Journal of Hydraulic and Water Engineering (JHWE), Vol. 1, NO. 1, 89-98

Shahrood University of Technology

Journal of Hydraulic and Water Engineering (JHWE)

Journal homepage: https://jhwe.shahroodut.ac.ir/

Iranian Hydraulic Association (IHA)

Evaluation of Different Spatial Interpolation Methods for IMERG Precipitation Zoning in Neyshabur Basin

M. Baseri [1] and E. Mahjoobi [2,*]

[1] Ph.D. student of Water Resource Management and Engineering, Faculty of Civil Engineering, Shahrood University of Technology, Shahrood, Iran.
[2] Department of Water and Environmental Engineering, Faculty of Civil Engineering, Shahrood University of Technology, Shahrood, Iran.

Article Info

Article history:

Received: 29 April 2023
Received in revised form: 9 July 2023
Accepted: 15 September 2023
Published online: 17 September 2023

DOI: 10.22044/JHWE.2023.13019.1012

Keywords:
Precipitation
Kriging
IDW
Spline
Interpolation
IMERG

Abstract

Direct measurement of continuous spatial data is almost impossible due to the limitations in data collection. To address this issue, many methods have been proposed for data interpolation. Determining the most suitable interpolation method to describe the amount of precipitation at any point on a certain time scale and in a specific region has been the concern of many researchers. In this study, different spatial interpolation methods were evaluated for predicting the rainfall zoning of IMERG satellite products in Iran's Neyshabur basin for a statistical period of 17 years. For this purpose, 11 interpolation methods including Inverse Distance Weighting (IDW), Spline, Ordinary Kriging (OK), and Universal Kriging (UK) with different techniques were used. Regarding the results of statistical indices for the UK method with quadratic drift (with Bias = -2.85, CC = 0.98, KGE = 0.97, and RMSE = 8.1), this method was the best interpolation technique for zoning rainfall in the Neyshabur basin. Finally, the choice of interpolation method depends on the spatial scale, density of stations, and variability of the data.

1. Introduction

One of the most significant meteorological factors is precipitation. Reliable rainfall data is a crucial input for hydrological modeling, forecasting of the extreme rainfall events (such as drought and flood), and evaluation of the quantity and quality of water resources (Shadeed *et al*., 2022). Precipitation is also used in a variety of fields such as water resource management, climate change, water balance estimation, etc. Therefore, accurately estimating local rainfall from point monitoring stations is crucial (Taher, 2020). Precipitation is measured pointwise at rain gauge stations. Because of the dispersion of stations and significant variations in rainfall in arid and semi-arid areas, it is crucial to take station location into account when making estimates (Esmaeili and Vafakhah, 2022). Establishing sufficient stations in each region is always associated with limitations due to economic and operational issues. Mountainous and arid

* *Corresponding author:* emahjoobi@shahroodut.ac.ir, Tel:+989129348103

regions are more affected by these limitations (Rata *et al*., 2020). Therefore, estimating precipitation in the areas between the stations and knowing its spatial distribution should be considered in various projects and research works due to the impossibility of complete coverage of the area by rain gauge stations (Golian *et al*., 2018).

Direct measurement of continuous spatial data is almost impossible due to the limitations in data collection. Faced with these problems, many methods have been proposed for data interpolation. Determining the most suitable interpolation method to describe the amount of precipitation at any point on a certain time scale and in a specific region has been the concern of many researchers (Adhikary *et al*., 2017; Vieux, 2001). An interpolation process is required to determine a new value based on a given set of data over an area. This work is often done by determining the interpolation equations by weighting the studied stations (Hassim *et al*., 2020; Seyyed Nezhad Golkhatmi *et al*., 2013).

When measured values are not available, interpolation is a method or mathematical function that estimates the values (Keblouti *et al*., 2012). The spatial interpolation method simulates the predicted values of unknowable points or regions based on the measured values of sampling points (Berger *et al*., 2001). In other words, it is possible to convert point data in a region into continuous data using spatial interpolation methods by having the point data of a variable such as precipitation, temperature, groundwater level, etc. (Zandkarimi and Mokhtari, 2018).

Several methods can be used to calculate the spatial distribution of meteorological variables. Interpolation methods can be broadly categorized into two groups: geostatistical (Kriging, Cokriging), and deterministic (Inverse Distance Weighting, IDW; Radial Basis Functions, RBF; Global Polynomial, GP; Local Polynomial, LP) (Shadeed *et al*., 2022). The fundamental idea behind deterministic methods is that as the distance from the prediction location increases, the relative importance of the observed value reduces. Geostatistical methods, which are based on the theory of regionalized variables, supply a set of statistical tools to incorporate the spatial correlation of the observations in data processing (Goovaerts, 2000; Mendez and Calvo-Valverde, 2016). Different geostatistical methods show different accuracy depending on the type of variable.

