Journal of Applied Sciences

ISSN 1812-5654

science alert

ANSInet
an open access publisher
http://ansinet.com

Journal of Applied Sciences 9 (1): 128-134, 2009
ISSN 1812-5654

Appraisal of the Geostatistical Methods to Estimate Monthly and Annual Temperature

M.H. Mahdian, S. Rahimi Bandarabady, R. Sokouti and Y. Norouzi Banis
Soil Conservation and Watershed Management Research Institute,
P.O. Box 13445-1136, Tehran, Iran

Abstract: Three geostatistical methods were evaluated for estimation of monthly and annual temperature. These methods consist of Thin Plate Smoothing Splines (TPSS) with and without co variable, Weighted Moving Average (WMA) and Kriging (ordinary and cokriging). Moreover, the elevation was used as co variable. Cross Validation technique was used for comparison of the above-mentioned methods. Based on the results obtained in this study, regression coefficients between elevation and monthly or annual temperature was greater than 0.8. Variography analysis shows good spatial correlation for monthly and annual temperature in these regions. The TPSS method with power of 2 and with elevation as co variable was recognized as the most precise method in estimating monthly and annual temperature. Mean absolute error values for annual and monthly temperature was calculated 1.02 and 1.45°C, respectively). Also, the Cokriging method is ranked as the second method in estimating temperature with MAE = 1.5°C in this study.

Key words: Assessment, cross validation, geostatistics, temperature

INTRODUCTION

Geostatistical methods are very important owing to considering data position, spatial structure and correlation of data. Selection of a method over looking its accuracy may result to erroneous outcome. Due to the large scatter of meteorological stations and high spatial variation of temperature in arid and semi-arid regions, more accuracy is needed in the application of these methods to estimate temperature. In view of the capability of these methods to be used in conjunction with GIS and expansion of the application of GIS, the importance of selecting a suitable method came understood.

Price *et al*. (2000) compared two methods Thin Plate Smoothing Spines, (TPSS) and GIDS for spatial interpolation of monthly and annual temperature and rainfall in east and west of Canada. Results showed, RMSE value was low for TPSS in both of area. Both of methods showed high precision in east area, where lower topography and climatic variability was observed.

Hargrove (2001) used spline method with tension and smooth extension for estimating rainfall in 362 stations in Switzerland. Lynch (2001) surveyed on converting methods of point of daily rainfall onto a rectangular grid. Results showed Weighted moving average, Kriging, splines and inverse distance methods were the best.

Kestevn and Hutchinson (2001) studied on monitoring of climatic variability and spatial modeling of climatic variable on a continental scale in Oceania. They used the TPSS method for temperature and pressure. Their results showed a large movement for oscine climate. Jeffry *et al*. (2001) used TPSS for spatial interpolation of daily climatic variables and Ordinary Kriging for interpolation of monthly and annual rainfall in Australia.

Carrera-Herna´ndez and Gaskin (2007) used some interpolation methods to estimate daily rainfall and minimum and maximum air temperature. The interpolation methods were Ordinary Kriging (OK), Kriging with External Drift (KED) and Block Kriging with External Drift (BKED), Ordinary Kriging in a local neighborhood (OKl) and Kriging with External Drift in a local neighborhood (KEDl). According to the analysis presented, the use of KEDl is recommended to estimate daily rainfall, while KED is recommended to undertake the interpolation of minimum and maximum temperature.

Zhao *et al*. (2005) compared the built linear regression model with geostatistical methods of ordinary kriging, splines and inverse distance weight. Generally; the predictions errors obtained by the geostatistical methods were larger than that by regression method. Ordinary kriging yielded smaller prediction errors than the linear regression of temperature. The worst results were produced by splines.

