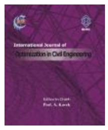
INTERNATIONAL JOURNAL OF OPTIMIZATION IN CIVIL ENGINEERING
Int. J. Optim. Civil Eng., 2016; 6(4):581-594



OPTIMAL SELECTION OF NUMBER OF RAINFALL GAUGING STATIONS BY KRIGING AND GENETIC ALGORITHM METHODS

A. Adib*, † and M. Moslemzadeh
Civil Engineering Department, Engineering Faculty, Shahid Chamran University, Ahvaz, Iran

ABSTRACT

In this study, optimum combinations of available rainfall gauging stations are selected by a model which is consist of geo statistics model as an estimator and an optimized model. At the first, watershed is approximated to several regular geometric shapes. Then kriging calculates the variance of the estimation error of different combinations from available rainfall gauging stations using inside and outside stations of watershed. In each combination, n is number of considered stations and N is number of available stations (N>n). At the end, the best combination is selected by genetic algorithm (the error variance of this combination is minimum). For optimal set with one sample point (station) estimator model and optimize model select station that locates near to center of watershed. While for two stations case, these models select two stations that locate in boundaries face to face. Also for combination n stations of N stations, selected stations have good and proportional distribution in watershed. These results show correctness of research methodology.

In this study, effects of variations of parameters of theoretical variogram and number of blocks in block estimation of kriging method are evaluated too. The variance of the estimation error from block estimation with 8*8 blocks has showed the acceptable results.

This research shows a linear relation between variations of error variance and scale of variogram. Optimum combination does not vary with variations of scale of variogram but it varies with variations of range of variogram. Increasing of nugget effect of variogram would raise the variance but does not vary optimum combinations.

Keywords: kriging; geo statistics; genetic algorithm; combination of rainfall gauging stations.

Received: 20 February 2016; Accepted: 4 May 2016

*Corresponding author: Civil Engineering Department, Engineering Faculty, Shahid Chamran University, Ahvaz, Iran
†E-mail address: arashadib@yahoo.com (A. Adib)

1. INTRODUCTION

Precipitation is the most important part of hydrological cycle. Therefore accurate measurement of depth and intensity of rainfall is essential for hydrological studies. Rainfall gauging stations measure depth of rainfall in a point (location of station). The depth of rainfall in different stations can be generalized to entire of watershed by some of the famous method such as thiessen from geometric or kriging from geo statistics methods. In kriging the variogram function determines spatial similarity between different points and this function is base of kriging estimator model. Appropriate distribution of weights between different points of watershed is calculated by some mathematical operations with the variogram basis. Thus accurate selection of this function is first step for using of kriging estimator model.

Cheng *et al*. [1] stated that determination of variogrom function is the first important stage for estimation of depth of rainfall by kriging estimator model. They estimated depth of rainfall for several storm events and compared their results to measured data. They concluded that using of a common variogram function (for entire of storm events) is wrong and suitable time step is less than one storm event time for use. In the other word variogram function should vary in several time sections and decreasing of time of this segmentation for variogram function will improve results of kriging estimator model. Goovaerts [2] studied about effects of roughness of watershed on depth of annual and monthly rainfalls. His case study was a watershed in south of Portugal with 36 rainfall gauging stations. The area of this watershed is 5000 Km^2. He constructed a DEM model and applied ordinary kriging, simple kriging and kriging with drift external for estimation of depth of rainfall. He considered two states for evaluation of topographic effects:

1. With considering variations of height of watershed
2. Without considering variations of height of watershed

He observed that considering variations of height of watershed improves estimation of depth of rainfall especially in mountain watersheds.

Tsintikidis *et al*. [3] studied about relation between increasing of number of rainfall gauging stations and improvement of estimation of depth of rainfall by kriging estimator model and decreasing the variance of the estimation error. They showed that because of the existence of variogram function role, estimation in a case of high number of stations may not improve estimation of depth of rainfall by kriging estimator model. This subject shows that estimation by kriging estimator model is a function of distance and position of stations. Of course in kriging estimator model the estimation will improve if several rainfall gauging stations are added to available rainfall gauging stations.

Saghafian & Bondarabadi [4] applied thin plate smoothing splines (TPSS), weighted moving average (WMA), ordinary kriging and cokriging models for estimation of depth of rainfall in mountain regions of south west of Iran. They prove that cokriging is the best method for estimation of depth of rainfall in these regions.

Barca *et al.* [5] considered Apolia watershed in the south of Italy. This watershed has 30 constant rainfall gauging stations and 15 moving rainfall gauging stations. Meteorologists wanted to develop 5 new rainfall gauging stations in this watershed. They utilized two estimator models (Average Shortest Path Distance Minimization (ASPDM) model and kriging model with an optimization model (SA model). They wanted to minimize the variance of the estimation error. They divided watershed to 10m*10m pixels by using GIS. They finally determined four optimum combinations of rainfall gauging stations for four seasons of year and showed moving stations and constant stations with tolerance 5Km.

