Assessment of Three Spatial Interpolation Models to Obtain the Best One for Cumulative Rainfall Estimation (Case study: Ramsar District)

Hasan Zabihi, Anuar Ahmad, Mohamad Nor Said
Department of Geoinformation,
Faculty of Geoinformation and Real Estate,
Universiti Teknologi Malaysia,
81310 UTM, Johor Bahru, Johor, Malaysia

Abstract—There has been an increasing use of predictive spatial distribution of rainfall patterns for planning and regional management decisions. This study focused on three interpolation techniques including ordinary kriging module, linear regression method and inverse distance weighted (IDW) that were used to obtain the reliable spatial distribution of cumulative rainfall parameters in the study area. First, we selected twenty two meteorological stations in and around the study area. Second, rainfall data were analyzed in geographical information systems (GIS) to determine the accuracy of three models. Third, the distribution pattern was also validated by field investigations. Finally, the reliability map of rainfall was produced in GIS software. The results demonstrate that linear regression significantly performed better than inverse distance weighted (IDW) and ordinary kriging model.

Keywords— Assessment; Cumulative rainfall; Spatial interpolation Models; RMS; GIS.

I. INTRODUCTION

The effective of rainfall patterns are basically one the essential phenomenon in agricultural production as successful cultivation areas all over the world [1, 2]. Lack of meteorological station in regions is one of the main problems for future management and estimation of rainfall in given areas. However, using prediction methods to estimate amount of rainfall can help manager for future planning. On the other hand, the accuracy and reliability of these models also are very crucial for policy maker and managers. Geostatistical models are reported in numerous textbooks such as Kriging (plain geostatistics); environmental correlation (e.g. regression-based); and hybrid models (e.g. regression-kriging) [10, 11]. To present the accuracy of models, three methods including kriging, regressions and IDW were conducted to compare and benefits. Objective of this study focuses in comparison with three methods to obtain the best-performed one. The results show that regression method is well adopted with field real data set in study area and reliable one. In previous study, [7] used the geostatistical methods of kriging and cokriging to estimate the sodium adsorption ratio in an agricultural field. [8] produced a radon distribution map using the kriging and GIS techniques.

II. STUDY AREA

The region is located in the northern part of Iran. Ramsar region is situated in the west of Mazandaran province, borders The Caspian Sea to the north and The Alborz Mountains range to the south. This region is one of the most important agricultural areas in Iran. The geographic coordinates of the study area are located between latitudes 36°32′00″ to 36°59′11″ N and longitudes 50°20′30″ to 50°47′12″ E. The total study area covers approximately 729.7 km^2. The altitude of Ramsar County starts at a height of -20 meters near The Caspian Sea to 3620 meters above sea level. A map of the study area is shown in Fig. 1.