Corresponding Author: Mohammad Hossein Mahdian, Soil Conservation and Watershed Management Research Institute, P.O. Box 13445-1136, Tehran, Iran

Air temperature was estimated using 5 geostatistical and two regression models (Benavides *et al*., 2007). The geostatistical models include the (OK), developed in the xy plane and in the x, y and z-axis (OKxyz), with zonal anisotropy in the Z-axis, Ordinary Kriging with External Drift (OKED) and universal kriging, using the Ordinary Least Squares (OLS) residuals to estimate the variogram (UK1) or the Generalized Least Squares (GLS) residuals (UK2). The OKED, UK1 and UK2 techniques were more satisfactory than OK in terms of standard prediction error and mean absolute error, which were inferior by 1.8°C, but OKxyz improved the results obtained with those techniques. Moreover, OKxyz, OKED, UK1 and UK2 improved slightly the results of a regression model with UTM coordinates and elevation data as independent variables in terms of bias whereas, a complex regression model, which includes altitude, latitude, distance to the sea and solar radiance as independent variables, showed better results in terms of mean absolute error, under 0.168°C.

Irmak and Ranade (2008) showed that Kriging is the best method of interpolation for estimation of temperature. Kriging showed better acceptability, better ability to produce observed variance and better agreement between observed and predicted values. For all three-interpolation techniques, crucial statistical parameters couldn't identify better option between interpolation with 6 and minimum 4 surrounding stations and interpolation with 4 and minimum 3 surrounding stations (data not shown).

The results of Yan Hong *et al*. (2005) showed interpolation errors for monthly temperatures varying within 0.42-0.83°C and 8-13% for monthly precipitation. These estimates are superior to results produced by methods commonly used in China.

Yan bing (2002) indicated that Spherical, Exponential and Linear models perform as smoothing interpolator of the data, whereas Gaussian and Rational quadratic models serve as an exact interpolator. Spherical, Exponential and Linear models tend to underestimate the values. On the contrary, Gaussian and Rational quadratic models tend to overestimate the values.

In this study, three methods of Thin Plate Smoothing Splines (TPSS), Ordinary Kriging and Weighted Moving Average (WMA) were evaluated and compared for estimating of monthly and annual temperature.

MATERIALS AND METHODS

Study area: This study was conducted at 2007. The study area is located in the South-East of Iran, between 51°, 0' to 62°, 50' E and 25°, 17' to 34°N. Mean annual rainfall is about 188 mm, with a minimum and maximum value about 41.3 and 486 mm, respectively. The elevation is variable, which varied between 0 and 437. The No. of stations is 45 with 22 years data in the study area (Fig. 1).

Variogram analysis: The most important part of the geostatistic is the calculation of semi variance (Bohling, 2005). Its equation is as following:

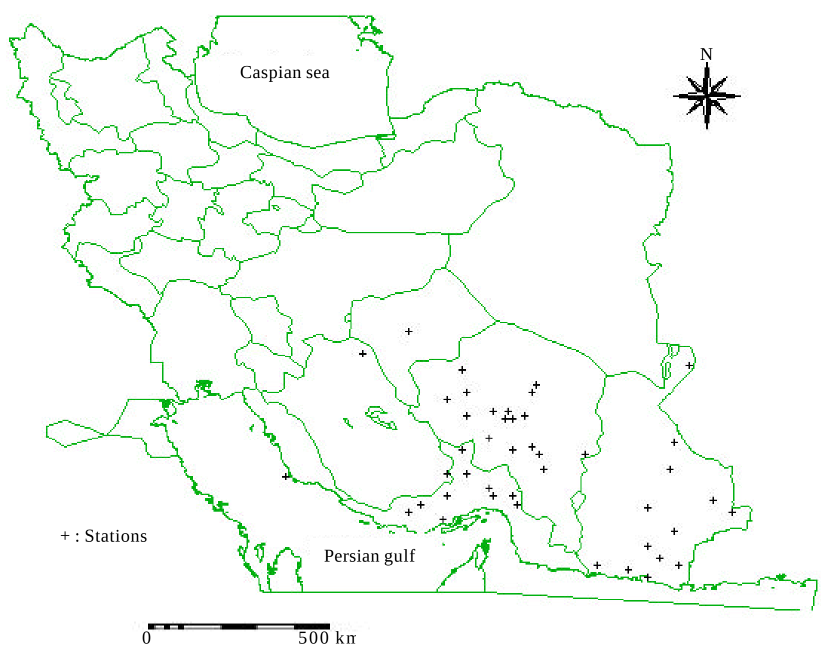

Fig. 1: Study area and position of stations in Iran

$$\gamma(h) = \frac{1}{2n(h)} \sum_{i=1}^{n(h)} \left[z(xi) - z(xi+h)\right]^2 \quad (1)$$

Where:
γ (h) = Semi variance
n (h) = No. of pairs that have h distance
z (xi) = Observation value of the variable x
z (xi+h)= Observation value of x that has h distance from x

Evaluation method: Interpolation methods were evaluated based on Cross Validation (CV) technique. In this technique, initially one point is removed temporarily and that point is estimated, then that point is restored and other point is removed. This is repeated for all of point. In the end, 2 columns obtained, i.e., observed and estimated values. Then Mean Absolute Error (MAE) and Mean Bias Error (MBE) was calculated using the following equations:

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |Z^*(xi) - Z(xi)| \quad (2)$$

$$MBE = \frac{1}{n}\sum_{i=1}^{n} (Z^*(xi) - Z(xi)) \quad (3)$$

Where:
z* = The estimated value
z = The observed value
n = No. of data

MAE and MBE show the precision and bias of estimation, respectively. The closeness of the values indicates the accuracy of the method estimated value is equal to observed value, if MAE and MBE become zero.

Interpolation methods: In this study, TPSS with and without covariable, Kriging and Cokriging and WMA were used to simulate monthly and annual temperature. Table 1 shows the abbreviation for each method. The general equation for different interpolation methods is in the form of the Eq. 4. The difference between of these methods is in the estimation of weighting factor.

$$Z^*(x_i) = \sum_{i=1}^{n} \lambda_i . Z(x_i) \quad (4)$$

Where:
z*(xi) = The estimated value of variable x
z = Observed value of variable x
λi = Weighting value of observed data
n = No. of data
I = Observation point

Table 1: Interpolation methods used in this study

Symbol	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
COK	Co ordinary kriging
TPSS	Thin plate smoothing splines
Without co variable	
TPSS-2	a = 2
TPSS-3	a = 3
TPSS-4	a = 4
TPSS-5	a = 5
With co variable	
TPSS-CO 2	a = 2
TPSS-CO 3	a = 3
TPSS-CO 4	a = 4
TPSS-CO 5	a = 5

a*: Power

For unbiased estimation the $\sum_{i=1}^{n} \lambda_i = 1$ should be maintained.

In WMA, λi is calculated as:

$$\lambda i = \frac{Di^{-\alpha}}{\sum_{i=1}^{n} Di^{-\alpha}} \quad (5)$$

Where:
D_i = The distance between observed and estimated point
α = Power
n = No. of data

In Ordinary Kriging, λi can be obtained as:

$$K.\ \lambda i = B \quad (6)$$

Where:
K = The matrix of covariance between observed data
B = Matrix of covariance between observed and estimated point

For solving this equation, the parameter of semivariogram must be determined.

TPSS is other method used in this study. TPSS is a kind of Kriging with the following covariance function:

$$C(h) = h^k. \operatorname{Log}(h)$$
$$C(o) = \theta \quad (7)$$

Where:
θ = Smoothing parameter
k = m-1
m = Degree of derivation of data
h = Distance between data

RESULTS AND DISCUSSION

Variography analysis: For investigating interpolation methods, initially, the histogram of monthly and annual temperature was drowning. As an example Fig. 2 showed histogram of annual temperature. Then the normality hypothesis is checked using the SPSS software. Normality hypothesis for months of spring and summer and annual temperature with confidence level of 95% and for fall and winter months with confidence level of 99% is not rejected. Skewness values are presented in Table 2. In order to determine the covariable, the correlation coefficients of annual and monthly temperature with elevation were calculated (Table 2); the results showed that the value of these coefficients are high (r>0.8).