Karamouz *et al*. [6] designed network for stated pollution measuring stations in the Karun River. They used kriging as simulator model and genetic algorithm as an optimization model. They determined variogram function by using location and value of pollution of 35 stations.

Other researchers applied different kriging methods for estimation depth of rainfall. For example Lebel *et al*. [7] utilized spherical variogram function. Barancourt *et al*. [8] and Bastin *et al*. [9] divided the variance of estimation error into two parameters (time scale parameter and time independent variance parameter). Petty [10] applied remote sensing technique and kriging estimator model for estimation of depth of rainfall. Zoubeida & Afef [11] applied kriging with external drift and ordinary kriging in case of two and three dimension varigrams. Dimensions were rainfall intensity, position of rainguage stations and duration time. They showed that in ordinary kriging there is no difference between 2D and 3D cases.

Pardo-Igúzquiza [12] applied thiessen, ordinary kriging, kriging with external drift and cokriging methods for estimation of depth of rainfall in south of Spain. External drift kriging showed the best results. Also Russo *et al*. [13] utilized inverse distance and kriging methods for estimation of depth of rainfall in the Rome city. Results of kriging were better than inverse distance method. Kottegoda, & Kassim [14] used simple kriging method for designing optimal rainfall gauging stations network. Zidek *et al*. [15] utilized Gaussian variogram function and found optimum monitoring network. Nunes *et al*. [16] optimized groundwater monitoring network and found the best combination of stations in several time. They considered error variance and cost of stations. Finally they reached one combination for whole time.

Wu *et al*. [17] applied GA method and kriging simulator model for optimization of pollution monitoring network. They minimize error variance. Jimenez *et al*. [18] optimized a monitoring network for lakes and reservoir dams by GA method. They determined location of new stations that must be added to network. Ruiz-C´ardenas *et al*. [19] applied Hybrid Genetic Algorithm (HGA), GA method and SA method for optimization of monitoring network. HGA is combination of GA method and stochastic search algorithms. Also in recent years Ahmadianfar *et al*. [20, 21] and Adib & Samandizadeh [22] applied GA methods for optimization of volume of released water from dam reservoir.

The purpose of this research is selection a number of rainfall gauging stations from entire of available stations so that this combination can give best estimation of depth of rainfall. This subject is a practical problem. Because of expensive equipment of rainfall gauging stations, shortage of skillful experts and high cost and hardness of maintain and repair of them, a number of rainfall gauging stations may be eliminated in developing countries. For determination of optimum combination, authors of this research apply kriging estimator model and genetic algorithm and create a mixed optimization model.

2. THE RESEARCH METHODOLOGY

Optimization methods want to minimize error estimation (the difference between estimated value and observed data). Kriging method utilizes principles of geo statistic science. For calculation of the variance of estimation error by kriging, this method uses variance, covariance and few statistic assumptions. Kriging method gets partial differential from error variance function relative to unknown weighted coefficient of different stations and equalize obtained equations to zero for minimization the error variance. For solution of non poised equation sets (there is always one additional equation because summation of weights would be one) kriging method utilizes Lagrange coefficient method (constrained optimization) and calculates weighted coefficient of different stations and these lead to the variance. In kriging the variance of estimation error could be representative as quality of estimation.

Variogram function of each storm event is unique. For simplification, variogram function must divide to two parts (time invariant and space invariant). Therefore variogram function $\gamma(t_j, h)$ converts to:

$$\gamma(t_j, h) = \alpha(t_j) g(h, \beta) \tag{1}$$

where:

$\alpha(t_j)$: Time scaling parameter (time dependent and space invariant)

β : Shape parameter (time invariant and space dependent)

$g(h, \beta)$: Scaled climatologically varigrom (time invariant and space dependent)

$g(h, \beta)$ can be exponential variogram, spherical variogram, Gaussian variogram or etc. Error variance is:

$$\sigma_E^2(t_j) = \alpha(t_j)\delta_E^2 \tag{2}$$

where:

δ_E^2 : Scaled estimation variance

The estimation variance can be expressed (Journel and Huijbergts 1978):

$$\delta_E^2 = \mu + \sum_{i=1}^{N} \lambda_i \bar{g}(h_{iA}) - \bar{g}(h_{AA}) \tag{3}$$

where:

μ : Lagrange parameter

λ_i : Weighted coefficient of station i

$\bar{g}(h_{iA})$: The average of variogram function that is dependent to distance between station *i* and other stations

$\bar{g}(h_{AA})$: The average of variogram when both extremes of the vector *h* describe independently the area A.

