This score is equal to the minimum score of the Spherical model. This result signifies that the Spherical and Gaussian models perform well in producing a spatial rainfall map that fulfills the criteria explained by Johnson et al. [5]. As for the rest, the performances of the variogram models are average and are ranked by the AHP method in the bottom place.

This result was compared with the findings by Othman et al. [4]. They conducted an analysis of a storm event on spatial rainfall distribution in the same study area. By using a total of 28 rain gauge stations, they found that the Gaussian model had a slightly better performance in producing spatial rainfall estimation compared to the Spherical and Exponential models. There is good agreement in the findings regarding the performance of the Gaussian and Exponential models. But, contradiction arises for the Spherical model. Applying the AHP method ranked the Spherical model at the top, indicating that the model has a good performance.

This is possibly due to 2 reasons, one of which is the number of rain gauge stations selected. There are 71 stations selected in this study, and at least 55 stations were used in the analysis based on data availability, which is twice more than the study by Othman et al. [4]. Thus, network density in this study is much greater. In geostatistical analysis, the accuracy of spatial rainfall distribution relies on the density of

Table 7. Results of variogram model ranking based on the final score value.

Pairwise matrix set		Set 1		Set 2		Set 3	
	1	**Spherical**	**3.78**	**Spherical**	**3.56**	**Spherical**	**3.67**
	2	Pentaspherical	3.11	Gaussian	3.56	Gaussian	3.56
Variogram model rank position	3	Gaussian	2.89	Tetraspherical	3.00	Tetraspherical	2.78
	4	Tetraspherical	2.89	Exponential	2.56	Pentaspherical	2.67
	5	Exponential	2.33	Pentaspherical	2.33	Exponential	2.33

Table 8. Results of variography parameter values from geostatistical analysis.

Date	Variogram model	Range (km)	Nugget	Semi variance/sill
24 March 2008	Spherical	38.974	231.19	931.97
	Gaussian	38.974	367.72	993.35
3 February 2009	Spherical	14.714	205.05	652.8
	Gaussian	12.592	272.31	654.89
3 March 2009	Spherical	11.949	127.07	861.10
	Gaussian	10.245	252.02	862.73
18 September 2011	Spherical	18.956	385.66	658.18
	Gaussian	16.772	429.15	662.97
13 December 2011	Spherical	7.144	585.02	1069.82
	Gaussian	6.085	660.12	1070.51
7 March 2012	Spherical	13.535	272.92	3776.92
	Gaussian	11.162	719.44	3768.34
18 April 2012	Spherical	16.057	44.624	712.594
	Gaussian	14.160	150.79	722.88
2 May 2012	Spherical	9.381	696.38	1078.51
	Gaussian	8.350	751.74	1084.02
21 August 2012	Spherical	14.557	38.027	917.687
	Gaussian	13.317	182.77	941.37

the network used [15]. The higher the density of the rain gauge network is, the more accurate the spatial rainfall map produced will be.

The second reason is the storm event selected for analysis. In this study, the rainfall data were extracted based on recent flood events on a cumulative, daily basis covering the years 2008 until 2012. This is to ensure that the data extracted for analysis are data with a good response from a hydrological perspective. However, Othman et al. [4] selected storm events from the same year. According to the findings of Kamel et al. [16], the estimated variance and interpolation system in geostatistical analysis are sensitive to season and region. As aforementioned, the climate feature in this study area is influenced by the monsoon seasons.

The different performances between the Spherical and Gaussian can be explained by how the model presents the correlation of the semivariogram within the range value. The Gaussian has high response to the nugget effect, but the Spherical is not. The nugget effect contributes to high error, while the model is used in spatial interpolation process at unmeasured points. This factor caused the Gaussian to perform less than Spherical model. Thus, it is important to assess the

variography structure of these two models that is represented by variography parameters to justify the result.

The variography parameters values of the data sets for Spherical and Gaussian models are tabulated in Table 8. The nuggets recorded by Gaussian are higher than Spherical, but the sill values are not much different. The range values are higher for Spherical model. This result shows that the Spherical model has better variography structure than Gaussian, because the Spherical has lower nugget and higher range value. This means that the Spherical has more spatial correlation at longer distance and less measurement error. By using the variography structure, semivariogram of Spherical and Gaussian for two selected data sets is plotted in Figure 4. It shows that the Spherical model fits the data set slightly better than Gaussian model.