The experimental semivariogram was calculated, then, the best model fitted with experimental variogram. Based on different semivariogram, a spatial correlation for monthly and annual temperature was shown in South-East of Iran (Table 3).

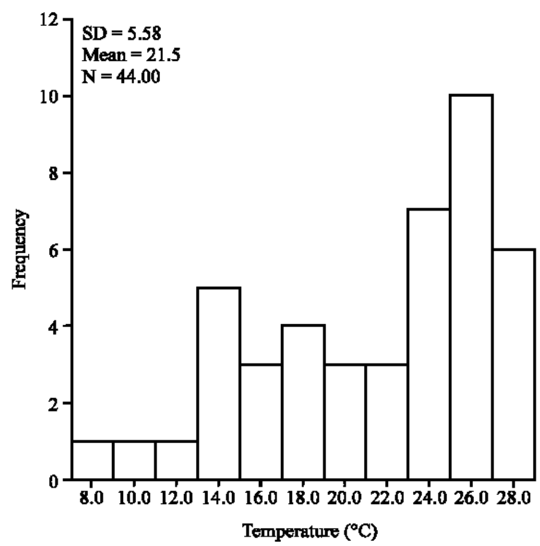

Fig. 2: Histogram of annual temperature in the study area

Table 2: Skewness and correlation coefficient of monthly and annual temperature

Month	Skewness	Correlation coefficient
April	-0.65	-0.96
May	-0.76	-0.96
June	-0.96	-0.95
July	-1.10	-0.91
August	-1.03	-0.90
September	-0.82	-0.90
October	-0.70	-0.93
November	-0.43	-0.94
December	-0.12	-0.90
January	-0.20	-0.93
February	-0.30	-0.94
March	-0.40	-0.96
Year	-0.70	-0.96

Semivariogram was calculated for 12 month of year. In addition, the crossvariogram of temperature and elevation shows spatial correlation. Although, because of negative correlation between elevation and temperature, semivariogram shows negative trend. Table 4 shows the parameters of cross variograms. Range of influence is ranged between 1.5 and 5° and semivariogram models, except November and December, are spherical. Model of annual temperature is exponential with a range about 2°. Range of cross variogram ranged from 1 to 3°. The largest range belongs to December and the smallest belongs to January, February and March. These are in agreement with those published by Yan bing (2002) generally but he was not considered monthly and annual semivariogram models separately. Figure 3 shows variogram and annual temperature model. In the most of months and year the C_0 value is little relative Sill, that meaning deterministic variance is low. In the other words, variograms have high confidence level. C_0 value is between 0.01 and 3°C^{-2} that is lower than obtained by Yan Hong *et al.* (2005) and Benavides *et al.* (2007).

Table 3: Parameters of experimental variogram for monthly and annual temperature

	OK			
Month	C_0 (°c^{-2})	C (°c^{-2})	R (degree)	Model
April	0.10	33.9	2.45	Spherical
May	0.10	38.5	2.40	Spherical
June	0.10	32.2	2.03	Spherical
July	0.01	25.0	1.55	Spherical
August	0.01	30.0	2.20	Spherical
September	0.01	28.1	1.70	Spherical
October	7.00	40.0	5.00	Spherical
November	0.10	38.7	1.08	Exponential
December	0.10	38.9	1.10	Exponential
January	0.10	41.3	1.67	Exponential
February	0.10	39.7	2.79	Spherical
March	0.10	35.4	2.69	Spherical
Year	3.00	37.0	2.00	Exponential

Table 4: Parameters of cross variograms for monthly and annual temperature with elevation