The scaled estimation variance depends on three factors: scaled climatologically variogram, the number of stations and their locations. This parameter is adopted for variance reduction technique in this paper.

Different combinations of stations produce different error variances. The best combination has the least error variance. Kriging method only determines error variance for each combination. For finding the best combination, an optimization model must be applied. The number of combinations (subsets) with n members is:

$$\binom{N}{n} = \frac{N!}{n!(N-n)!} \tag{4}$$

where:

N is the number of entire of members (available stations).

In this research an empirical variogram is calculated by using data from location of stations and value of variables at them. Then a theoretical model is fitted on empirical variogram. Kriging estimator model can calculate error variance of different combination by theoretical model. Kb2d program are applied by authors as kriging estimator model. The programming language of kb2d is Fortran. This program was prepared from geo statistic software library (g.s.lib) and can perform ordinary kriging and simple kriging. Authors linked kb2d to GA (genetic algorithm) toolbox of MATLAB. Kb2d can estimate variables at points that locate inside and outside the stations limit. While other soft wares (such as Arc GIS) that perform estimation by kriging method can not estimate variables outside the stations limit. Estimation region of those soft wares is shown in Fig. 1.

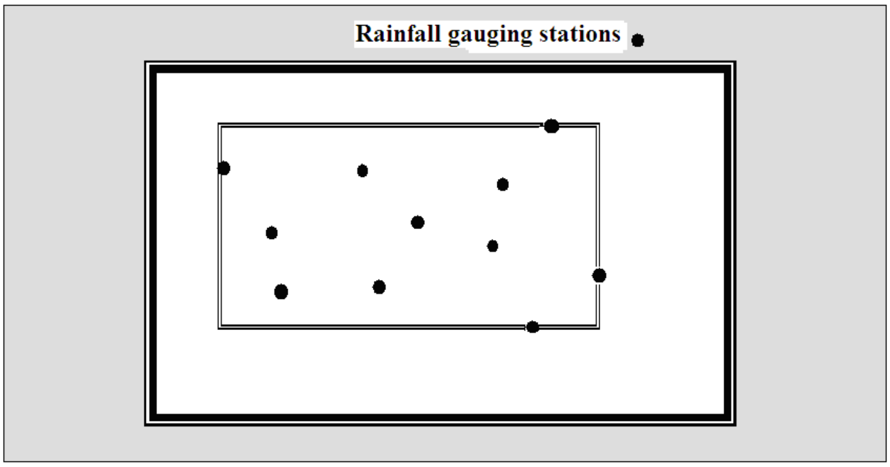


Figure 1. Estimation region of soft wares such as Arc GIS

For using of kb2d program, the region is considered to several rectangular and square. Kriging utilizes information of inside and outside of region for estimation. Therefore region and limitation boundaries which stations are effective for estimation of depth of rainfall should be distinguished. Stations that locate out of limit must be eliminated in optimization procedure. In this purpose a great penalty allocates to these stations and error variance can not be the minimum by considering them. Also each station must not be considered more than one time in optimization procedure. Therefore a great penalty allocates to this state too. For estimation, the effective radius (maximum of search radius) must be determined using variogram. The points that their distance from considered point is more than effective radius are not suitable for estimation. Effective radius is equal to effective variogram's range or is less than it (2/3 effective variogram's range). Using the radius decreases the number of used point for estimation but increases accuracy of estimation and the process will run faster. In this program if the kriging uses unreasonable data for estimation (less than minimum required data or more than its maximum which are determined by user), kb2d will assume an especial value. For elimination of this state, a great penalty allocates to objective function which is the variance of combination.

For decreasing of time solution, error variance from series of combinations are calculated (no total combinations) according to optimization model. Therefore an explorer method such as GA method must be applied for optimization. Link between kriging and GA methods is shown in Fig. 2.