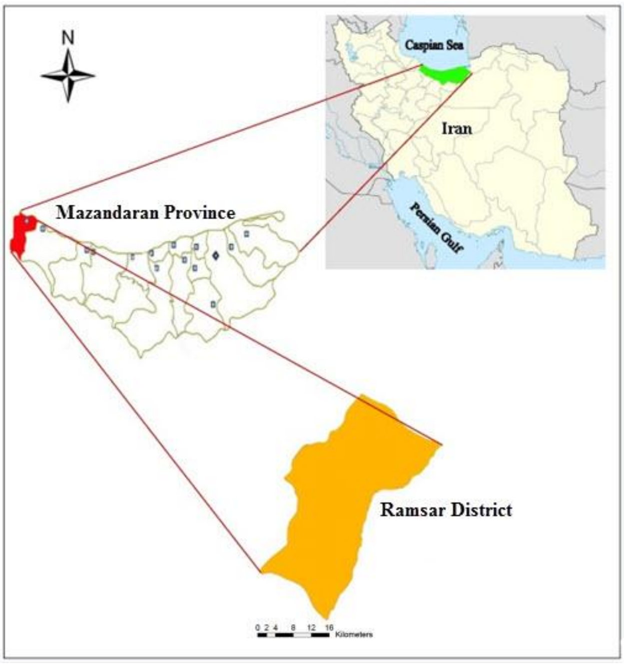


Fig.1 The location of study area

III. METHODOLOGY

Three conventional interpolation methods including linear regression, kriging and IDW are applied in this study to determine the best-performed one. Their principles are well explained elsewhere [3, 4]. Rainfall data was obtained from the Mazandaran Meteorological organization of 30 years (1980–2010). Rainfall data analysis was conducted by using curve expert software version 1.4 as a comprehensive curve fitting system. It employs a large number of regression models (both linear and nonlinear) as well as various interpolation schemes in the most precise and convenient way. In addition, the user may define any customized model desired for use in a regression analysis. The next step, elevation as independent imported in X axis or Y axis to get regression (linear fit) and coefficient data based on following equation "(1)":

$$Y = a + b\,x \qquad (1)$$

Moreover, kriging module and IDW are extracted in Geostatistical analyst into ArcGIS version 10.1. The coefficients give the values of the relevant parameters for the current model. These coefficients are always expressed as a and b. In other words, coefficients give the most vital information about the model in the graph. Twenty two meteorological stations were used that are given in Table 1. Rainfalls' prediction, measure, errors and regression equation are shown in Table 2, 3, 4 and 5.

Table 1: Top selection meteorological stations in and around study area (Climatic Atlas of Iran, I.R. of Iran Meteorological Organization, 2012)

Name	Lat.	Long.	Station Class	Elev. (m)	Total rainfall
Chaboksar	36.96	50.58	Raingage	-10	1012.1
Mianlat	36.9	50.6	Raingage	350	633.5
Sar limak	36.85	50.68	Raingage	200	659
Azarak andokoh	36.85	50.7	Raingage	100	713.6
Galesh mahaleh	36.81	50.73	Raingage	74	739.5
Chapar sar	36.82	50.79	Raingage	5	862.6
Soleiman abad	36.8	50.8	Raingage	25	616.5
Tonekabon	36.81	50.87	Raingage	-15	747.2
Golali abad	36.7	50.85	Raingage	50	801
Lirasar	36.68	50.89	Raingage	300	688.5
Balaoshtoj	36.64	50.76	Raingage	800	806.9
Shahnetrash	36.64	50.72	Raingage	1550	668
Javaherdeh	36.85	50.48	Raingage	2000	510
Tomol	36.64	50.41	Raingage	2010	551.9
Holoo-an	36.58	50.83	Raingage	900	732.2
Tole lat	36.98	50.3	Raingage	40	902.5
Malekot	36.89	50.11	Raingage	1428	471.2
Zar abad	36.49	50.43	Raingage	1790	466.4
parchkouh	36.63	50.17	Raingage	1600	554.3
Ramsar	36.9	50.66	Synoptic	-21	1207.1
Khorram abad	36.78	50.87	Climatology	50	1079
Jannat roudbar	36.75	50.58	Raingage	1700	650.5

To assess accuracy, we calculated the root mean square error (RMSE) is shown in following formula that is reported by [6] as in equation "(2)".

$$\text{RMSE} = \sqrt{\frac{i}{n}\sum_{i=1}^{n}\left(X(observed) - X(estimated)\right)^2} \qquad (2)$$

