Kriging Interpolation: Alternative Approach for Ungaged Daily Rainfall Estimation

S.Tantanee
Civil Engineering Department, Engineering Faculty, Naresuan University
sarintipt@nu.ac.th, satnatnee@yahoo.com

***Abstract*-- According to the problem of lacking daily rainfall data for the studied area, the method of rainfall estimation for the ungaged site is needed. Although, there are several processes to interpolate point rainfall from the available measured records of surrounding raingages, these methods provide the average values without the statistical consideration. Kriging interpolation which is the geostatistic process then is applied to estimate the daily rainfall data. The study was undertaken to investigate the Kriging interpolation application for daily rainfall estimation by using 2005 daily rainfall records of 10 stations distributed over eastern part of Bangkok. Comparing the obtained estimated values among Kriging, the simple arithmetic mean, and the spatial interpolation of Inverse Distance Weighing (IDW), it apparently shows that Kriging interpolation provides lowest level of Root Mean Square Error (RSME) and good preservation on the statistical characteristics, as well.**

***Index Terms*-- Interpolation, Rainfall, Spatial Interpolation, Kriging**

I. INTRODUCTION

Hydrologists and engineers always have the problem of lacking daily rainfall data in the studied station. The daily rainfall data is the essential input for several hydrological models. Based on the assumption that measured rainfall at a station is point rainfall, there are some traditional approaches to estimate the missing daily rainfall data from available records of surrounding raingages. Rainfall estimates at ungaged station can be obtained from linear or non-linear combinations of neighboring gauged measured values with the appropriate weighing factor (Lanza et. al., 2001). Among these approaches, the simplest methods are arithmetic averaging process (Wanielista et. al., 1997) and spatial based interpolation of Inverse Distance Weighing (IDW) (Dingman, 2002). However, these processes are the roughly estimation without the considering of statistical point of view. Rainfall is realized as a high degree of variability both space and time, thus, the stochastic interpolation theory has been introduced to estimate the daily rainfall (Lanza et. al., 2001). The objective of this study is to study the alternative process on determining the missing daily rainfall data for a particular site with the geostatistics method of Kriging interpolation. The studied process is undertaken by using the wet season daily rainfall data of the eastern Bangkok area obtained from Thai Meteorological Department (TMD).

II. STUDIED DATA

The 2005 daily rainfall records of 10 stations distributed over the eastern area of Bangkok are selected to study because of the completion of the records. These stations are RS07, RS15, RS16, RS20, RS25, RS26, RS27, RS29, RS32 and RS33. According that more than 80% of rainfall occurs during wet period, only wet season daily rainfall of 6 months (May- October) are analyzed in this study. The locations of the studied raingages are shown in Figure 1.

III. STUDIED MODEL

A. Kriging interpolation

There are several types of interpolation procedures such as: point/aerial interpolation, global/local interpolation, exact/approximation interpolation and stochastic/deterministic interpolation. Kriging is one of the spatial interpolations which is the procedure to estimate unknown variables at a particular point (unmeasured point) by using the measured values from the other point. (Beers and Kleijnen, 2004) Kriging is the geostatistic model, depended on models of spatial autocorrelation, which is formulated as fitting model of covariance or semivariogram. The Kriging of *ordinary Kriging* was used in this study. The concepts of ordinary function are: (1) the random process model consisting of constant mean (μ) and error term ($\delta(x)$) and, (2) the prediction for a point (x_{n+1}) is a weighted linear function of all the observed output data. The value of unmeasured point, $\hat{Y}(x_{n+1})$ can be estimated by a weighted linear combination of all measured point, x_i.

$$\hat{Y}(x_{n+1}) = \sum_{i=1}^{n} \lambda_i Y(x_i) = \lambda' \cdot \mathrm{Y} \qquad (1)$$

Where $\sum_{i=1}^{n} \lambda_1 = 1$, $\lambda = (\lambda_1, \ldots, \lambda_n)'$ and $\mathrm{Y} = (Y(x_1), \ldots, Y(x_n))'$ are random variables. The weights λ_i are chosen so as to minimize the prediction variance, which provide a Best Linear Unbiased Estimator (BLUE) of the output value for a given input. In ordinary Kriging assumption, the output covariances are considered as the function of the distance between the input data locations. The process of a covariance is second order stationary of which