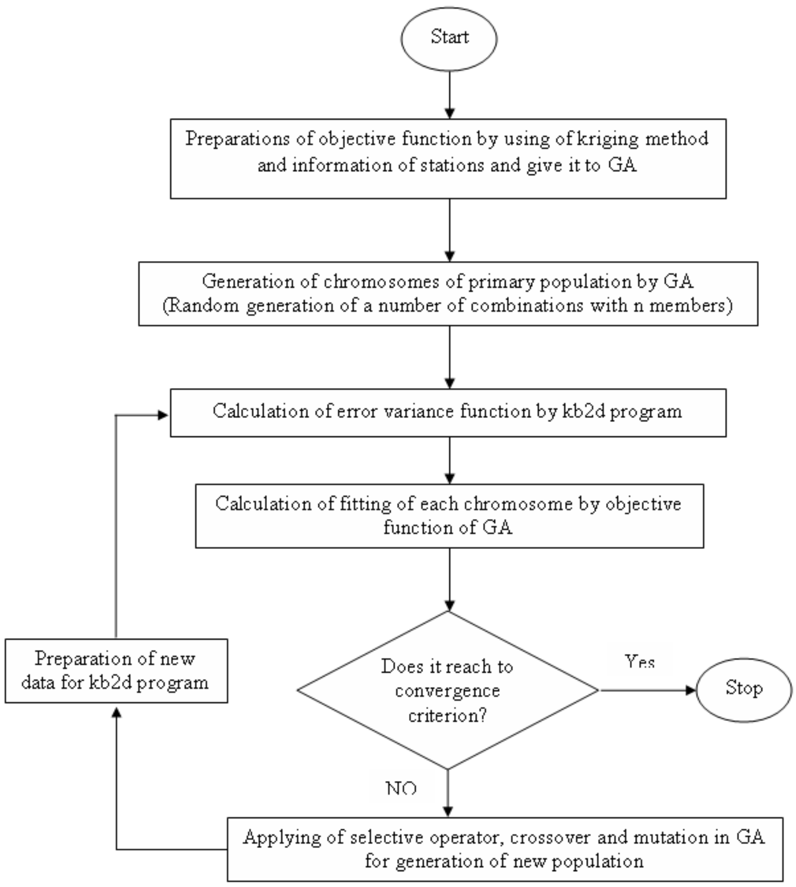


Figure 2. Link between kriging and GA methods

3. RESULTS AND DISCUSSION

In this project twenty five stations are considered which are shown in Fig. 3. These stations are located inside and outside the estimation region. The coordinates of four corners of estimation region are (4, 5), (25, 5), (4, 16) and (25, 16).

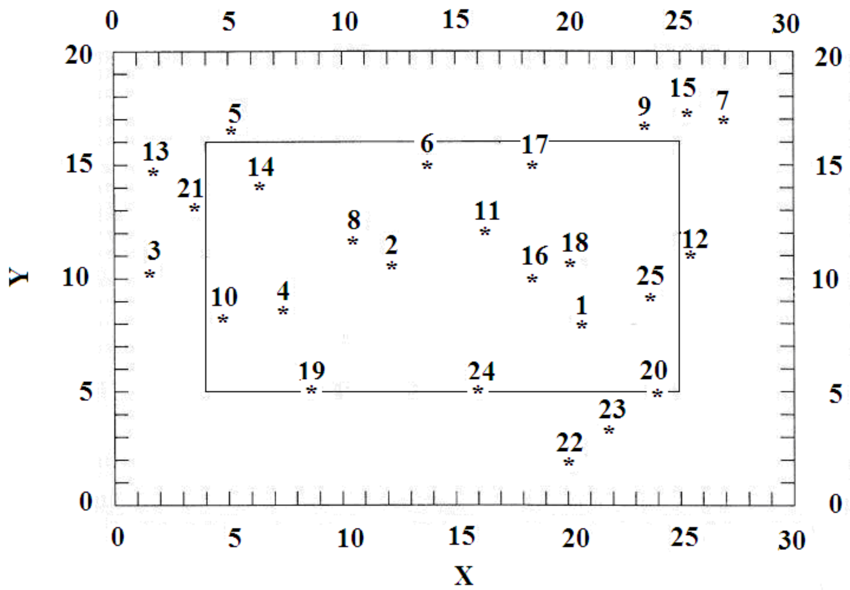


Figure 3. The shape of estimation region and location of 25 stations

The characteristics of variogram function are: anisotropic, spherical, effective range in the X axis is 10 units, effective range in the Y axis is 5 units and the value of sill is 100. At the first, characteristics of region, location of stations, characteristics of variogram function, type of kriging method (ordinary, simple and etc.), address of output file, maximum of search radius, minimum and maximum of number of stations for different combinations should be introduced to kriging estimator model by a parameter file. Maximum of research radius is equal to 10 units and minimum and maximum of number of stations for different combinations is considered 1 and 25 stations respectively.

Ordinary kriging utilizes static assumptions. The values of error variance of different combinations of stations are shown in Table 1 for some different ordinary block kriging cases. These values are concern to one state (location of stations is constant for each combination). The numbers of blocks in block kriging method in different cases are 1*1, 2*2, 4*4, 6*6, 8*8 and 10*10 blocks.

By increasing of number of blocks, error variance decreases but running time of estimation process would be increased. If number of blocks is 10*10, error variance will decrease slightly while time of estimation will increase very greatly. Thus the number of blocks is used 8*8 for kriging estimator model.

After using of kriging estimator model for several states from each combination , these results introduce to GA optimized model and GA determines optimum state for each combination.

