Survey on Application of Geostatistical Methods for Estimation of Rainfall in Arid and Semiarid Regions in South West Of Iran

Ebrahim Hosseini Chegini
Mohammad Hossein Mahdian
Sima Rahimy Bandarabadi
Mohammad Mahdavi
Soil Conservation Watershed Management Research Center, P.O. Box: 13445-1136, Tehran, Iran

ABSTRACT

Estimation of rainfall data is necessary in many natural resources and agricultural studies. There are several methods to estimate rainfall along which interpolation are useful in this regards. In this research, three interpolation methods were used to estimate monthly and annual rainfall. These methods are: thin plate smoothing splines (TPSS), with and without co-variable (TPSS-CO), weighted moving average (WMA), and kriging (ordinary, cokriging and logkriging). The data of 167 climatic stations in south west of Iran with 22 years records were used in this study. Based on elevation, the studied area was divided into sub catchments. In the cases which a few stations exist in the catchments or regression coefficient between elevation and rainfall for two sites was the same, those two sub catchments was combined. Finally, 3 sub regions were used for analysis. Regression coefficient between rainfall and elevation was low in one of the sub catchments (R=0.12) and in two others was better (0.72 and 0.82). In general, for monthly data, this coefficient was less than 0.5. These results show that the regression coefficient between rainfall and elevation must be grater than 0.6 in order to improve interpolation accuracy using elevation as co variable. These results also show that TPSS with elevation as co-variable is the most precise method to estimate annual rainfall (MAE=13-38 mm). TPSS with power of 2 was the most precise method to estimate monthly (MAE=1.1 to 20.2 mm) and annual rainfall (MAE=13-38) mm.

INTRODUCTION

Knowledge about the rainfall is essential for environmental projects, and policy-makers in such activities as pattern and trend analysis of rainfall, and monitor network design. According to Boughton (1981), there are two main source of possible error in the estimation of aerial catchment rainfall, the measurement of point samples and extrapolation /interpolation from point to aerial rainfall. Rainfall variability is a major factor limiting the accuracy of rainfall runoff models. Errors may be additive because of the compounding effects of measurement deficiencies due to wind, and biased because of location of gauges at lower elevations leading to underestimating catchment rainfall in hilly or mountainous terrain (Boughton, 1981).

Aerial estimates of rainfall based on point measurements should be regarded, as an index of the true mean rainfall over a catchement and errors between 10 to 20 percent can be considered as normal. Where strong wind effects or mountainous catchements are being considered, errors up to 60 percent can be experienced (Hall and Barklay, 1975).

Several methods can be used in estimating annual and monthly rainfall data. The method such as Thiessen, isotheral method, multiple regression and correlation; and geostatistical methods can be used in this regard. Contrary to all methods, the structural analysis and estimation of errors can be determined with the geostatistical methods.

Abtew et al. (1994) used seven different interpolation methods for estimating monthly rainfall in south Florida. Based on their results, and standard errors as a evaluation criteria (SE), ordinary kriging (SE=1.13) universal kriging (SE=1.17) and bivariate interpolation (SE=1.32) was recognized as the most suitable methods in estimating monthly rainfall. Kriging and multivariate were compared using hourly rainfall of satellite by Borgan and Vizzaccaro (1997). In the conditions where the concentration of data was high, the kriging method gave the reliable results and in the regions with low concentration of data, the bivariate method was recognized better than kriging. Ahmed and De Marsily (1987) compared ordinary kriging, universal kriging and cokriging for estimating transmissivity of an aquifer and found that ordinary kriging gave the best estimate. Cook et al. (1993) studied the suitability of several interpolation methods to estimate steady state infiltration and found that inverse distance method to be the most accurate when compared with kriging, triangulation, and minimum surface curvature.

Studies by Hosseni et al. (1994) showed that the method of thin plate smoothing splines was the most accurate method for interpolation of soil salinity. Voltz and Goulard (1994) compared the performance of three interpolation methods (nearest neighbor, ordinary kriging, and cokriging) for soil moisture retention curves and found that ordinary kriging and cokriging were the most accurate.