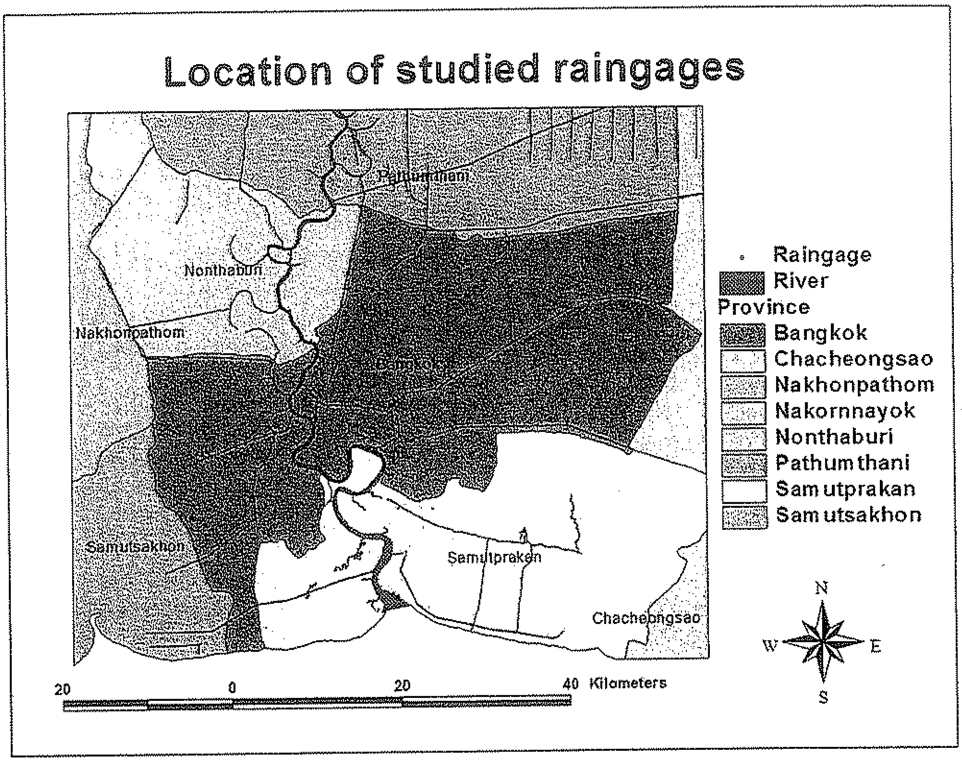


Fig. 1. The location of studied raingage stations

the mean (μ_Y^2) and variances (σ_Y^2) are constant and output covariances depends on distance (h) between the inputs. There are several types of covariance function used such as: exponential covariance functions with parameter θ, Guassian function, and linear covariance function. (Beers and Kleijnen, 2003) This covariance function, sometimes, is called variograms, which shows the characteristic of stationary stochastic process.

B. The traditional interpolation procedures for estimating missing daily rainfall

Based on the assumption that measured rainfall at station point rainfall is point rainfall, the interpolation methods and aggregation techniques can be used to estimate the aerial rainfall and the missing daily rainfall data from available records in watershed. Such techniques may be the simple process of arithmetic averaging (Wanielista et. al., 1997) or the spatial based interpolation such as: Inverse Distance Weighing (Dingman, 2002); isohyets method (Linsley et. al., 1949); Thiessen polygons ((Linsley et. al., 1949); and the splines process (Lanza et. al., 2001). According to the simplicity and widely being used, the most common interpolation techniques of arithmetic averaging and the Inverse Distance Weighing (IDW) method are selected to compare with the Kriging process.