	OK-CO			
Month	C_0 (°c^{-2})	C (°c^{-2})	R (degree)	Model
April	-10	-4752	2.39	Spherical
May	-10	-5039	2.38	Spherical
June	-10	-4562	2.28	Spherical
July	-10	-3795	1.90	Spherical
August	-10	-3620	1.44	Spherical
September	-10	-4025	2.10	Spherical
October	-10	-4655	2.34	Spherical
November	-10	-4877	2.41	Spherical
December	-67	-7241	3.00	Exponential
January	-10	-4895	1.05	Exponential
February	-10	-5003	0.98	Exponential
March	-10	-4819	0.97	Exponential
Year	-10	-4557	2.30	Spherical

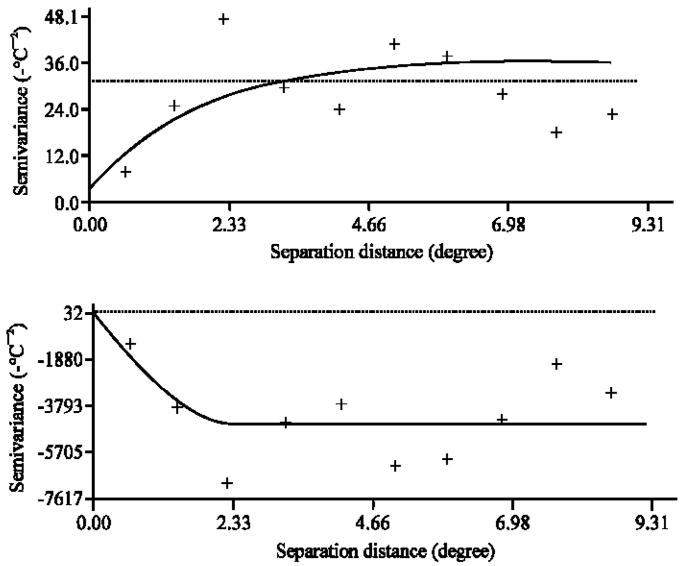

Fig. 3: (a) Variogram and (b) cross variogram calculated for annual temperature

Table 5: Result of comparison different methods in estimating monthly and annual temperature

	March		February		January		December		November		October		September	
Method	MAE	MBE	MAE	MBE	MAE	MBE	MAE	MBE	MAE	MBE	MAE	MBE	MAE	MBE
WMA-1	2.6	-0.6	2.7	-0.6	2.6	-0.5	2.7	-0.5	2.6	-0.5	2.7	-0.5	2.3	-0.8
WMA-2	1.9	-0.4	1.9	-0.3	1.8	-0.3	2.0	-0.2	1.8	-0.3	2.0	-0.4	2.2	-0.6
WMA-3	1.8	-0.3	1.9	-0.3	1.7	-0.2	2.0	-0.1	1.6	-0.2	1.9	-0.4	2.1	-0.4
WMA-4	1.8	-0.2	1.9	-0.2	1.7	-0.2	2.0	-0.1	1.0	-0.2	2.0	-0.4	2.2	-0.3
OK	1.5	-0.1	1.6	-0.1	1.4	-0.1	1.9	-0.1	1.5	-0.1	3.8	-0.6	2.1	-0.2
OK-CO	1.4	0.0	1.5	0.0	1.2	0.0	1.6	0.0	1.3	0.0	1.5	-0.1	2.0	-0.5
TPSS-2	1.5	0.0	2.0	-0.1	1.4	0.3	3.1	-3.0	4.7	-4.6	1.2	-0.0	1.7	0.2
TPSS-3	1.6	0.1	3.9	3.1	2.0	0.1	3.7	-3.2	4.9	-4.7	2.2	-0.1	2.6	0.2
TPSS-4	4.5	-1.6	8.8	1.4	2.0	0.1	6.0	-4.7	7.0	-6.1	6.2	2.3	6.8	-1.3
TPSS-CO 2	1.3	-0.1	3.1	3.0	1.0	0.3	1.5	0.2	1.1	0.0	1.0	0.1	1.5	0.2
TPSS-CO 3	1.4	-0.2	3.0	3.0	1.3	0.4	1.8	0.3	1.2	-0.1	1.3	-0.1	1.4	0.1
TPSS-CO 4	1.8	0.1	-	-	2.4	-0.5	1.8	0.4	1.2	-0.1	1.8	-0.4	1.7	0.4