Goovaerts (2000) evaluated the methods of inverse square distance, Thiessen polygon and ordinary kriging for estimating annual and monthly rainfall in the Algarve region, Portugal. Cross validation is used to compare the prediction performances of these interpolation methods. Large prediction errors are obtained for the inverse square distance and the Thiessen polygon methods that ignore both elevation and rainfall records at surrounding stations. Ordinary kriging and linear regression stresses the importance of accounting for spatially dependent rainfall observations in addition to the allocated elevation. Ordinary kriging yields more accurate predictions than linear regression when the correlation between rainfall and elevation in moderate (less than 0.75).

There is not a reliable distribution of meteorological stations in Iran, therefore, several regions can be found with any meteorological stations. In certain of them, the number of year with acceptable data is very short, so they must be neglected in the analysis. Therefore, the objective of this paper is to evaluate the applicability of different interpolation methods in estimating monthly and annual rainfall in south west of Iran.

MATERIALS AND METHODS

Description of the study area

In this study, arid and semiarid regions of south west of Iran were selected as study area (Fig. 1). This region includes the province of Yazd, Kerman, Sistan and Blochestan; and a part of the provinces of Fars, Esfahan, Khorasan and Hormozgan. The region has a longitude between 50° 0' 0" and 62º 49' 48" and latitude between 25º

16' 48'' and 43º 0' 0''. Elevation in this region has a high variation that varied between zero and 4376 m above sea level. The normal mean annual rainfall of 167 stations is 188 mm (from 41.3 to 486 mm).

INTERPOLATION METHODS

There are several geostatitiacl methods for estimating a variable in considered region. Table 1 shows the interpolation methods used in this study. All these method use the following equation:

$$* = \sum_{I=1}^{N} l_{I} \times Z_{I}$$

Where:

Z* = estimated value,

Zi = observed value,

N = number of observation,

I = index for the number of stations.

Table 1. Interpolation methods used in this study.

Abbreviation	Method
WMA	Weighted Moving Average
WMA-1	a=1
WMA-2	a=2
WMA-3	a=3
WMA-4	a=4
WMA-5	a=5
OK	Ordinary Kriging
OK-LO	with logaritme of data
OK-CO	elevation as covariable
TPSS	Thin plate smoothing splines
	Without covariable
TPSS-2	a=2
TPSS-3	a=3
TPSS-4	a=4
TPS-5	a=5
	With covariable

TPSS-CO2	a=2
TPSS-CO3	a=3
TPSS-CO4	a=4
TPS-CO5	a=5

EVALUATION CRITERIA

The performance of interpolation methods was evaluated by cross validation. In this procedure, an observed value is temporarily discard from the sample data set, and one estimated value at that location is determined using the other sample points. This results in a series of observed and estimated values that can be used to assess the validity of the interpolation method. In this study, estimated and observed values were compared using two criteria: the mean bias error (MBE) and the mean absolute error (MAE):

$$MBE = \frac{\sum_{i=1}^{n}(Psi - Poi)}{n}$$

$$MAE = \frac{\sum_{i=1}^{n}\left|Psi - Poi\right|}{n}$$

Where:

n = number of values,

Ps = simulated rainfall (mm),

Po = observed rainfall (mm),

i = index for the number of data.

RESULTS AND DISCUSSION

Variography analysis

Elevation has an important effect on rainfall, therefore, it has been considered as a co variable in this study. Statistic at the 161 stations showed that elevation has approximately a normal distribution (Fig. 2). Correlation between rainfall and elevation was calculated 0.08 for these stations. Variography analysis showed that elevation has a high spatial structure with the range of influence (R) about 400 km, the nugget effect (C_o) of 100000 m and the sill (C) of 1100000 m. Based on the results of this study, rainfall data has approximately a normal distribution (Fig. 3) but using logarithm of data, gave a good normal distribution. However, variography analysis showed that corresponding semivariogram has the following characteristic: R= 450 km, $C_{o=}$ 3800 mm, and C=9800 mm. Therefore, it can be said that measured rainfall data in the study area has an error about 40% (C_o /C). After the transformation of data (logarithm), a higher spatial structure was obtained (C_o /C=33%) with a spherical model: R=460 km, C_o =0.08 mm and C=0.24 mm.