Error variances of kriging estimator model and GA optimized model for different combinations (from one station to 10 stations) are shown in Table 2.

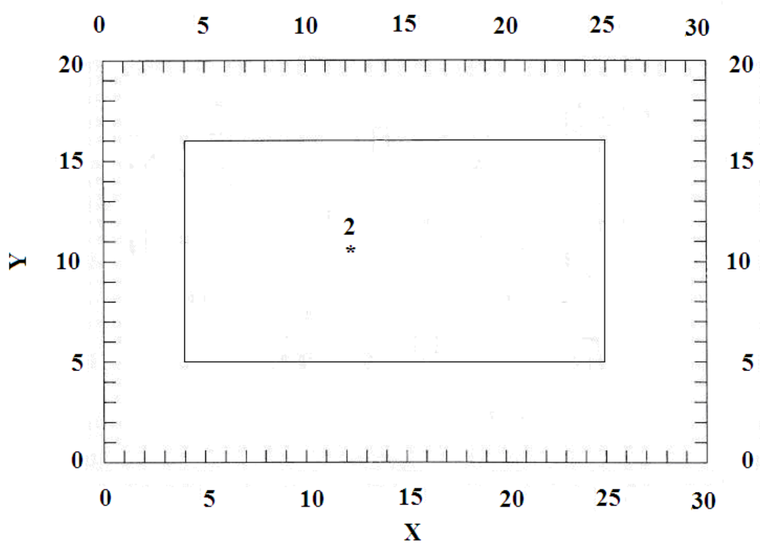
Table 1: The values of error variance of different combinations of stations

Number of stations	1*1 blocks	2*2 blocks	4*4 blocks	6*6 blocks	8*8 blocks	10*10 blocks
1	187.55	89.37	88.47	86.81	86.79	86.75
2	89.96	105.47	85.45	83.96	83.63	83.48
3	200	102.19	108.06	104.33	103.8	103.67
4	188.4128	92.72	88.25	88.14	87.69	87.65
5	200	120.71	105.67	104.18	103.79	103.65
6	190.3	112.01	99.68	96.21	93.83	93.72
7	199.09	125	110.56	108.16	107.73	107.58
8	201	95.79	84.6	84.37	84.01	83.93
9	130.31	122.27	106.9	105.39	105.01	104.86
10	200	108.66	97.37	96.07	95.64	95.61

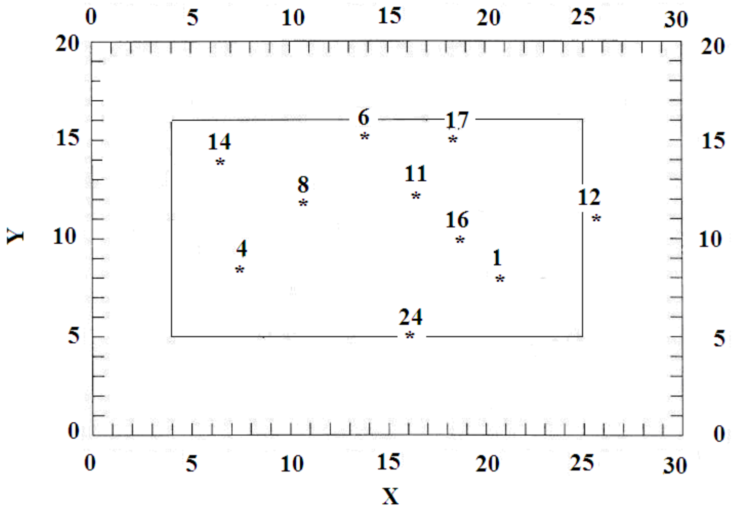
Table 2: The values of error variance of kriging estimator model and GA optimized model for different combinations of stations

Number of stations	1	2	3	4	5	6	7	8	9	10
error variance	83.63	34.89	20.53	15.04	9.57	7.75	6.47	5.66	4.67	4.2

For example, optimum spatial combination of stations (optimized by GA method) is illustrated in Fig. 4 for one station and ten stations.



(a)



(b)

Figure 4. (a) Optimum combination for one station (b) Optimum combination for ten stations

Characteristics of GA optimized model are:

Population size 10-50, Crossover Probability 0.5-0.8, Mutation Probability 0.001-0.08 and the number of generations 500.

For validation of GA method, the total of combinations in one station case (25 states) and two stations case (300 states) were considered by kriging estimator model. Minimum error variance of combinations in one station is 83.63 (equal to error variance resulted by GA) and optimum selection is station 2. Minimum error variance of combinations in two stations is 34.89 (equal to error variance of GA) and optimum selection is stations 8 and 18.

3.1 Effect of theoretical variogram parameters

1- Scale of variogram:

Fig. 5 shows two similar variogram functions with different scales. The sill of first variogram function is 100 units while sill of the second variogram function is 200 units.

