Trends in
Applied Sciences Research

ISSN 1819-3579

Academic
Journals Inc.

www.academicjournals.com

Trends in Applied Sciences Research 7 (6): 467-475, 2012
ISSN 1819-3579 / DOI: 10.3923/tasr.2012.467.475

Assessing the Performance of Spatial Interpolation Methods for Mapping Precipitation Data: A Case Study in Fars Province, Iran

[1]S. Bazgeer, [2]E.A. Oskuee, [2]M. Hagigat and [3]A.R. Darban Astane

[1]Department of Physical Geography, College of Geography, University of Tehran, Tehran, Iran
[2]Islamic Republic of Iran Meteorological Organization, Tehran, Iran
[3]Department of Human Geography, College of Geography, University of Tehran, Tehran, Iran

Corresponding Author: Saeed Bazgeer, Department of Physical Geography, College of Geography, University of Tehran, Azin Alley, Enghelab Ave., Tehran, Iran

ABSTRACT

The Geographic Information System (GIS) is an effective and reliable tool in estimation of spatial distribution of environmental variables. It is well-known that spatial interpolation methods are commonly used for estimating temperature or precipitation when climate stations are few and widely separated. This research aimed to implement and compare the accuracy of different interpolation methods including Inverse Distance Weighted Averaging (IDWA), regularized and tension spline, spherical, circular, exponential and Gaussian kriging for interpolating yearly precipitation of the Fars province in the south of Iran. Long term average of yearly precipitation of synoptic meteorological and rain gauge stations of the study area have been used. As no single method yields an optimum estimation for all regions, Cross Validation (CV) was used to check which method provides the best estimations. Various statistics including Mean Absolute Error (MAE), Mean Bias Error (MBE), Root Mean Square Error (RMSE) and Correlation Coefficient (r) were considered to evaluate the performance of the interpolation methods used. The results revealed that the exponential kriging was associated with less errors compared to other methods and that it could reasonably predict long term average precipitation (RMSE = 83.5 mm and MAE = 60.0 mm). The tension spline, circular and spherical kriging were also performed well and could be used as the second order of priority. The results showed that the application of the Gaussian kriging, IDWA and regularized spline approaches generated more errors compared to other methods.

Key words: Spatial interpolation methods, precipitation estimation, model performance

INTRODUCTION

The Geographic Information System (GIS) is an effective and worthwhile tool in the estimation of the spatial distribution of environmental variables (Dogan, 2007a; Ordu and Demir, 2009; Rabah *et al.*, 2011). The spatial interpolation methods are commonly used for estimating temperature or precipitation when climate stations are not dense (Maciej, 1998; Dogan, 2007b). Rainfall regime is the most important climatic parameter strongly influencing agricultural activities, natural and water resources in hydrological studies (Tzimopoulos *et al.*, 2008). It can considerably vary with time and space; even within a few kilometers distance and very short time scales (Ayanlade, 2008). Hence, having rainfall/precipitation maps is of essential needs for earth science researchers, environment planners and watershed managers (Ninyerola *et al.*, 2007).

Interpolation can be undertaken utilizing simple mathematical models (e.g., inverse distance weighting, trend surface analysis and splines and Thiessen polygons), or more complex models (e.g., geo-statistical methods, such as kriging) (Negreiros *et al*., 2011). There are many studies showing the application of interpolation methods for estimating temperature or precipitation. The methods include distance weighting (Eischeid *et al*., 1995; Lennon and Turner, 1995; Ashraf *et al*., 1997; Dodson and Marks, 1997), interpolating polynomials (Stewart and Cadou, 1981; Tabios and Salas, 1985; Eischeid *et al*., 1995), kriging (Phillips *et al*., 1992; Hammond and Yarie, 1996; Holdaway, 1996; Ashraf *et al*., 1997) and splines (Hulme *et al*., 1995; Lennon and Turner, 1995). Among the methods, the geo-statistical interpolation method has found grounds in climatology because it is based on the spatial variability of variables of interest and can quantify the estimation uncertainty (Martinez-Cob, 1996; Holawe and Dutter, 1999; Sun *et al*., 2008).