Even though there are plenty of interpolation methods that can be used to estimate weather data, selecting the most effective technique can be difficult (Zhang *et al*., 2016). In this situation, the ArcGIS software can be useful because it makes it possible for users to compare different interpolation methods, including deterministic and geostatistical methods (Amini *et al*., 2019; Xiao *et al*., 2016). As part of spatial analysis in GIS, spatial interpolation includes the software and user attempting to estimate reasonably unknown or unmeasured values in continuous data (Moradi, 2020; Taher, 2020).

Many researchers have evaluated various methods for the interpolation of meteorological data. For example, Keblouti *et al*. (2012) used IDW, Spline, and ordinary kriging methods to study the spatial interpolation of annual rainfall over 29 years in Annaba City, Algeria. They showed that the ordinary kriging and spline are the less efficient interpolation methods, and the IDW is the most accurate method for describing the distribution of rainfall.

Fadavi (2015) investigated several interpolation methods including IDW, kriging, cokriging, kriging-regression, multiple regression, and spline in the regionalization of daily minimum temperature data from 30 meteorological stations in 1992 and 54 meteorological stations in 2007 in Isfahan, Iran. They found that multiple regression and kriging regression were the most accurate methods.

Zandkarimi and Mokhtari (2018) evaluated the performance of 10 different interpolation methods in the Kurdistan Province rainfall estimation based on the statistical indices of root mean square error, bias, coefficient of determination (R^2), and mean absolute error. Additionally, they demonstrated that the ordinary kriging method is more reliable and appropriate for mountainous areas.

Amini *et al*. (2019) assessed and compared six spatial interpolation methods including IDW, natural neighbor, regularized spline, tension spline, Ordinary Kriging (OK), and Universal Kriging (UK) in the Zayandeh-Rud River basin over 45 years. According to the findings, the natural neighbor was the most effective method for mapping monthly precipitation and maximum temperature, while ordinary kriging with a Gaussian semi-variogram was the best method for mapping minimum temperature.

Esmaeili and Vafakhah (2022) examined interpolation methods including deterministic methods (IDW, RBF, GP, and LP) and geo-statistical methods (kriging and cokriging) using GIS to estimate monthly and annual rainfall in the Namak lake watershed. They used data from 110 meteorological stations over a statistical period of 30 years. The results showed that according to Root Mean Square Error (RMSE), the cokriging method has the best annual rainfall estimation. RBF, cokriging, and kriging methods had also the best estimation for monthly rainfall data.

The past research can be summarized by saying that interpolation methods exhibit different performances in various regions. Also so far, these methods have not been applied to satellite products in Iran. Therefore, this study evaluates different interpolation methods, and chooses the best method for rainfall zoning of IMERG satellite products in the Neyshabur basin, which is one of the important plains of the agricultural hub in Khorasan Razavi Province. It should be noted that the interpolation methods were selected based on the methods available in the Spatial Analyst>Interpolation toolbox of the GIS software. Four statistical indicators including bias, Correlation Coefficient (CC), Kling-Gupta Efficiency (KGE), and RMSE were used.

2. Material and Methods

2.1. Studied area

Neyshabur basin is located in Khorasan Razavi Province in Iran with a maximum height of 3316 meters and a minimum height of 1049 meters above sea level. The area of this basin is about 9500 km^2. About half of the basin area is used for agriculture, dry farming, and orchards. The average annual rainfall of the basin is about 230 mm and the average annual temperature is about 13 degrees Celsius. Figure 1 shows the location of this basin.

2.2. Datasets

In this study, IMERG satellite precipitation products (in short: IMERG products) corresponding to the locations of 30 rain gauge stations inside and outside of the Neyshabur basin were used. The GPM-IMERG V06 Final Run daily precipitation product at the location of the meteorological station was obtained from http://apdrc.soest.hawaii.edu/. The specifications of the rain gauges in the study area can be seen in Table 1.

Precipitation data was collected monthly for a statistical period of 17 years, from the beginning of 2001 to the end of 2017, and then converted into annual data. Finally, the long-term mean annual precipitation data of IMERG products was used to evaluate different interpolation methods. To compare interpolation methods, several different methods, including IDW, Spline, and Kriging, were evaluated and compared with each other. In the following, the mentioned interpolation methods are briefly introduced.