IV. RESULTS AND DISCUSSION

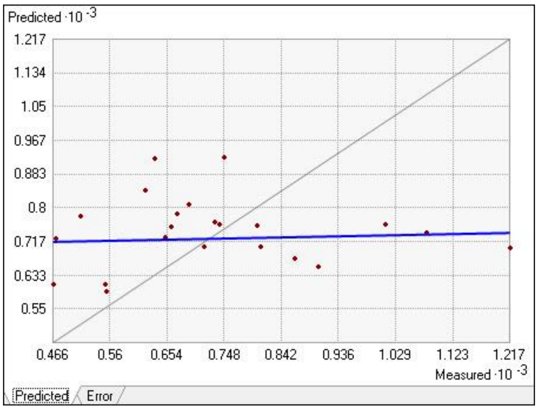
As a result of this analysis and to assess accuracy, we calculated the RMS of three methods and relationship between rainfall, altitude and stations. The root mean square error (RMS) indicates the spread of how far the computed values deviate from the observed. Based on Table 2, 3 and 4 higher RMS was obtained by linear regression and IDW. A lowest prediction error was achieved by kriging model. However, IDW can produce "bulls eyes" around data locations. Cumulative rainfall is classified into five classes. With regard to different interpolation analysis, it obvious that linear regression well performed for prediction of rainfall distributions in study area. The results of spatial interpolation into IDW method, ordinary kriging and linear regression are illustrated in the following Table 2, 3, 4 and 5. Moreover, Table 2 demonstrates cumulative rainfall into IDW interpolation method. Using IDW, the weight of any known point is set inversely proportional to its distance from the estimated point. It is calculated as basic formula follows:

$$v = \frac{\sum_{i=1}^{n}\frac{1}{d_i}v_i}{\sum_{i=1}^{n}\frac{1}{d_i}} \qquad (3)$$

v = value to be estimated
v i = known value
di..., dn= distances from the n data points to the point estimated n.
Root mean square value = 205.93.

Table 2 Spatial prediction of rainfall using IDW method

Source ID	Included	Measured	Predicted	Error
19	Yes	1217	702.1...	-514.8663625315157
20	Yes	1079	738.8...	-340.1736390699001
0	Yes	1012.1	759.1...	-252.9508188777345
15	Yes	902.5	654.8...	-247.64838621920023
5	Yes	862.6	675.2...	-187.3473348291352
10	Yes	806.9	705.9...	-100.93302258402798
8	Yes	801	756.3...	-44.650651517297206
3	Yes	713.6	705.0...	-8.501875295677905
4	Yes	739.5	759.8...	20.35880549394892
14	Yes	732.2	766.6...	34.468424608757346
18	Yes	554.3	595.2...	40.933976218801604
13	Yes	551.9	611.9...	60.041472322999084
21	Yes	650	727.9...	77.94771792748077
2	Yes	659	753.6...	94.60059798149314
11	Yes	668.8	786.9...	118.10252367038481
9	Yes	688	808.7...	120.7001059292379
17	Yes	466.4	612.4...	146.06863239058828
7	Yes	747	926.4...	179.4136104801944
6	Yes	616.5	845.2...	228.70882366012142
16	Yes	471.2	726.1...	254.96415791208352
12	Yes	510	781.5...	271.536480650195
1	Yes	633.5	924.2...	290.7339751784075



Regression function= 0.0296094924757434 * X + 701.926303235849

The second method of spatial interpolation performed in this study was kriging. Kriging is used to estimate unknown values from data observed at known locations. Among the various existing Kriging techniques, both exact and inexact methods could be selected, depending on the measurement error model [5, 9]. A more detailed explanation of the kriging method is given by [12] and [13]. Kriging are based on statistical models that include autocorrelation. In reality, there are the statistical relationships among the measured points. Because of this, geostatistical techniques not only have the capability of producing a prediction surface but also provide some measure of the certainty or accuracy of the predictions [14, 15]. Table 3 shows ordinary kriging module using cumulative rainfall of twenty two meteorological stations.

The general formula for both interpolators is formed as a weighted sum of the data:

$$\hat{Z}(s_0) = \sum_{i=1}^{N} \lambda_i Z(s_i) \qquad (4)$$

where:

$Z(s_i)$ = the measured value at the ith location
λ_i = an unknown weight for the measured value at the ith location
s_0 = the prediction location
N = the number of measured values
Root mean square value = 180.21.