	April		May		June		July		August		Yearly	
Method	MAE	MBE	MAE	MBE	MAE	MBE	MAE	MBE	MAE	MBE	MAE	MBE
WMA-1	2.7	-0.8	2.9	-0.8	2.9	-0.8	2.8	-0.8	2.6	-0.7	2.5	-0.5
WMA-2	1.9	-0.5	2.2	-0.6	2.2	-0.6	2.2	-0.7	2.1	-0.6	1.8	-0.4
WMA-3	1.7	-0.3	2.0	-0.4	2.0	-0.4	2.1	-0.6	2.0	-0.5	1.7	-0.3
WMA-4	1.7	-0.3	2.0	-0.3	2.0	-0.3	2.2	-0.5	2.1	-0.5	1.8	-0.3
OK	1.4	-0.1	1.8	-0.1	1.8	-0.1	2.0	-0.2	15.6	-1.9	3.4	-0.6
OK-CO	1.3	0.0	1.7	-0.0	1.7	0.0	1.9	0.0	1.8	0.0	1.4	0.0
TPSS-2	5.0	-0.4	1.9	-0.1	2.1	0.0	2.1	0.1	1.4	0.2	1.2	0.1
TPSS-3	5.3	-0.6	2.0	-0.1	2.3	-0.4	2.2	0.1	2.7	-0.1	1.9	-0.1
TPSS-4	8.1	-2.0	5.9	-1.5	2.5	-1.7	2.8	-2.1	6.8	-1.2	5.7	-1.8
TPSS-CO 2	1.0	-0.1	1.3	-0.1	1.5	0.0	1.6	-0.1	1.7	0.1	1.0	-0.2
TPSS-CO 3	1.0	0.0	1.3	-0.1	1.3	0.1	1.8	0.1	1.5	-0.1	1.1	-0.1
TPSS-CO 4	-	-	-	-	-	-	1.8	0.2	1.5	-0.1	1.5	-0.2

Evaluating interpolation methods: The WMA method was executed using the different number of neighborhood points, which 9 neighborhood points were performed the best precision. The mentioned method was accomplished with power 2 to 5. In many months, this method with power 3, gave high accuracy (WMA between 1.6-2.1°C). So, it can be concluded the WMA method (Ryan, 2000) with power of 2 is more accurate for estimation of annual temperature (Table 5).

The TPSS method with power of 2 gave the best precision (with MAE about 1.2 to 5°C and MBE 0.01 to -3°C) in simulating monthly and annual temperature. For