The study of precipitation on a global scale has found interests and there have been attempts to map worldwide precipitation. In this respect, there are many studies using various methods ranging from more classic approaches such as the Times Survey Atlas of the World (Bartholomew, 1922) to more recent initiatives that utilize GIS techniques such as the Digital Atlas of the World Water Balance.

To validate a predictive model and choose appropriate algorithm Cross-Validation (CV), which is a statistical method, has been widely used (Arlot and Celisse, 2010). Past applications of the CV approach have given a range of results which have not always been consistent (Falivene *et al*., 2010). For instance, the comparison between two widely used algorithms such as kriging and IDWA have resulted in inconsistent conclusions; some authors have concluded that kriging yields better interpolations (Weber and Englund, 1994; Goovaerts, 2000; Teegavarapu and Chandramouli, 2005), some have not found any significant difference in the results (Dirks *et al*., 1998; Moyeed and Papritz, 2002) and others have advocated that the IDWA methods yields better interpolations (Weber and Englund, 1992; Lu and Wong, 2008).

This research aimed to implement and compare the accuracy of different interpolation methods including Inverse Distance Weighted Averaging (IDWA), regularized and tension spline, spherical, circular, exponential and Gaussian kriging. The accuracy assessment was implemented using Cross Validation (CV) on the interpolated annual precipitation of the Fars Province, Iran.

MATERIALS AND METHODS

Study area and data: The study area is located at the south and south west of Iran (50°36'-55° 35' E and 27°03'-31°40' N), Fars Province, covering an area of about 133000 km^2 (Fig. 1). Agriculture and animal husbandry are the main economic activities of people. The region has various climates, including arid (in south and south-east), temperate (in the center) and mountainous (in north and north-west). The mean annual temperature is 18°C and annual average rainfall is about 285 mm, of which 55% is received in the winter, 27% in the fall, 16% in the spring and 2% in the summer (Http: llirimo.ir). Twenty years historical rainfall data (from 1990-91 to 2009-10) published by the I.R. of Iran Meteorological Organization (IRIMO) were used to develop different spatial interpolation models. Rainfall data were observed from 92 locations at synoptic and rain gauge meteorological stations. In order to reduce the influence of outliers, 6 meteorological stations were added to database of the 92 stations, two stations (Bushehr and Bandar Abbas stations), located near the shores of Persian Gulf, were excluded. This provided adjusted and homogenized climatic data as these two stations were located close to large lakes.

Geospatial interpolation: ArcGIS and GS+ softwares were used for the visualization of model outputs and the geostatistical analysis, respectively (Negreiros *et al*., 2010). Different interpolation