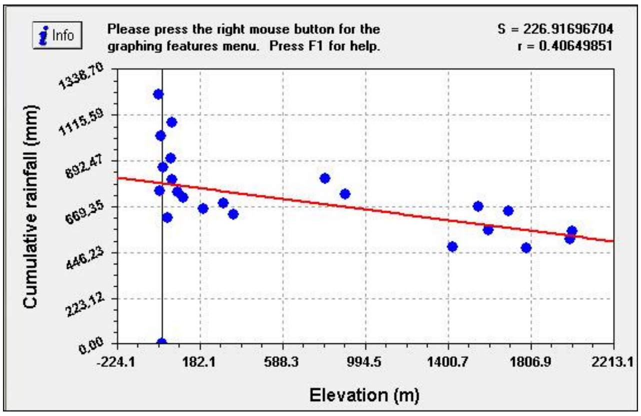
Table 3 Spatial prediction of rainfall using ordinary kriging method

Source ID	Included	Measured	Predicted	Error
19	Yes	1217	748.9909673752457	-468.00903262475435
20	Yes	1079	737.5829019810699	-341.4170980189301
0	Yes	1012.1	759.1787871984873	-252.92121280151275
15	Yes	902.5	651.2217570963477	-251.2782429036523
5	Yes	862.6	787.7771808613974	-74.82281913860265
10	Yes	806.9	711.1394436269975	-95.76055637300249
8	Yes	801	762.5956490304284	-38.404350969571624
7	Yes	747	783.1284049740465	36.128404974046475
4	Yes	739.5	796.3186527754882	56.818652775488204
14	Yes	732.2	732.7560548323822	0.5560548323821877
3	Yes	713.6	792.5712762617584	78.97127626175836
9	Yes	688	790.877460684812	102.87746068481204
11	Yes	668.8	703.4515793136753	34.65157931367537
2	Yes	659	777.3915149263781	118.39151492637814
21	Yes	650	702.6386314079576	52.63863140795763
1	Yes	633.5	800.892498990472	167.39249899047195
6	Yes	616.5	806.846158017586	190.34615801758605
18	Yes	554.3	551.008754200805	-3.2912457991949395
13	Yes	551.9	603.2201406101872	51.320140610187195
12	Yes	510	742.9165505493342	232.91655054933415
16	Yes	471.2	697.816868992722	226.616868992722
17	Yes	466.4	620.3926650471302	153.99266504713023

Regression function= 0.0371076672402633 * X + 717.536544660987

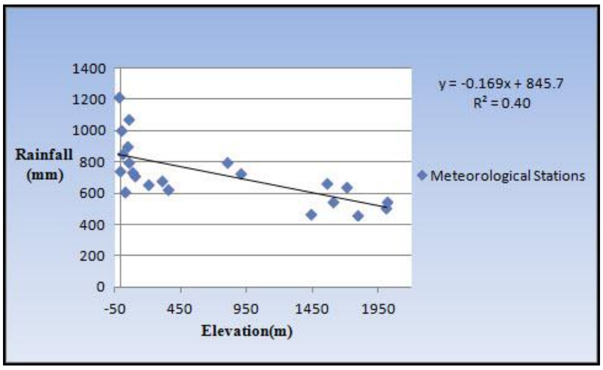
The last method that was used in this study was regression approach. Regression of cumulative rainfall is addressed into Curve expert 1.4 and Excel that are shown in Table 4 and 5.

Table 4 Regression trend between cumulative rainfall and elevation using Curve Expert software



Y= -0.1691* X + 845.46, R^2 = 0.40 and standard error = 226.91.