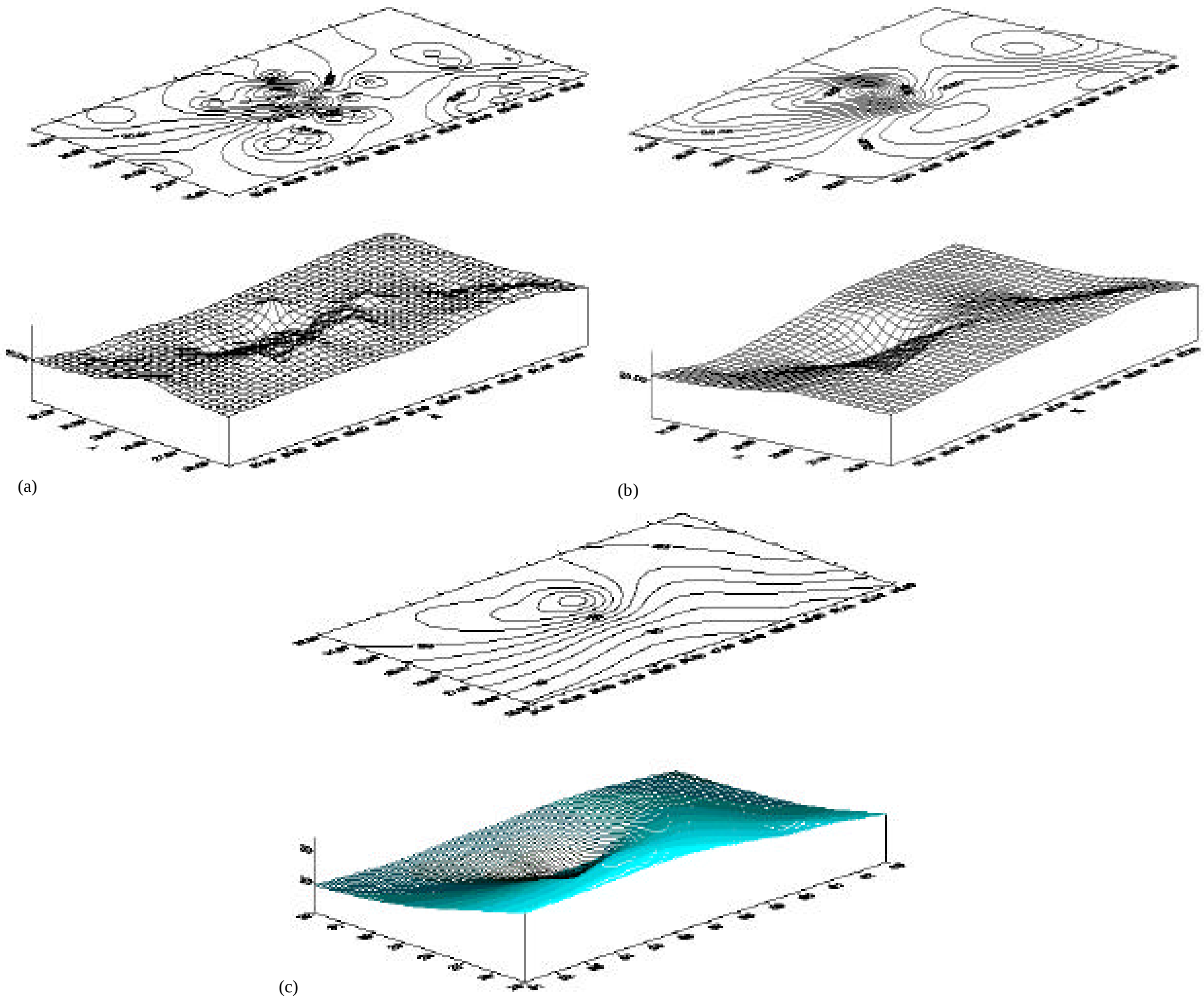

Fig. 4: Maps of annual temperature (a) WMA, (b) Kriging and (c) TPSS

TPSS-CO, the power 2 is more suitable. However, TPSS-CO_2 is better than TPSS-2 (Table 5). This method was not used by other reviews mentioned before and Jeffry *et al.* (2001) used OK to estimate of annual rainfall.

After variogram analysis, Ordinary kriging and cokriging was evaluated using Cross validation technique. Based on the value of MAE and MBE, COK is more accurate than OK in all months.

In brief, the comparison of these methods showed that, TPSS-CO 2 is the best method and COK method is ranked as the second method for estimation of monthly temperature and TPSS for annual temperature (Table 5). This is the case against results obtained by Kestevn and Hutchinson (2001), Zhao *et al.* (2005) and Ayse Irmak and Ranade (2008). Difference between WMA-3, Cokriging and TPSS-CO 2 is low (maximum difference value of MAE is about 1°C). Difference between mean of observed data and estimated value in three methods is very low (maximum difference value of MAE is about 0.3°C). For that, it can be used every three methods, if mean of temperature is considering and this is in agreement with the results of Lynch (2001) but he have used general WMA and TPSS method. Therefore, these results indicate the spatial regularity of climatic data but it differs in different regions.