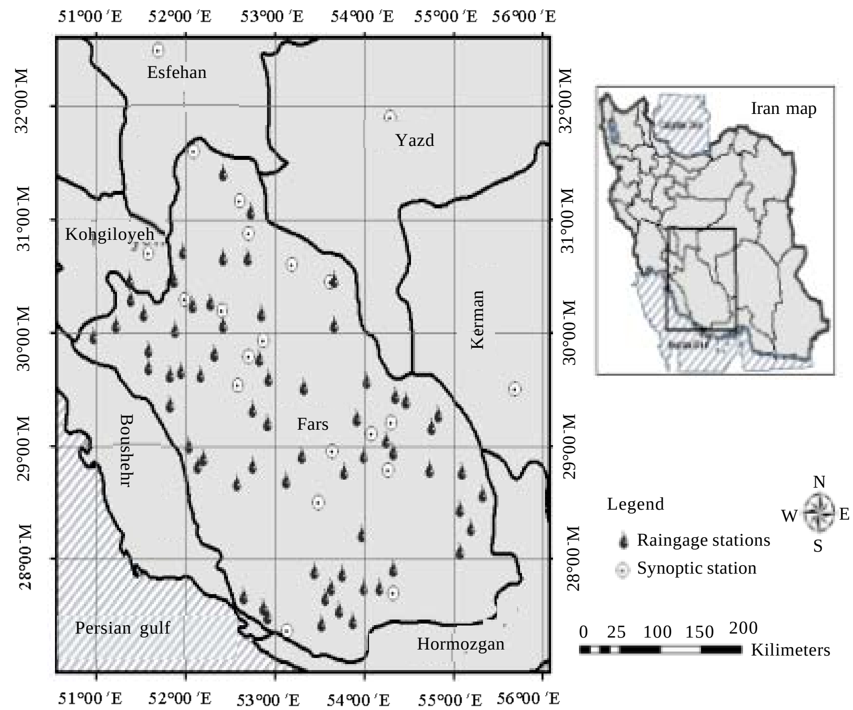

Fig. 1: Location of the study region and meteorological stations

methods including Inverse Distance Weighted Averaging (IDWA), regularized and tension spline as the determinist methods and spherical, circular, exponential and Gaussian kriging as the stochastic methods were used for interpolating yearly rainfall. A brief description of the methods used is provided below:

Inverse distance weighted averaging (IDWA): Inverse distance weighted averaging estimates the variable of interest by assigning more weight to closer points (Burrough and McDonnell, 1998; Goula Bi *et al.*, 2011):

$$\lambda_i = \frac{1/d_i^p}{\sum_{i=1}^{n} 1/d_i^p} \tag{1}$$

where, d_i is the distance between x_0 and x_i, p is a power parameter and n represents the number of sampled points used for the estimation.

The main variable affecting the accuracy of IDWA is the value of the power parameter. Weights diminish as the distance increases, especially when the value of the power parameter increases (Isaaks and Srivastava, 1989; Robinson and Metternicht, 2006). The powers (1 to 5) were examined to select those models with lower error (i.e., IDWA (1), IDWA (2)... IDWA (5).

Spline: The spline method can be thought of as fitting a rubber-sheeted surface through the known points using a mathematical function. Splines consist of polynomials, which describe pieces of a line or surface and they are fitted together so that they join smoothly (Anderson, 2002):

$$Z_{(x,y)} = T_{(x,y)} + \sum_{j=1}^{N} \lambda_j R(r_j) \tag{2}$$

where, N is the number of sampled points used for the estimation, λ is Coefficient of linear equations, r_j is distance from the Sample of point and T is determined by the user.

Kriging: Kriging weights are derived from a statistical model of spatial correlation expressed as semivariograms that characterize the spatial dependency and structure in the data. The experimental variogram measures the average degree of dissimilarity between un-sampled values and a nearby data value and consequently can depict autocorrelation at various distances. The value of the experimental variogram for a separation distance of h (referred to as the lag) is half of the average squared difference between the value at $z(x_i)$ and the value at $z(x_i + h)$ (Robinson and Metternicht, 2006):

$$\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [z(x_i) - z(x_i + h)]^2 \tag{3}$$

where, N (h) is the number of data pairs within a given class of distance and direction.

If the values at $z(x_i)$ and $z(x_i + h)$ are auto-correlated, the result of Eq. 3 relative to an uncorrelated pair of points will be small, Using an analysis of the experimental variogram, a suitable model (i.e., spherical, circular, exponential and Gaussian) is fitted. This is made using weighted least squares and relevant parameters (e.g., range, nugget and sill) are then used in the kriging procedure. The examination of the annual precipitation data indicated that they were not significantly different from a normal distribution.