The simple arithmetic averaging is the common used procedure for estimating daily totals relied on the data from three adjacent stations. The locations of the selected adjacent stations are closed to and approximately spaced around the considered missing data site. Equation (2) and (3) is used to estimated the missing daily rainfall in case that the average annual rainfall difference between the considered site and each of three adjacent stations are less than 10% and greater than 10%, respectively.

$$P_X = \frac{1}{3}(P_A + P_B + P_C) \tag{2}$$

$$P_X = \frac{1}{3}\left(\frac{N_X}{N_A}P_A + \frac{N_X}{N_B}P_B + \frac{N_X}{N_C}P_C\right) \tag{3}$$

Where

P_X = estimated daily rainfall at missing data site (depth)

P_A, P_B, P_C = daily rainfall at the adjacent stations A, B and C (depth)

N_X = average annual rainfall at missing data site (depth)

N_A, N_B, N_C = average annual rainfall at the adjacent stations A, B and C (depth)

The spatial based procedure of Inverse Distance Weighing (IDW) is one of the distance weight interpolation methods. IDW can be used to generate spatially interpolated rainfall using observed surrounding stations. The considered missing data site would be approximately at the center of the locations of these measured stations. The weight is in function of the distance between the adjacent stations and the considered site. The closer the station is , the higher weighing factors is obtained. The weighing functions is given by:

$$W_i = \frac{1}{d_i^2} \quad (4)$$

The estimated daily rainfall can be calculated from:

$$P_X = \frac{\sum_{i=1}^{n} W_i P_i}{\sum_{i=1}^{n} W_i} \quad (5)$$

Where

W_i = weighing function for station i

d_i = distance between the missing data site and the adjacent station i

P_X = estimated daily rainfall at missing data site (depth)

P_i = daily rainfall at the station i (depth)

IV. Methodology

The 2005 daily rainfall data of 10 distributed stations over eastern area of Bangkok are used in this study. The RS25 raingage station, which locates approximately in the center of the area, is selected to be the missing daily rainfall site. The interpolation procedure of arithmetic averaging, IDW and Kriging are applied to the measured rainfall in order to generate the estimated missing daily rainfall at RS25. The selected involving measured stations for each procedure are determined depending on the condition of the procedure. For the arithmetic averaging procedure, the three adjacent stations of RS20, RS26 and RS29 are selected to involve in the missing values computation because they are closer to the considered site than the remaining stations. As the annual rainfall differences between these stations are more than 10%, the arithmetic mean with the normal ratio of equation (3) is used. Considering the locations and the distance between the missing data site and the surrounding stations, the four measured stations of RS20, RS26, RS29 and RS33 are selected to be analyzed by the IDW procedure.

For Kriging procedure, rainfall records of all remaining 9 stations are used to estimate the missing daily rainfall. As the Kriging procedure involved with the huge amount of spatial data, there are several sophisticated GIS software products that provide this functions. In this study, the extension of Kriging interpolation in Arcview has been applied in the analysis of Kriging process.

The comparing of statistical characteristics preservation among three models have been done by considering statistical parameters of *mean, standard error, standard deviation, sample variance, skewness* and *maximum estimated values*. The comparing of estimation error among these models are considered by Root Mean Square Error (RSME), which is calculated by:

$$RSME = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left[\hat{X}(z_i) - X(z_i)\right]^2} \quad (6)$$

Where $\hat{X}(z_i)$ is the estimated value at time i and $X(z_i)$ is the actual values at time i.