A good spatial rainfall map can be produced by a good semivariogram structure. Apparently, based on the semivariogram structure, the Spherical model is able to produce good spatial rainfall map. Spatial rainfall maps for the same selected events are generated and shown in Figure 5. It is found that the spatial rainfall map produced using the Spherical model is smoother than with the Gaussian model, especially
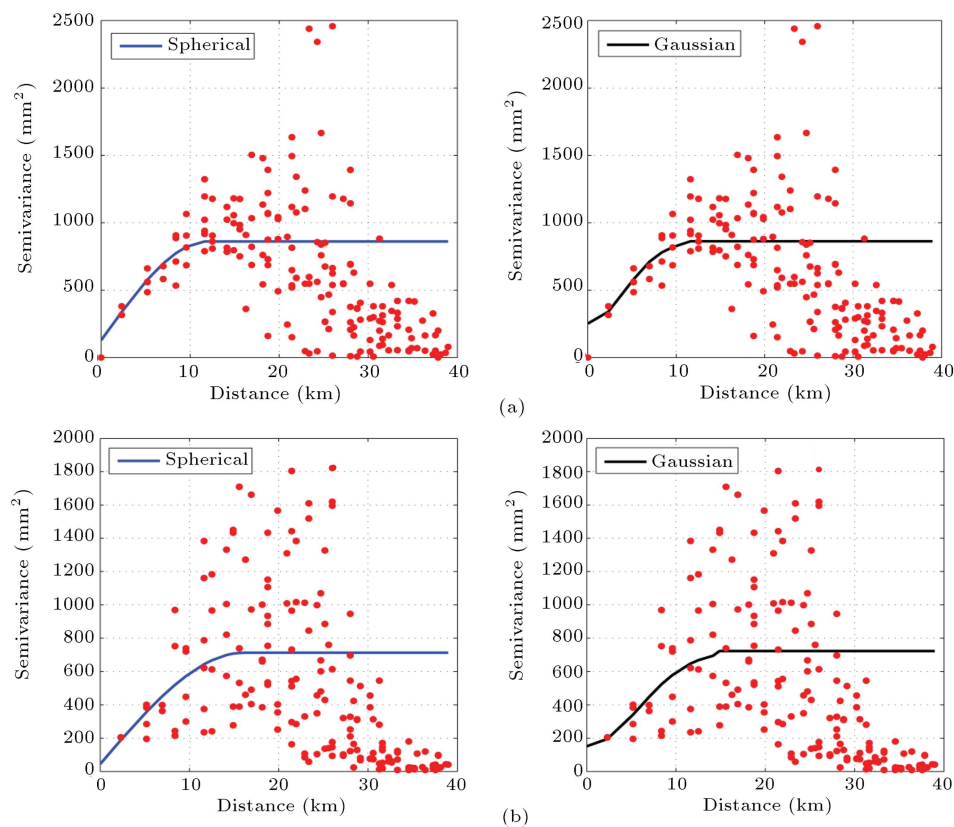
Figure 4. Semivariogram generated by Spherical and Gaussian models: (a) Rainfall data on 3 March 2009, and (b) rainfall data on 18 April 2012.
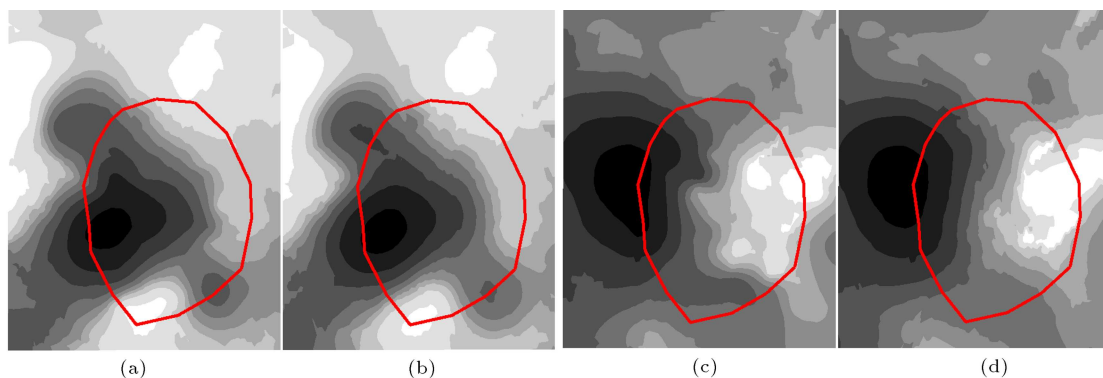
Figure 5. Spatial rainfall map generated by (a) spherical model for rainfall data on 3 March 2009, (b) Gaussian model for rainfall data on 3 March 2009, (c) spherical model for rainfall data on 18 April 2012, and (d) Gaussian model for rainfall data on 18 April 2012.

in the area marked with red boundary. The smooth line in Spherical is produced from the accurate spatial interpolation. The smooth line justifies that the model has better spatial interpolation as explained in the previous paragraph.

5. Conclusion

This study was conducted to present the application of a Multi-Criteria Decision-Making method, namely the AHP, to evaluate the performance of a variogram model in producing the best spatial rainfall map. Geostatistical analysis was applied to study the spatial structure of cumulative daily rainfall data in the upper Klang River basin. Four criteria were determined as performance indicators of a spatial rainfall map: root-mean-square error, average standard error, mean standardized error, and root-mean-square standardized error, to assess the five variogram model candidates, i.e. Spherical, Tetraspherical, Pentaspherical, Exponential, and Gaussian.