Models validation: The common validation method in climatological studies has variously been termed as cross-validation (Nalder and Wein, 1998). Cross validation (leaving-one-out method) is based on removing one data point at a time and performing the interpolation for the location of the removed point using the remaining samples. At the final step of cross validation, the difference (residual) between observed and predicted values of that point are calculated. The leaving-one-out approach is repeated until every sample has been, in turn, removed (Davis, 1987) and estimates are calculated for each point. The overall performance of each interpolation method is calculated using Mean Bias Error (MBE), Mean Absolute Error (MAE) and Root Mean Square Error (RMSE), (Solaimani, 2009). Moreover, correlation coefficient (r), a quantitative index of association between observed data (O) and predicted values (P) for N number of cases are calculated. It is described by Pearson's product-moment correlation coefficient (Willmott, 1982):

$$MBE = N^{-1} \sum_{i=1}^{N} (P_i - O_i) \tag{4}$$

$$MAE = N^{-1} \sum^{N} |P_i - O_i| \tag{5}$$

$$RMSE = [N^{-1} \sum_{i=1}^{N} (P_i - Oi)^2]^{0.5} \tag{6}$$

RESULTS AND DISCUSSION

A summary of errors obtained from the application of interpolation models using cross validation tests together with correlation coefficient (r) is presented in Table 1. It is expected that an optimal interpolation algorithm should be basically associated with minimum cross-validation error (Falivene *et al.*, 2010). Similarly, the application of interpolation models using inconsistent and un-homogenized climatic data may produce considerable errors in predictions. Hence, two stations of Bushehr and Bandar Abbas due to their vicinity to Persian Gulf and inconsistency and un-homogeneity that might give to the data base were removed from the considered stations. The inclusion of these stations produced extremely high errors whereas their removal reduced MAE and RMSE in all the models Table 1.

The results showed that exponential kriging produced the most accurate interpolated precipitation data (RMSE = 83.5, MAE = 60.0, MBE = 0-3.7 and r = 0.902). This method was therefore chosen as the most reliable model for interpolating precipitation and used for the final map production (Fig. 2). On the contrary, the worst performance has been observed where the

Table 1: Difference measures of precipitation interpolation model performance

	DM			
SIM	MBE	MAE	RMSE	r
KC	-2.6	60.3	85.8	0.898
KE	-3.7	60.0	83.5	0.902
KG	0.8	74.3	105.8	0.840
KS	-2.7	60.1	85.5	0.898
IDW(2)	-11.3	66.0	95.5	0.873
IDW(3)	-7.8	63.5	89.1	0.889
SR	-8.8	66.4	91.7	0.881
ST	0.1	63.9	85.5	0.899

DM: Difference measure, SIM: Spatial interpolation model, MAE: Mean absolute error, RMSE: Root mean square error, r: Correlation coefficient, KC, KE, KG and KS, circular, exponential, Gaussian, spherical kriging, respectively; IDWA (2) and (3): Inverse distance weighted averaging with power 2 and 3, SR and ST: Regularized and tension spline