Figure 4 shows map of annual temperature with different methods. Comparison of maps, showed map of drowned by TPSS-CO 2 method is more accurate and smoother than others.

CONCLUSIONS

Different interpolation methods were evaluated for estimating annual and monthly temperature in this study. Based on the results of this research, the following conclusions can be obtained:

- Monthly and annual temperature has spatial structure and their spatial variation conform the spherical and exponential models. Spatial correlation range is about 2.3°, with respect to cross effect elevation and temperature
- In the most months, the C_0 value is little relative to Sill, that meaning variograms have high confidence level. C_0 value varied between 0.01 and 3°C^{-2}
- Cross variograms have negative trend that is because of negative relation elevation with temperature
- TPSS-CO 2 is more accurate method for interpolation of monthly and annual temperature in South-East of Iran compare to other methods. However, Cokriging method is suitable for monthly temperature estimation
- Because of low difference between the value of MBE for WMA-3, Cokriging and TPSS-CO 2 methods, can used from every three methods, if mean of temperature is considering

ACKNOWLEDGMENT

Authors would like to thank Soil Conservation and Watershed Management Institute for its support of this research.

REFERENCES

Benavides, R., F. Montes, A. Ubio and K. Osoro, 2007. Geostatistical modeling of air temperature in a mountainous region of Northern Spain. Agric. For. Meteorol., 146: 173-188.

Bohling, G., 2005. Introduction to geostatistics and variogram analysis. C and PE 940, 17 October 2005. http://people.ku.edu/~gbohling/cpe940/Variograms.pdf.

Carrera-Herna´ndez, J.J. and S.J. Gaskin, 2007. Spatio temporal analysis of daily precipitation and temperature in the Basin of Mexico. J. Hydrol., 336: 231-249.

Hargrove, W.W., 2001. Interpolation of rainfall in Switzerland using a regularized with tension. Geographic Information and Spatial Technologies Group Oak Ridge National Laboratory. hnw@fire.esd.ornl.gov.

Irmak, A. and P.K. Ranade, 2008. GIS based estimation of spatial distribution of temperature and evapotranspiration in Nebraska. Published by the American Society of Agricultural and Biological Engineers, St. Joseph, Michigan.

Jeffrey, S.J., J.O. Carter, K.B. Moodie and A.R. Beswick, 2001. Using spatial interpolation to construct a comprehensive archive of Australian climate data. Environ. Modell. Software, 16: 309-330.

Kestevn, J. and M.F. Hutchinson, 2001. Spatial modeling of climate variable on a continental scale. Center for Resource and Environmental Studies Institute of Advanced Studies. Australian National University.

Lynch, S.D., 2001. Converting point estimates of daily rainfall onto a rectangular grid. Department of Agricultural Engineering. University of Natal. South Africa. http://gis.esri.com/library/userconf/proc98/proceed/TO200/PAP196/P196.HTM.

Price, D.T., D.W. Mckenny, I.A. Nedler, M.F. Hutchinson and J.L. Kesteven, 2000. A comparison of two statistical methods for interpolation. Canadian monthly mean climate data. Agric. For. Metrol., 101: 81-94.

Ryan, T.P., 2000. Statistical Methods for Quality Improvement. 2nd Edn., Wiley, New York.

Yan Hong, H., A. Nix, F. Mike Hutchinson and H. Trevor Booth, 2005. Spatial interpolation of monthly mean climate data for China. Int. J. Climatol., 25: 1369-1379.

Yan-bing, T., 2002. Comparison of semivariogram models for Kriging monthly rainfall in eastern China. J. Zhejiang Univ. Sci., 3: 584-590.

Zhao, C., Z. Nan and G. Cheng, 2005. Methods for modeling of temporal and spatial distribution of air temperature at landscape scale in the southern Qilian Mountains, China. Ecol. Model., 189: 209-220.