The AHP results indicate that the Spherical and

Gaussian models have good performances to produce a spatial rainfall map that fulfills the performance indicators. Two out of three sets of priority weight used in this study had both these models ranked at the top place with scores of at least 3.56. Moreover, the Spherical model was found to be slightly better than the Gaussian model because all priority weight sets had similar results that ranked the Spherical model in the first place. The smoothness of the contour map represents the spatial rainfall distribution generated by the Spherical model, justifying its performance.

Acknowledgment

References

1. Meng, Q., Liu, Z. and Borders, B.E. "Assessment of regression Kriging for spatial interpolation - comparisons of seven GIS interpolation methods", *Cartogr. Geogr. Inform. Sci.*, **40**(1), pp. 28-39 (2013).

2. Bargaoui, Z.K. and Chebbi, A. "Comparison of two kriging interpolation methods applied to spatio temporal rainfall", *J. Hydrol.*, **365**(1-2), pp. 56-73 (2009).

3. Ly, S., Charles, C. and Degré, A. "Geostatistical interpolation of daily rainfall at catchment scale: the use of several variogram models in the Ourthe and Ambleve catchments, Belgium", *Hydrol. Earth Syst. Sc.*, **15**(7), pp. 2259-2274 (2011).

4. Othman, F., Akbari, A. and Samah, A.A. "Spatial rainfall analysis for an urbanized tropical river basin", *Int. J. Phys. Sci.*, **6**(20), pp. 4861-4868 (2011).

5. Johnson, K., Hoef, J.M.V., Krivoruchko, K. and Lucas, N., *ArcGIS 9 : Using ArcGIS Geostatistical Analyst*, ESRI, United State of America (2003).

6. Journel, A.G. and Huijbregts, C.J., *Mining Geostatistics*, Academic Press, New York (1978).

7. Goovaerts, P., *Geostatistics for Natural Resources Evaluation*, Oxford University Press, New York (1997).

8. Chiles, J.P. and Delfiner, P., *Geostatistics: Modelling Spatial Uncertainty*, John Wiley & Sons, New York (1999).

9. Saaty, T.L. "How to make decision: The analytic hierarchy process", *Cent. Europ. J. Oper. Re.*, **48**, pp. 9-26 (1990).

10. Saaty, T.L., *The Analytic Hierarchy Process*, McGraw-Hill International, New York, NY, U.S.A. (1980).

11. Saaty, T.L. and Vargas, L.G., *Prediction, Projection, and Forecasting: Applications of the Analytic Hierarchy Process in Economics, Finance, Politics, Games, and Sports*, Kluwer Academic Publishers, Boston (1991).

12. Coulter, E.D., Coakley, J. and Sessions, J. "The analytic hierarchy process: A tutorial for use in prioritizing forest road investments to minimize environmental effects", *Int. J. For. Eng.*, **17**(2), pp. 51-69 (2006).

13. Ishizaka, A. and Lusti, M. "How to derive priorities in AHP: a comparative study", *Cent. Europ. J. Oper. Re.*, **14**(4), pp. 387-400 (2006).

14. Saaty, T.L. "Decision making with the analytic hierarchy process", *Int. J. Serv. Sci.*, **1**(1), pp. 83-98 (2008).

15. Goovaerts, P. "Performance comparison of geostatistical algorithms for incorporating elevation into the mapping of precipitation", *Proc. 4th Int. Conf. on Geo Comput.*, Mary Washington College Fredericksburg, Virginia, USA (1999).

16. Kamel, H.F.E., Slimani, M. and Cudennec, C. "A comparison of three geostatistical procedures for rainfall network optimization", *Int. Renew. Energy Congr.*, Sousse, Tunisia, pp. 260-267 (2010).

Biographies

Mohd Zaharifudin Muhamad Ali was born in 1978 in Perak, Malaysia. He obtained his BEng (Hons) degree in Civil and Structural Engineering from Universiti Kebangsaan Malaysia in 2000 and his MEng degree in Hydrology and Water Resources Engineering from Universiti Teknologi Malaysia in 2009. He is a Professional Engineer registered with Board of Engineer Malaysia since 2010. Currently, he is working with Drainage and Irrigation Department of Malaysia as a Senior Engineer and working on rain gauge network optimization for his PhD degree program in University Malaya of Malaysia.

Faridah Othman obtained her BS degree in Civil Engineering from the University of Missouri Kansas City, USA, MS degree in the field of Hydraulic Engineering, and PhD degree in Computational Hydraulics from the University of Newcastle Upon Tyne, UK. Her field of specialization is water resources, environmental hydraulics, river and water quality modeling. She has been involved in modeling exercise for more than fifteen years and is actively involved in the study of the water resources management, river and surface water quality studies, quality and performance of the water supply, flood mitigation works, and GIS application in water engineering. She has also been consulted by the rele-

vant agencies and private companies regarding projects related to water resources, hydraulics and hydrology, river water quality, and environmental issues.

Dr Faridah has supervised a number of PhDs, Masters as well as undergraduate students in their final year theses. She has authored and co-authored several technical papers in journals, conference proceedings, and other technical reports. She has also served as a referee for several papers published in international and local journals and proceedings.