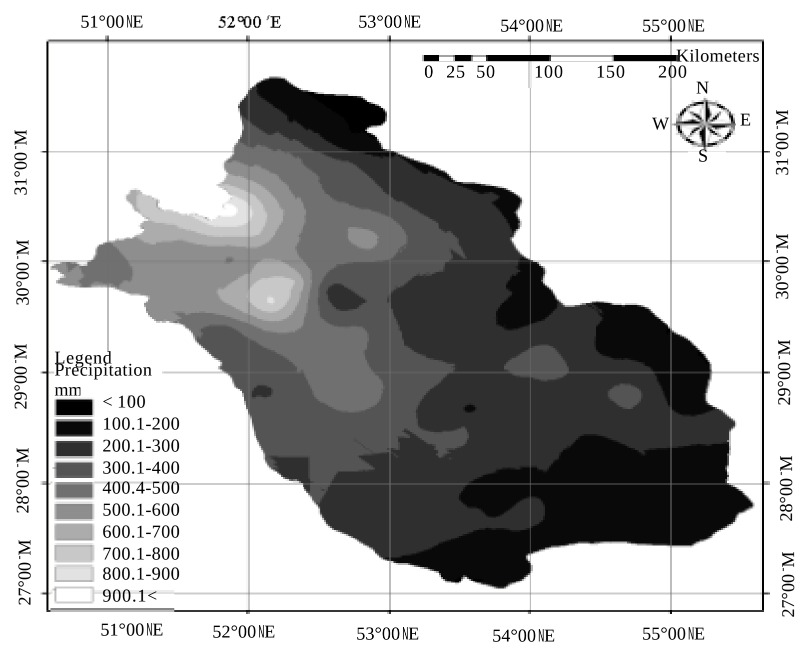

Fig. 2: Interpolated precipitation map using exponential kriging

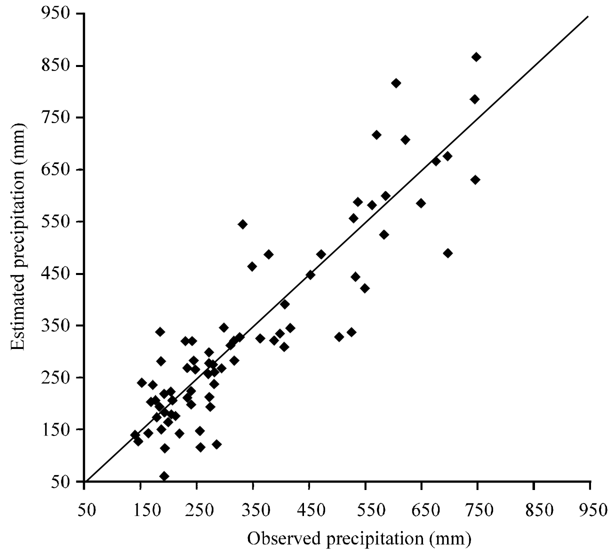

Fig. 3: Comparison between observed and estimated precipitation using exponential kriging (KE)

IDWA (2) and (3), Gaussian, kriging and regularized spline methods were used (Table 1). Previous studies confirm our findings in this research. For instance, Weber and Englund (1994) and Goovaerts (2000) showed that the kriging method produced more accurate results compared to IDWA. They justified this as the kriging method could take the relative positions of sampling points and their distance from the interpolated point into account.

The results also showed that mean error was low for all kriging models, except for Gaussian, but the lowest value of mean error was resulted from the exponential kriging (KE) application. Moreover, the highest value of r (r = 0.902) was obtained using the KE compared to all other models. On the basis of MBE, the tension spline and Gaussian kriging produced lower errors. This might stem from the fact that MBE would not provide enough diagnostic to justify its inclusion, over other measures, in an array of model evaluation measures (Willmott, 1982).

The lowest MAE and RMSE were obtained using the exponential, circular and spherical kriging methods. Willmott (1982) argued that RMSE and MAE could be among the best overall measures of the model performance as both measures summarize the mean difference in the units of observed and predicted values. In a comparison of methods for estimating annual precipitation, the exponential and Gaussian kriging methods created the highest (r = 0.902) and the lowest (r = 0.842) coefficient of correlation, respectively (Table 1). Figure 3 shows the comparison between observed and estimated annual precipitation. In overall, the Fars Province can be divided into the following rainfall-altitudes relationships parts:

- Altitudes of more than 2000 m, with annual rainfall of about 368 mm
- Altitudes range between 1500 and 2000 m, with annual rainfall of about 308 mm
- Altitudes range between 700 and 1500 m, with annual rainfall of about 282 mm
- Altitudes of less than 700 m, with annual rainfall of about 264 mm

CONCLUSION

This research aimed to implement and compare the accuracy of different interpolation methods using cross validation errors for interpolating total annual precipitation of the Fars Province, Iran.

The results showed that the application of the exponential kriging method produced less errors between observation and prediction (RMSE = 83.5 and MAE = 60.0 mm) and it was chosen to predict long term average precipitation. The tension spline, circular and spherical kriging methods were also reasonably performed well and could be used as the second order of priority. In contrast, the Gaussian kriging, IDWA and regularized spline methods were associated with an unacceptable level of errors and could not be justifiably used for interpolating long term average precipitation in the study area.

ACKNOWLEDGMENT

We wish to thank the I.R. of Iran Meteorological Organization (IRIMO) for providing us the precipitation data of meteorological stations in Fars Province.