Because of unreliable distribution of the climatic stations and a low correlation between rainfall and elevation (as co variable), the study area was divided into homogenous sub regions (Fig.1). The number of stations in the sub regions of 1, 2 and 3 was 25, 45 and 111, respectively. Based on the obtained results, annual rainfall in sub region 1 has a form of normal distribution. Correlation between rainfall and elevation was calculated 0.82 in this sub region. Variography analysis shows that there was not a high spatial structural in this sub region. In this regard, the application of the method such as kriging dose not gives the acceptable estimation of rainfall.

In the sub region of 2, using the logarithm of the data, a high spatial structure has been obtained. Correlation between rainfall and elevation was calculated 0.72 in this sub region. A spherical model was fitted to rainfall with C_o =0.6, C=0.5, R=460 km and measured error percentage =13. The calculated cross semivariogram for rainfall and elevation shows a good spatial structure. A spherical model was fitted for this cross semivariogram (C_0=0, C_1=0.1, R = 275 km).
In the largest sub region throughout the study area (sub region 3), the correlation between rainfall and elevation was calculated equal to 0.2. Annual rainfall dose not respect normal distribution, therefore, by using the transformation of data (logarithm), an acceptable semivariogram has been obtained with the following parameters: C_0=0.03, C_1=0.3, R=1000 km). Fig. 4 shows the semivariogram for annual rainfall in the south west of Iran.

RESULTS OF INTERPOLATION METHODS

ANNUAL RAINFALL

Table 2 presents the results of cross validation technique for different methods of interpolation. All methods include some bias for annual rainfall that carried between 1.1 mm for OK and 30.9 mm for TPSS-CO5. The accuracy of these methods has a vast variations ranged between 38.4 mm for TPSS-2 and 110 mm for the method of OK-CO. The interpolation methods are ranked based on MAE (Table 2), among of them TPSS-2, TPSS-3 and TPSS-CO2 give the most accurate results.

Results of evaluation of different interpolation methods for estimating of annual rainfall in different sub regions were shown in Table 3. In the sub region 1, the methods of TPSS with the power of 2 and elevation as co variable (TPSS-CO2) were the most precise methods with the MAE value of 13.2 mm. In sub region 2, the methods of TPSS-CO3, TPSS-CO4 and TPSS-CO5 were the most precise methods with the MAE value of 30.8 mm. Finally, in the sub region 3, TPSS-2 can be considered as a reliable method. As a conclusion, by dividing the study area in sub region, the suitable method of interpolation was the same as considering the study area as on region, but the precision of estimation improved considerably. Fig. 5 shows the isohyetal rainfall map for annual rainfall using the TPSS method in the south west of Iran.

MONTHLY RAINFALL

The results of evaluation of different interpolation methods for estimating monthly rainfall in the study area are shown in Table 4. The results of this table present the average of 12 months of the year. All methods overestimated monthly rainfall, except OK-LOG that presents an underestimation of 1.35 mm. However, methods include some bias for monthly rainfall that carried between –1.35 mm for OK-LOG and 3.49 mm for TPSS-CO5. The average of MAE values is 6.0 mm that ranged from 4.9 to 9.3 mm. The interpolation methods are ranked based on MAE, among of them TPSS-2 and TPSS-3 show the most precise results with the MAE value of 4.9 and 5.1 mm, respectively.

By dividing study area into sub regions, the results of cross validation of different method of interpolation and for 2, 5, 8 and 12 th month along the year are presented in tables 5 and 6 for sub regions of 2 and 3, respectively. In this situation, the TPSS-2 method also furnishes the most accurate results for estimating monthly rainfall. Fig. 6 shows the iso-contour rainfall map for 2 th, 5 th, 8 th and 11 th month the year using the TPSS method in the south west of Iran.

