Accuracy of Various Interpolation Methods in Estimating Rainfall Values to Select the Most Optimal Algorithm (Case Study: Kurdistan Province)

Arash Zandkarimi
MA in Remote Sensing, University of Tabriz, Tabriz, Iran
Davood Mokhtari⁎
Associate professor of Geomorphology, University of Tabriz, Tabriz, Iran

Received: 05/06/2015 Accepted: 09/01/2018

Extended abstract

Introduction

Continuous spatial data of environmental variables are often required for environmental sciences and management. However, information for environmental variables is usually collected by point sampling. Thus, the methods generating such spatially continuous data from the point samples become essential tools for many environmental analyses. Spatial interpolation is the procedure of estimating the value of un-sampled points using existing observations. The methods for spatial interpolation can be classified into two main categories as deterministic and geostatistical. Deterministic interpolation techniques including Inverse Distance Weighting (IDW), Radial Basis Function (RBF), and so on calculate the values of un-sampled areas based on the known values of the sampled points and create surfaces from measured points. However, Geostatistical interpolation techniques, e.g. Kriging use statistical properties of the measured points to quantify the spatial autocorrelation among the measured points and account for the spatial configuration of the sample points around the estimation location.

Materials and methods

In this study, to assess the accuracy of the various interpolation methods to estimate the rainfall distribution of Kurdistan Province, we have used data from rain gauge stations, and synoptic and climatology data. After reviewing the statistical situation of the stations, statistical period of 2001-2013 has been selected. Among all stations in the basin, the stations which had 12 years of full or recyclable statistics until 2013 were selected for the study. It must be noted that the selection of the stations was according to their statistics rebuilt by the application of the highest correlation with the adjacent stations. Finally, normality of data quality and data sets were recorded and evaluated using Komogorov-Smirnov and Chi square X2 statistical tests. In addition, we have used the digital elevation model data collected by the SRTM satellite sensors with spatial resolution of 30 m. We have also employed analytic functions of the ArcGIS 10.2.2, Surfer 11, and IBM SPSS Statistics 22. After reviewing data on existing stations (77 Rain gauge

⁎E-mail: d_mokhtari@tabrizu.ac.ir Tel: +989144156714

stations, 22 synoptic and climatology stations of the Meteorology Organization, and 76 Rain gauge stations and evaporation stations of the Department of Energy), we have rebuilt the missing rainfall data and selected the minimum length of the statistical common period for less than 10 years. The normality of data has been evaluated to select 145 stations for the analysis of the interpolation methods and choosing the best method. The methods used in this study were IDW, Spline (with Tension, Thin Plate and Completely Regular functions), Ordinary Kriging (with Circular, Spherical, Exponential, and Gussian functions), and Universal Kriging (with Rational Quadratic and Liner functions).

Results and discussion

In order to conduct this statistical analysis, the data should be normal. Therefore, the average annual rainfall values of the Kurdistan Province were tested for normality. In this research, Komogorov - Smirnov and Chi square X2 tests in SPSS have revealed that the data have the normal distribution. In order to analyze the accuracy of the various interpolation methods, the models were implemented using the ArcGIS application. By applying each of these models on rainfall data, the maps were obtained. In order to evaluate and determine the most optimal model, the validity and the accuracy of the maps were evaluated. As it mentioned in the previous section, the more the Mean Absolute Error (MAE) and the closer Mean Bias Error (MBE) to zero, the accuracy of the model is higher. On the other hand, the less the Root Mean Square Error (RMSE) and the higher the correlation coefficient (R2), the less is the model error is. The results also show the error rate of implementation of the interpolation methods. Based on the findings of the study, the lowest error is related to the Ordinary Kriging Method with the circular function; and after that, it is related to the General Kriging Method with the Quadratic Variogram. In general, the Kriging method provides results with higher accuracy than those of other methods.

Conclusion

In this study, the models of the deterministic and geostatistical methods were compared with each other to find the most suitable spatial interpolation method in Kurdistan Province. To compare the actual results, the same conditions were used to assess the accuracy. Then, the most important methods of the validity were extracted and identified: Mean Absolute Error (MAE), Mean Bias Error (MBE), Root Mean Square Error (RMSE) and correlation coefficient (R2). Ordinary Kriging Method of interpolation with the circular function had the highest accuracy compared with the other methods. One of the most important factors to achieve high accuracy in this method is its ability to depict the non-bias linear estimation. However, other methods, especially the Universal Kriging with Quadratic function, due to the use of local procedures offers an acceptable accuracy.

Keywords: accuracy assessment, Kordestan Province, estimated distribution of precipitation, interpolation.