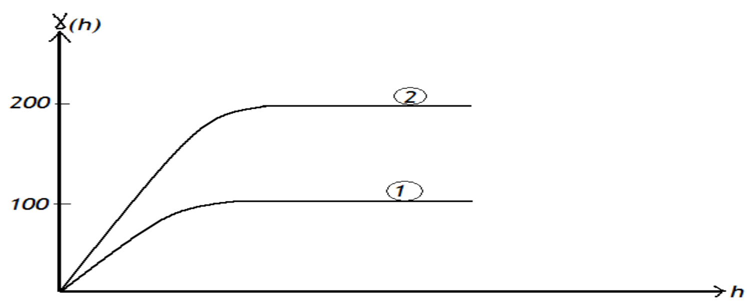


Figure 5. Two similar variogram functions with different scales

Results of kriging estimator model for combinations in one station type are shown in Fig. 6. Fig. 6 compares error variance calculated by those two variogram functions.

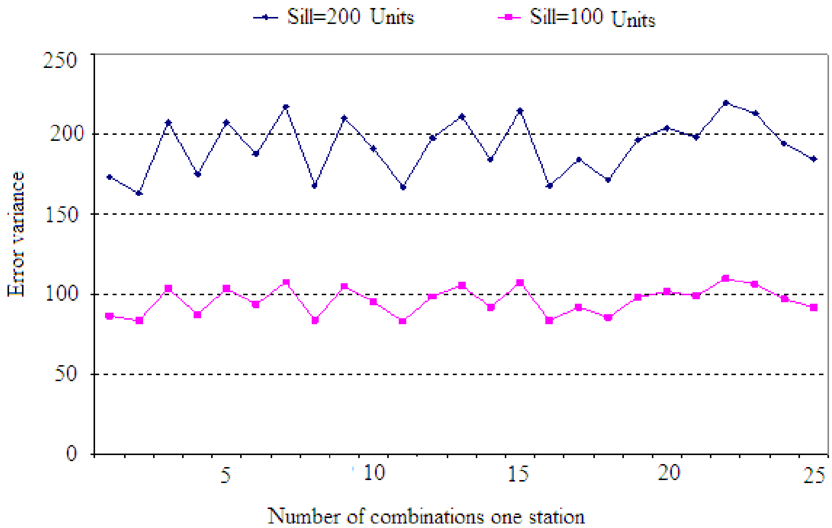


Figure 6. Effects of variations of scale on error variance

Fig. 6 shows those error variances will twice if scale of variogram becomes double but optimum combination does not vary. Thus scale of variogram is not effective for selection of optimum combination of stations.

2- Range of variogram:

Fig. 7 shows two similar variogram functions with different ranges in X direction. The range in X direction of first variogram function is 10 units while range in X direction of the second variogram function is 20 units.

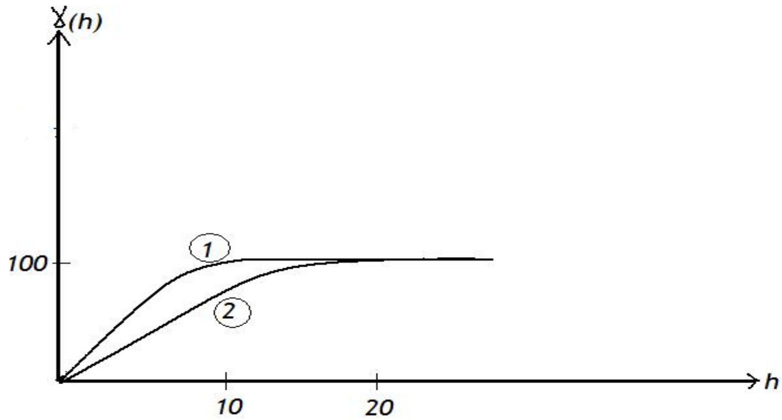


Figure 7. Two similar variogram functions with different ranges in X direction

Results of kriging estimator model for combinations one station are shown in Fig. 8. Fig. 8 compares error variance of two variogram functions.