Table 2. Results of evaluation of different methods in estimating annual rainfall

In the study area.

METHOD	MBE	MAE	Rank
WMA-1	6.9	48.7	5
WMA-2	6.9	47.9	4
WMA-3	6.14	49.2	6
WMA-4	5.4	50.5	9
WMA-5	4.9	51.5	10
OK	1.1	49.9	8
OK-LOG	12.8	49.4	7
OK-CO	15.6	110.1	16
TPSS-2	2.9	38.4	1
TPSS-3	6.2	45.7	3
TPSS-4	9.4	56.7	13
TPSS-5	11.2	106.6	15
TPSS-CO2	23.9	45.6	2
TPSS-CO3	30.5	55.9	12
TPSS-CO4	29.7	58.7	11
TPSS-CO5	30.9	63.4	14

Table 3. Results of evaluation of interpolation methods for estimating annual rainfall in different sub regions.

METHOD	SUB REGION 1		SUB REGION 2		SUB REGION 3	
	MBE	MAE	MBE	MAE	MBE	MAE
WMA-1	-0.8	28.5	1.8	51.0	11.4	53.5
WMA-2	-0.3	26.7	2.9	51.1	10.2	51.9
WMA-3	-0.2	25.7	3.0	54.0	8.6	53.0
WMA-4	-0.3	25.6	2.7	56.2	7.5	54.0
WMA-5	-0.2	25.6	2.4	57.7	6.8	54.6
OK	-0.5	28.8	1.4	50.2	3.02	50.3
OK-LOG	-6.3	31.6	9.4	48.4	9.3	49.8
OK-CO	1.9	48.2	2.3	48.3	-	-
TPSS-2	6.01	28.8	-2.9	47.2	-6.5	44.7
TPSS-3	18.4	48.1	-3.1	46.8	-3.7	57.0
TPSS-4	52.7	122.7	-3.1	46.8	-8.5	61.7
TPSS-5	52.7	122.7	-3.1	46.8	-8.5	61.7
TPSS-CO2	4.7	13.2	-1.0	38.1	-	-
TPSS-CO3	7.3	15.6	-8.9	30.8	-	-
TPSS-CO4	13.8	27.5	-8.9	30.8	-	-
TPSS-CO5	13.8	27.5	-8.9	30.8	-	-

Table 4. Results of cross validation for monthly rainfall in the study area

METHOD	MBE	MAE
WMA-1	0.14	5.5
WMA-2	0.32	5.6
WMA-3	0.44	5.8
WMA-4	0.48	6.0
WMA-5	0.51	6.2
OK	0.16	5.6
OK-LOG	-1.35	5.3
TPSS-2	0.33	4.9
TPSS-3	0.26	5.1
TPSS-4	1.27	6.9
TPSS-5	3.49	9.3

Table 5. Results of evaluation of different methods for estimating monthly rainfall in sub region 2.

METHOD	2 th month		5 th month		8 th month		11 th month	
	MBE	**MAE**	**MBE**	**MAE**	**MBE**	**MAE**	**MBE**	**MAE**
WMA-1	0.7	4.4	0.0	1.3	-0.5	2.3	-1.0	11.8
WMA-2	0.5	4.4	0.0	1.2	-0.4	2.3	0.1	13.0
WMA-3	0.5	4.5	0.3	1.3	-0.4	2.3	0.7	13.9
WMA-4	0.4	4.7	0.3	1.3	-0.4	2.3	1.2	14.6
WMA-5	0.4	4.9	0.3	1.4	-0.4	2.3	1.6	15.2
OK	0.2	4.3	0.0	1.3	0.0	1.4	0.4	13.2
OK-LOG	-0.2	4.6	-0.5	1.1	-0.3	1.5	-4.0	11.9
TPSS-2	-1.7	3.2	-0.1	1.1	0.1	1.4	-2.2	11.2
TPSS-3	-2.2	4.3	-0.1	1.1	0.1	1.6	-2.0	10.0
TPSS-4	-0.8	4.8	9.2	9.8	0.2	1.6	-5.0	14.9
TPSS-5	-0.8	4.8	9.2	9.8	0.2	1.6	12.1	26.6

Table 6. Results of evaluation of different methods for estimating monthly rainfall in sub region 3.