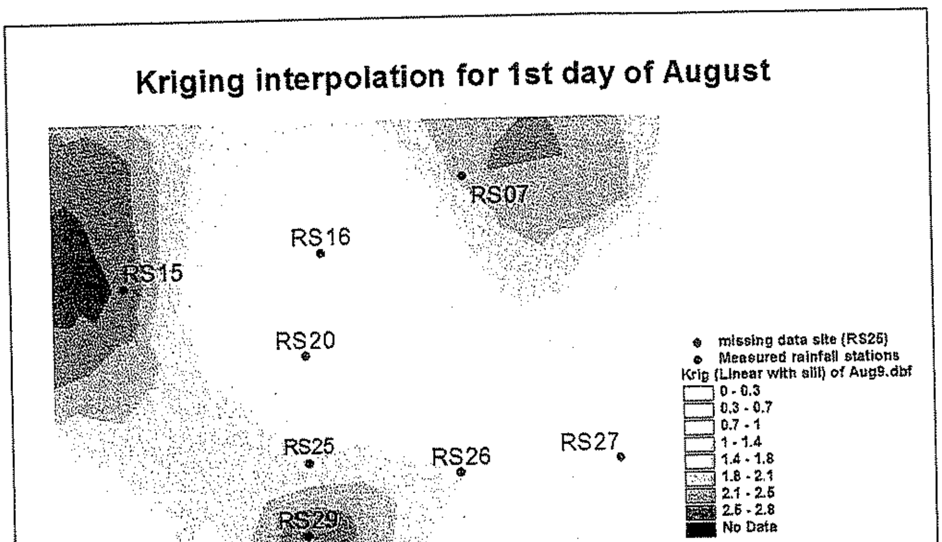
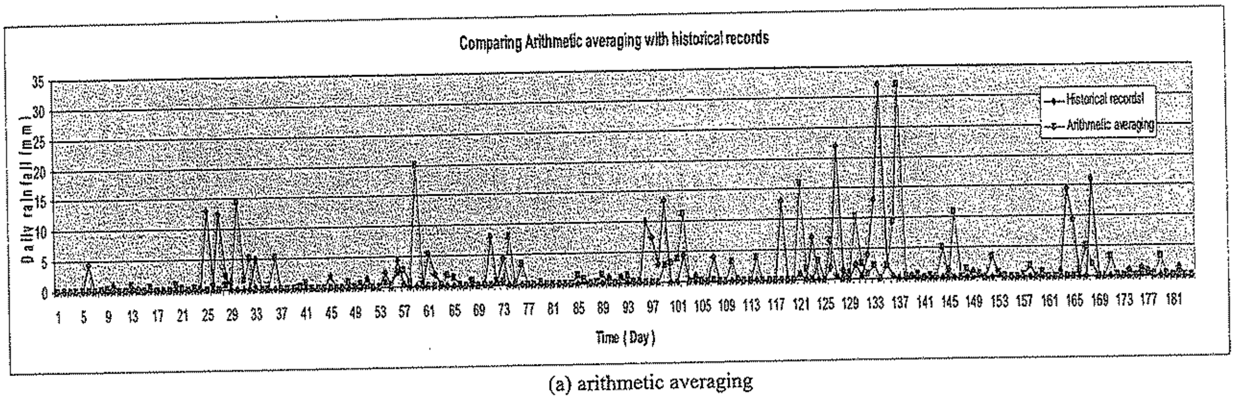
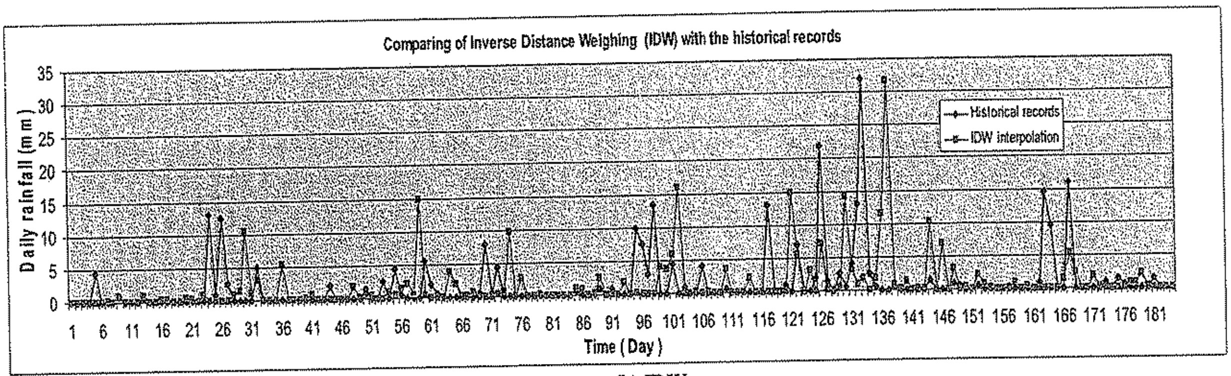