Table 5 Regression trend between cumulative rainfall and elevation using Excel 2007



V. CONCLUSION

Root mean squared (RMS) of regression method is higher than kriging and IDW methods in terms of rainfall patterns. In reality, the methods that have the lowest root-mean-square prediction error when performing cross-validation have been adopted. Indeed, the smaller the root-mean-square prediction error, the closer the predictions are to their true values and better are the interpolation method [5]. As a comparison of methods, however; an ordinary kriging method is the most accurate than linear regression and IDW methods in terms of lower error. But, linear regression method is reliable due to altitude. it is also considered as an essential spatial characteristic. In other words, based on Table 5 in Ramsar district, rainfall trend would be decreased by increasing altitude (e.g. R^2 value < 0.50) across the study region. R2 describes the proportion of the variance in measured data explained by the model. R^2 ranges from 0 to 1, with higher values indicating less error variance, and typically values greater than 0.5 are considered acceptable [16, 17]. Hence, this suggests that IDW and kriging perform very poorly in spatial interpolation, which may not be suitable for interpolation analysis to this purpose. Thereby, in spatial prediction models, regression performs accurately rainfall map than IDW and kriging interpolation performance in study area.

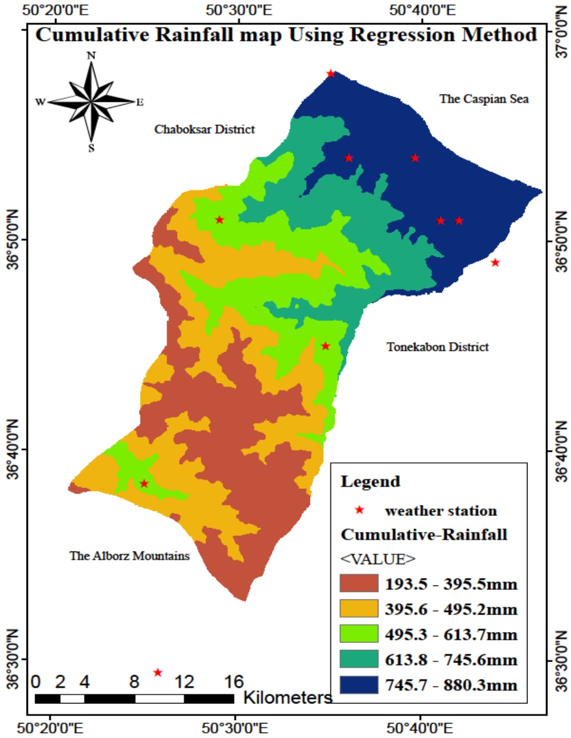
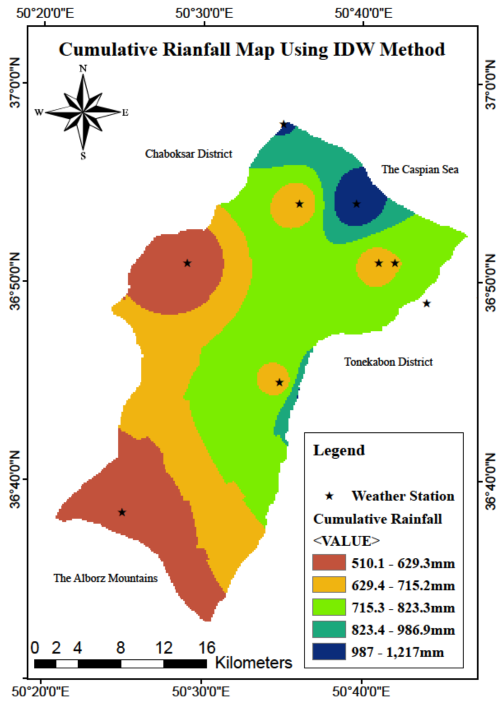
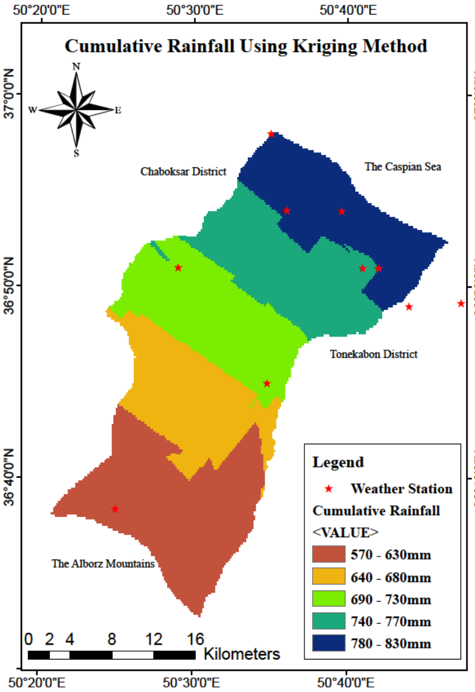