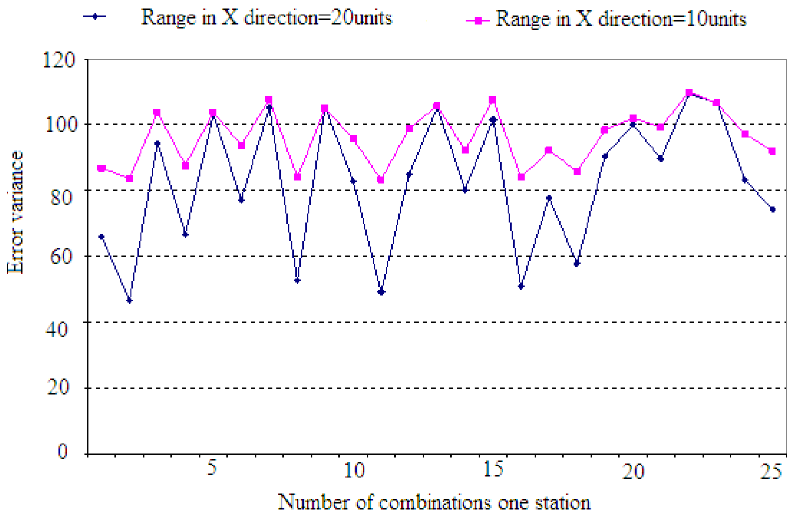


Figure 8. Effects of variations of range in X direction on error variance

Fig. 8 shows that variations of error variances are irregular if range of variogram varies. Also optimum combination varies with variations of range. For example optimum combination in two stations type are stations number 1 and 8 for variogram with range 20 units in X direction while optimum combination for variogram with range 10 units in X direction are stations number 8 and 18. Thus variation of range of variogram function is effective for selection of optimum combination of stations.

3- Nugget effect of variogram:

Fig. 9 shows two similar variogram functions with different nugget effect. The first variogram function has not any nugget effect while nugget effect of the second variogram function is half of sill (50 units).

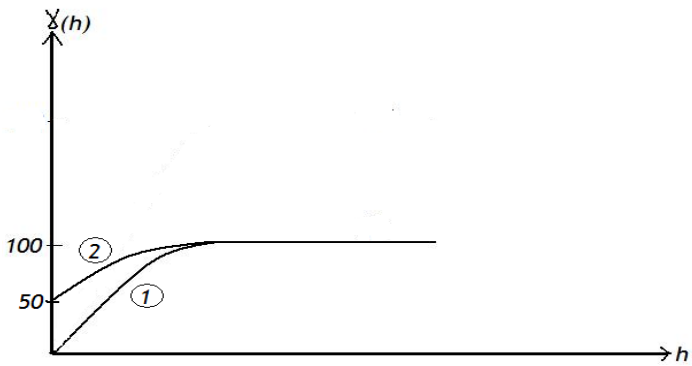


Figure 9. Two similar variogram functions with different nugget effect

Results of kriging estimator model for combinations in one station type are shown in Fig. 10. Fig. 10 compares error variance of two variogram functions.

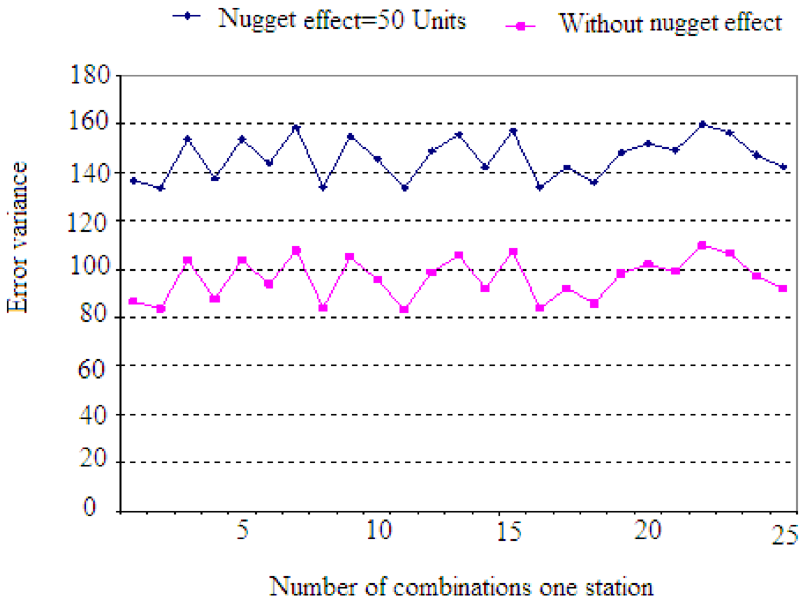


Figure 10. Effects of variations of nugget effect on error variance

Fig. 10 shows that error variance will vary if nugget effect of variogram varies but optimum combination does not vary. Thus any mistake in calculating nugget effect of variogram could be acceptable for selection of optimum combination of stations.

4. CONCLUSION

Kriging estimator model is a suitable method for calculation of error variance. This model must run for each combination separately. For decreasing of time solution and finding optimum combination, kriging estimator model must link to an optimized model. The most important stage is selection a suitable variogram function for determination of optimum combination. Wrong selection of variogram function produces a wrong combination of rainfall gauging stations. Sill and nugget effects of variogram function are not effective and range of variogram function is effective in selection of optimum combination.