REFERENCES

Anderson, S., 2002. An evaluation of spatial interpolation methods on air temperature in Phoenix, AZ. http://www.cobblestoneconcepts.com/ucgis2summer/anderson/anderson.htm

Arlot, S. and A. Celisse, 2010. A survey of cross-validation procedures for model selection. Stat. Surv., 4: 40-79.

Ashraf, M., J.C. Loftis and K.G., Hubbard, 1997. Application of geostatistics to evaluate partial weather station networks. Agric. For. Meteorol., 84: 225-271.

Ayanlade, A., 2008. Spatiotemporal interpolation of seasonal rainfall variability in Guinea Savanna part of Nigeria. Advances in Natural and Applied Sciences, pp: 233. http://www.thefreelibrary.com/Spatiotemporal+interpolation+of+seasonal+rainfall+variability+in...-a0215515488

Bartholomew, J.G., 1922. The Times Survey Atlas of the World. The Times, London, UK., Pages: 122.

Burrough, P.A. and R.A. Mc Donnell, 1998. Creating Continuous Surfaces from Point Data. In: Principles of Geographic Information Systems, Burrough, P.A. and M.F. Goodchild, R.A. McDonnell, P. Switzer, M. Worboys (Eds.)., Oxford University Press, Oxford, UK., pp: 535-550.

Davis, B.M., 1987. Uses and abuses of cross-validation in geostatistics. Math. Geol., 19: 241-248.

Dirks, K.N.D., J.E. Hay, C.D. Stow and D. Harris, 1998. High-resolution studies of rainfall on Norfolk Island: Part II: Interpolation of rainfall data. J. Hydrol., 208: 187-193.

Dodson, R. and D. Marks, 1997. Daily air temperature interpolated at high spatial resolution over a large mountainous region. Clim. Res., 8: 1-20.

Dogan, H.M., 2007a. Climatic portrayal of Tokat province in Turkey: Developing climatic surfaces by using LOCCLIM and GIS. J. Boil. Sci., 7: 1060-1071.

Dogan, H.M., 2007b. High resolution climatic surfaces of Nallihan ecosystem in Turkey; a convenient methodology to create climate maps. J. Applied Sci., 7: 654-662.

Eischeid, J.K., F.B. Baker, T.R. Karl and H.F. Diaz, 1995. The quality control of long-term climatological data using objective data analysis. J. Appl. Meteorol., 34: 2787-2795.

Falivene, O., L. Cabrera, R. Tolosana-Delgado and A. Saez, 2010. Interpolation algorithm ranking using cross-validation and the role of smoothing effect: A coal zone example. Comput. Geosci., 36: 512-519.

Goovaerts, P., 2000. Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. J. Hydrol., 228: 113-129.

Goula Bi, T.A., V. Fadika and G.E. Soro, 2011. Improved estimation of the mean rainfall and rainfall-runoff modeling to a station with high rainfall (Tabou) in South-Western Côte d`ivoire. J. Applied Sci., 11: 512-519.

Hammond, T. and J. Yarie, 1996. Spatial prediction of climatic state factor regions in Alaska. Ecoscience, 3: 490-501.

Holawe, F. and R. Dutter, 1999. Geostatistical study of precipitation series in Austria: Time and space. J. Hydrol., 219: 70-82.

Holdaway, M.R., 1996. Spatial modeling and interpolation of monthly temperature using kriging. Clim. Res., 6: 215-225.

Hulme, M., D. Conway, P.D. Jones, T. Jiang, E.M. Barrow and C. Turney, 1995. Construction of a 1961-1990 European climatology for climate change modelling and impact applications. Int. J. Climatol, 15: 1333-1363.

Isaaks, E.H. and R.M. Srivastava, 1989. An Introduction to Applied Geostatistics. Oxford University Press, New York, USA., Pages: 561.