METHOD	2 th month		5 th month		8 th month		11 th month	
	MBE	**MAE**	**MBE**	**MAE**	**MBE**	**MAE**	**MBE**	**MAE**
WMA-1	0.2	3.6	0.3	2.2	-0.2	3.8	1.6	14.8
WMA-2	0.3	3.6	0.4	2.2	0.0	4.0	1.7	14.1
WMA-3	0.3	3.7	0.5	2.2	0.1	4.2	1.6	14.3
WMA-4	0.2	3.8	0.4	2.2	0.2	4.4	1.6	14.6
WMA-5	0.1	3.9	0.4	2.2	0.2	4.5	1.5	14.9
OK	0.1	3.5	0.1	2.3	0.2	4.2	0.3	14.7

OK-LOG	-1.0	3.3	-1.1	2.2	-1.4	3.5	-2.3	14.6
TPSS-2	0.2	3.3	0.0	1.8	0.8	2.2	5.5	15.0
TPSS-3	0.6	3.8	0.1	2.4	0.7	2.5	4.9	15.0
TPSS-4	0.5	3.9	0.1	3.4	0.7	2.5	5.3	14.0
TPSS-5	0.5	3.9	-1.4	4.1	2.8	9.7	5.3	14.0

CONCLUSIONS

Several interpolation methods were evaluated for estimating monthly and yearly rainfall in the south west of Iran. Based on these results, the following conclusions can be extracted from this research:

1- Elevation can be used as a co variable in order to improve the precision of the estimation when the correlation between rainfall and elevation is grater than 0.6.
2- These results show that the range of influence for the elevation of climatic stations is about 400 km, therefore for larger distances, the relation between stations will be random.
3- These results also show that the range of influence for the rainfall of climatic stations is about 400 km, therefore for larger distance, any relation can be found between these stations.
4- The TPSS method is the most precise method in simulating monthly and yearly rainfall. Therefore, these results must be considered in the climatic study such as application of GIS for map preparation.
5- When the concentration of data in a region is low, by dividing of region into homogenous sub regions, the precision of estimation can be improved considerably.

REFERENCES

Abtew, W., J. Obeysekera and G. Shih. 1993. Spatial analysis for monthly rainfall in South Florida. Water Resources Bulletin 29(2): 179-188.

Ahmed, S. and G. De Marsily. 1987. Comparison of geostatistical methods for estimating transmissivity using data on transmissivity and specific capacity. Water Resources Research. 23(9): 1717-173.

Borgan, A. and A. Vizzaccaro. 1997. On the interpolation of hydrologic variables: formal equivalence of multiquadratic surface fitting and kriging. Journal of Hydrology Amesterdam 195☹1-4): 160-171.

Boughton, W.C. 1981. Rainfall variability in Rainfall-Runoff Modeling. Civil Engineering Transactions, I.E. Aust., Vol. CE30: 68-73.

Cook, R.A., S. Mostaghimi and J.B. Campbell. 1993. Assessment of methods for interpolating steady-state infiltration. Transaction of the ASAE 36(5).

Goovarets, P. 2000. Performance comparison of geostatistical algorithms for incorporating elevation into the mapping of precipitation. Journal of Hydrology 228:113-129.

Hall, A. J. and P.A. Barclay. 1975. Method of determining areal rainfall from observed data. In: Prediction in cathchment hydrology, T.G. chaoman and F.X. Dunin (eds), Aust. Acad. Sci: 46-57.

Hosseini, E., J. Gallichand and D. Marcotte. 1994. Theoritical and experimental performance of spatial interpolation methods for soil salinity analysis. Transaction of the ASAE 36(7): 1799-1807.

Voltz, M. and M. Goulard. 1994. Spatial interpolation of soil moisture retention curves. Geoderma 62:109-123.