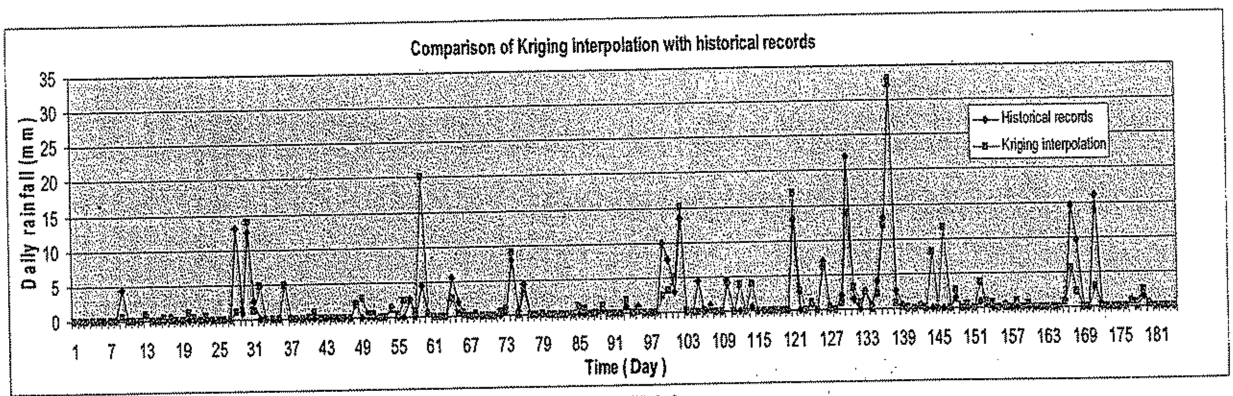
Fig. 2. Example of Kriging interpolation at 1st day of August



(a) arithmetic averaging



(b) IDW



(c) Kriging

Fig. 3. Comparison of daily rainfall historical records at RS25 with the estimated values from (a) arithmetic mean, (b) IDW and (c) Kriging

TABLE I
Comparison of statistical characteristics and Root Mean Square Error (RSME)
among models used

Statistics Parameter	Historical value	Estimated value from		
		Arithmetic mean	IDW	Kriging
Mean	1.459	1.339	1.389	1.465
Standard error	0.300	0.267	0.267	0.286
Standard deviation	4.064	3.616	3.622	3.887
Sample variance	16.519	13.078	13.122	15.107
Skewness	4.246	5.157	4.839	4.790
Maximum value	32.000	31.892	31.757	33.050
RSME	-	5.537	5.345	2.566

V. Results And Discussions

Figure 2 shows the example of Kriging interpolation at 1st day of August. The Figure also illustrates the variation of daily rainfall over the studied area in that day. The generating results of estimated daily rainfall from three studied interpolation procedures comparing with the historical records at RS25 are demonstrated in Figure 3. It can be seen that Kriging interpolation procedure provides the most compatible of rainfall time variation pattern with the historical rainfall records. The comparing of statistical characteristics and the error estimation between the estimated values and the historical records are summarized in Table 1. It is obviously shown that among three procedures Kriging provides the best of statistical characteristics preservation, whereas, the other two processes provide better results of the maximum estimated value. Comparing the obtained error of each procedures, Kriging also provides less error in estimation and the obtained error from arithmetic mean and IDW is around the same level.

VI. Conclusion

With the case study of Bangkok daily rainfall, it is summarized that Kriging interpolation can be utilized as an alternative method for the missing daily rainfall estimation. However, this study is the primary investigation using the ordinary Kriging interpolation for estimating ungaged site daily rainfall and comparing with only two common interpolation processes of arithmetic mean and IDW. In addition, only one year of the daily rainfall records have been analyzed in the study. Thus, it should be studied further on comparing Kriging process with the other spatial interpolation methods and also expand the studied area with the longer records of rainfall.

References

[1] Dingman, S.L. (2002). Physical Hydrology (2nd ed.). Prentice Hall.

[2] Lanza, L.G., Ramirez, J.A. and Todini, E. (2001). Stochastic rainfall interpolation and downscaling. J. Hydrol. Earth System Sci., 5(2), 139-143.

[3] Linsley, R.K., Kholer, M.A. and Paulhus, J.L.H., (1949). Applied Hydrology, McGraw Hill, New York.

[4] Van Beers, W.C.M. and. Kleijnen, J.P.C. (2003). Kriging for interpolation in random simulation. J. of the Oper. Res. Society, 54: 255-262.

[5] Van Beers, W.C.M. and Kleijnen, J.P.C. (2004). Kriging interpolation in simulation: A survey. Proceedings of the 2004 Winter Simulation Conference eds. R .G. Ingalls, M. D. Rossetti, J. S. Smith, and B. A. Peters.

[6] Wanielista, M., Kersten, R. and Eaglin, R. (1997) Hydrology: Water quantity and quality control. John Wiley & Sons, Inc. New York.