Lennon, J.J. and J.R.G. Turner, 1995. Predicting the spatial distribution of climate: Temperature in great Britain. J. Anim. Ecol., 64: 370-392.

Lu, G.Y. and D.W. Wong, 2008. An adaptive inverse-distance weighting spatial interpolation technique. Comp. Geosci., 34: 1044-1055.

Maciej, T., 1998. Spatial interpolation and its uncertainty using automated anisotropic Inverse Distance Weighting (IDW)-Cross-Validation/Jackknife approach. J. Geographic Inf. Decis. Anal., 2: 18-30.

Martinez-Cob, A., 1996. Multivariate geostatistical analysis of evapotranspiration and precipitation in mountainous terrain. J. Hydrol., 174: 19-35.

Moyeed, R.A. and A. Papritz, 2002. An empirical comparison of kriging methods for nonlinear spatial point prediction. Math. Geol., 34: 365-386.

Nalder, I.A. and R.W. Wein, 1998. Spatial interpolation of climatic normals: Test of a new method in the Canadian boreal forest. Agric. For. Meteor., 92: 211-225.

Negreiros, J., A.C. Costa and M. Painho, 2011. Evaluation of stochastic geographical matters: Morphologic geostatistics, conditional sequential simulation and geographical weighted regression. Trends Applied Sci. Res., 6: 237-255.

Negreiros, J., M. Painho, F. Aguilar and M. Aguilar, 2010. Geographic information systems principles of ordinary kriging interpolator. J. Applied Sci., 10: 852-867.

Ninyerola, M., X. Pons and J.M. Roure, 2007. Monthly precipitation mapping of the Iberian Peninsula using spatial interpolation tools implemented in a Geographic Information System. Theor. Appl. Climatol., 89: 195-209.

Ordu, S. and A. Demir, 2009. Determination of land data of Ergene basin (Turkey) by planning geographic information systems. J. Environ. Sci. Technol., 2: 80-87.

Phillips, D.L., J. Dolph and D. Marks, 1992. A comparison of geostatistical procedures for spatial analysis of precipitation in mountainous terrain. Agric. For. Meteorol., 58: 119-141.

Rabah, F.K.J., S.M. Ghabayen and A.A. Salha, 2011. Effect of GIS interpolation techniques on the accuracy of the spatial representation of groundwater monitoring data in Gaza strip. J. Environ. Sci. Technol., 4: 579-589.

Robinson, T.P. and G. Metternicht, 2006. Testing the performance of spatial interpolation techniques for mapping soil properties. J. Comput. Elect. Agric., 50: 97-108.

Solaimani, K., 2009. A study of rainfall forecasting models based on artificial neural network. Asian J. Applied Sci., 2: 486-498.

Stewart, R.B. and C.F. Cadou, 1981. Spatial estimates of temperature and precipitation normals for the Canadian Great Plains. Research Branch, Agriculture Canada, Ottawa, Canada.

Sun, R., Z. Baiping and T. Jing, 2008. A multivariate regression model for predicting precipitation in the Daging mountain. Mt. Res. Dev., 28: 318-325.

Tabios III, G.Q. and J.D. Salas, 1985. A comparative analysis of techniques for spatial interpolation of precipitation. Water Resour. Bull., 21: 365-380.

Teegavarapu, R.S.V. and V. Chandramouli, 2005. Improved weighting methods, deterministic and stochastic data-driven models for estimation of missing precipitation records. J. Hydrol., 312: 191-206.

Tzimopoulos, C., L. Mpallas and C. Evangelides, 2008. Fuzzy model comparison to extrapolate rainfall data. J. Environ. Sci. Technol., 1: 214-224.

Weber, D. and E. Englund, 1992. Evaluation and comparison of spatial interpolation. Math. Geol., 24: 381-391.

Weber, D.D. and E.J. Englund, 1994. Evaluation and comparison of spatial interpolation II. Math. Geol., 26: 589-603.

Willmott, C.J., 1982. Some comments on the evaluation of mode performance. Bull. Am. Meterrol. Soc., 63: 1309-1313.