By increasing the number of stations, error variance decreases. If the number of stations is relatively great, decreasing of error variance will become slightly. In optimum combination, selected stations have good and proportional distribution in region. For example for combination in one station type, selected station locates near the center of region.

Results of this research can be applied for decreasing of the number of rainfall gauging stations. By using of both kriging estimator model and an optimized model, the most effective stations are selected and other stations are eliminated.

REFERENCES

1. Cheng KS, Wei C, Cheng YB, Yeh HC. Effect of spatial variation characteristics on contouring of design storm depth, *Hydrol Process* 2003; **17**(9): 1755-69.
2. Goovaerts P. Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall, *J Hydrol* 2000; **228**(1-2): 113–29.
3. Tsintikidis D, Georgakakos KP, Sperfslage JA, Smith DE, Carpenter TM. Precipitation uncertainty and raingauge network design within Folsom Lake watershed, *J Hydrol Eng, ASCE* 2002; **7**(2): 175-84.
4. Saghafian B, Bondarabadi SR. Validity of regional rainfall spatial distribution methods in mountainous areas, *J Hydrol Eng, ASCE* 2008; **13**(7): 531-40.
5. Barca E, Passarella G, Uricchio V. Optimal extension of the rain gauge monitoring network of the apulian regional consortium for crop protection, *Environ Monit Assess* 2008; **145**(1-3): 375-86.
6. Karamouz M, Kerachian R, Akhbari M, Hafez B. Design of river water quality monitoring networks: a case study, *Environ Model Assess* 2009; **14**(6): 705-14.
7. Lebel T, Bastin G, Obled C, Creutin JD. On the accuracy of areal rainfall estimation: a case study, *Water Resour Res* 1987; **23**(11): 2123-34.
8. Barancourt C, Creutin JD, Rivoirard J. A method for delineating and estimating rainfall fields, *Water Resour Res* 1992; **28**(4): 1133-44.
9. Bastin G, Lorent B, Duque C, Gevers M. Optimal estimation of the average areal rainfall and optimal selection of rain gauge locations, *Water Resour Res* 1984; **20**(4): 463-70.
10. Petty GW. The status of satellite-based rainfall estimation over land, *Remote Sens Environ* 1995; **51**(1): 125-37.
11. Zoubeida KB, Afef Ch. Comparison of two kriging interpolation methods applied to spatiotemporal rainfall, *J Hydrol* 2009; **365**(1-2): 56-73
12. Pardo-Igúzquiza E. Comparison of geostatistical methods for estimating the areal average climatological rainfall mean using data on precipitation and topography, *Int J Climatol* 1998; **18**(9): 1031-47.
13. Russo F, Napolitano F, Gorgucci E. Rainfall monitoring systems over an urban area: the city of Rome, *Hydrol Process* 2005; **19**(5): 1007-19.
14. Kottegoda NT, Kassim AHK. Classification of storm profiles using crossing properties, *J Hydrol* 1991; **127**(1-4): 37-53.
15. Zidek JV, Sun W, Le ND. Designing and integrating composite networks for monitoring multivariate gaussian pollution fields, *J Royal Stat Soc: Series C (Applied Statistics)* 2000; **49**(1): 63-79.
16. Nunes LM, Cunha MC, Ribeiro L. Optimal space-time coverage and exploration costs in groundwater monitoring networks, *J Environ Monit Assess* 2004; **93**(1-3): 103-24.
17. Wu J, Zheng C, Chien CC. Cost-effective sampling network design for contaminant plume monitoring under general hydrogeological conditions, *J Contam Hydrol* 2005; **77**(1-2): 41-65.
18. Jimenez N, Mauricio Toro F, Ve´lez JI, Aguirre N. A methodology for the design of quasi-optimal monitoring networks for lakes and reservoirs, *J Hydro Inform* 2005; **7**(2): 105-16.

19. Ruiz-C´ardenas R, Ferreira MAR, Schmidt AM. Stochastic search algorithms for optimal design of monitoring networks, *Environmetrics* 2010; **21**(1): 102-12.
20. Ahmadianfar I, Adib A, Taghian M. A multi-objective evolutionary algorithm using decomposition (MOEA/D) and its application in multipurpose multi-reservoir operations, *Int J Optim Civil Eng* 2015; **5**(2): 167-87.
21. Ahmadianfar I, Adib A, Taghian M. Determination of optimal hedging rule using fuzzy set theory for multi-reservoir operation, *Int J Optim Civil Eng* 2016; **6**(1): 27-42.
22. Adib A, Samandizadeh MA. Comparison ability of GA and DP methods for optimization of released water from reservoir dam based on produced different scenarios by Markov chain method, *Int J Optim Civil Eng* 2016; **6**(1): 43-62.