Fig. 2 Three different rainfall prediction maps using regression, kriging and IDW spatial interpolation





ACKNOWLEDGMENTS

We are thankful to Universiti Teknologi Malaysia for generously funding the research. The authors also wish to express sincere thanks to Iran Citrus Research Institute for kind support data during this study.

REFERENCES

[1] S.S., Panda, G., Hoogenboom, J. Paz, "Distinguishing blueberry bushes from mixed vegetation land-use using high resolution satellite imagery and geospatial techniques". Computers and Electronics in Agriculture. 2009, 67: 51-59.

[2] W. Schlenker, and M. J. Roberts, "Nonlinear effects of weather on corn yields". Applied Economic Perspectives and Policy. 2006, 28: 391–398.

[3] C. D., Lloyd, "Assessing the effect of integrating elevation data into the estimation of monthly precipitation in Great Britain". J. Hydrol. 2005, 308: 1– 4.

[4] P., Goovaerts, "Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall". J. Hydrol. 2000, 228: 113–129.

[5] ESRI. Guideline for User of ArcGIS 9.2. ESRI Company, 2009.

[6] C. J., Willmott, "On the evaluation of model performance in physical geography". In: Gaile, G. L. Willmott, C. eds., spatial statistics and models. D. Reidel Publishing Company, Dordrecht. 1984, PP. 443-460.

[7] L., Pozdnyakova, and R. Zhang, "Geostatistical analyses of soil salinity in a large field". Precision Agriculture, 1999, 1: 153–165.

[8] H. C., Zhu, J. M., Charlet, and, A. Poffijn, "Radon risk mapping in Southern Belgium: An application of geostatistical and GIS techniques". The Science of the Total Environment, 2001, 272: 203–210.

[9] ESRI, Using Arc GIS Geostatistical Analyst, Printed in the USA, 2003.

[10] C.V., Deutsch, A. G., Journel, "Geostatistical Software Library and User's Guide," 2nd ed. Oxford University Press, New York, USA, 1998, pp. 369.

[11] T. Hengl, "A Practical Guide to Geostatistical Mapping" 2nd ed. University of Amsterdam, Amsterdam. 2009.

[12] J. K., Yamamoto,. "An alternative measure of the reliability of ordinary kriging estimates". Mathematical Geology, 2000, 32(4): 489–509.

[13] D., Mcgrath, and C., Zhang,. "Spatial distribution of soil organic carbon concentrations in grassland of Ireland". Applied Geochemistry, 2003, 18: 1629–1639.

[14] P. A.. Burrough, "Principles of Geographical Information Systems for Land Resources Assessment". New York: Oxford University Press. 1986.

[15] M. A. Oliver, "Kriging: A Method of Interpolation for Geographical Information Systems". International Journal of Geographic Information Systems, 1990, 4: 313–332.

[16] C. J. G., Santhi, A. J. R., Williams, W. A., Dugas, R., Srinivasan and L. M., Huack " Validation of the SWAT model on a large river basin with point and nonpoint sources". Journal of American Water Resources Association, 2001, 37 (5): 1169-1188.

[17] M. W., Van Liew, T. L., Veith, D. D., Bosch, J. G., Arnold, " Suitability of SWAT for conservation effects assessment project: A comparison on USDA-ARS experimental watersheds". Journal of Hydrologic Engineering, 2007, 12(2): 173-189.