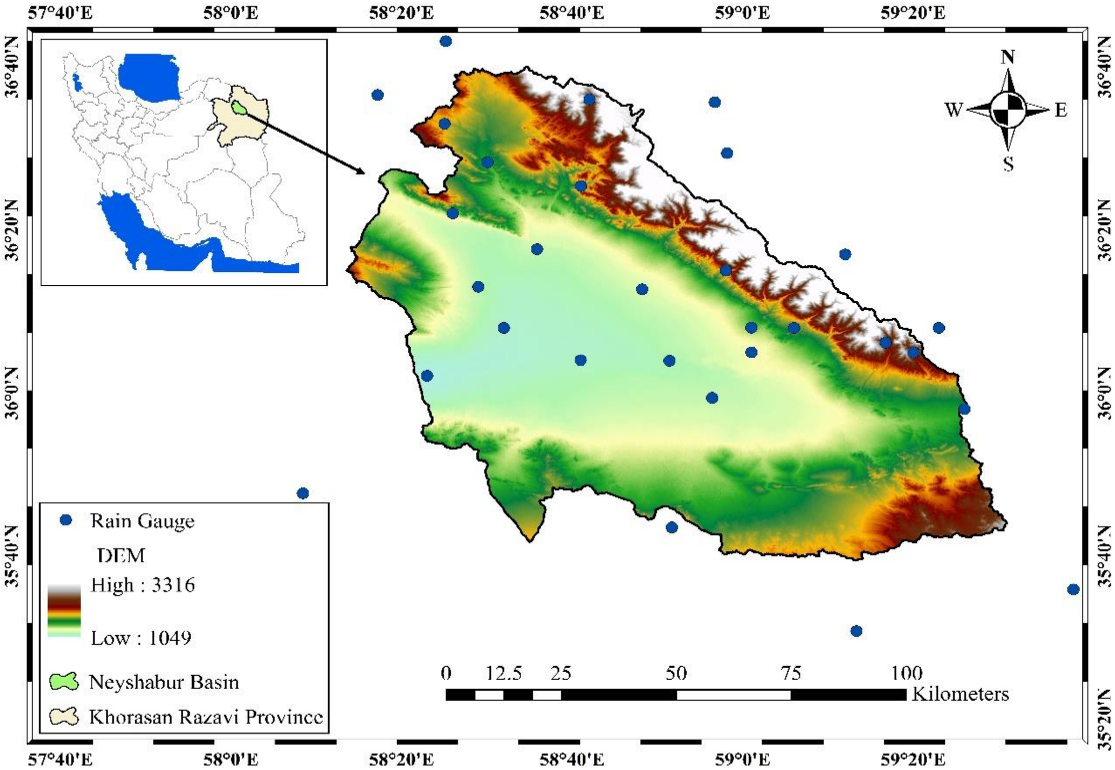

Figure 1. Studied area of Neyshabur basin and location of rain gauges.

Table 1. Specifications of the rain gauges.

Code	Name	Longitude (decimal degrees)	Latitude (decimal degrees)	Code	Name	Longitude (decimal degrees)	Latitude (decimal degrees)
P01	Emam_Taghi	59.433	35.967	P16	Edare_Neyshabur	58.803	36.212
P02	Pivehjen	59.333	36.083	P17	Hoseinabad_Jangal	58.383	36.035
P03	Darrood	59.1	36.133	P18	Barzanun	58.286	36.609
P04	Soleymani	58.433	36.367	P19	Roohabad	58.856	36.066
P05	Bojhan	58.967	36.25	P20	Sagh_Beyg	58.419	36.719
P06	Shoorgasht	58.483	36.217	P21	Dizabad_Olia	59.279	36.103
P07	Khayyam	59.017	36.083	P22	Karkhane_Ghand_Neyshabur	58.598	36.294
P08	Baghi	58.7	36.6	P23	Abghad_Farizi	58.969	36.490
P09	Kang	59.2	36.283	P24	Taghoon	58.684	36.423
P10	Hematabad_Zamani	58.533	36.133	P25	Karizno_Fariman	59.646	35.598
P11	Akbarabade_Peskal	59.0167	36.1333	P26	Sangerd_Sabzevar	58.140	35.795
P12	Eshghabad_Neyshabur	58.683	36.067	P27	Dahane_Akhlamad	58.945	36.594
P13	Sarchah	58.417	36.55	P28	Ghonchi	59.222	35.513
P14	Aghanj	59.383	36.133	P29	Zarandeh	58.501	36.471
P15	Esehaghabad	58.94	35.99	P30	Talkh_Bakhsh	58.861	35.725

2.2.1. IDW method

One of the most common methods for spatial interpolation is IDW. It is a deterministic interpolation method that calculates unsampled location values using a weighted average of the nearby sampled data points (Zhao *et al*., 2022). Weights of the sample points are determined based on their distance from the interpolation point so that the points closer to the unsampled points have a greater weight and impact on interpolation calculations. In accordance with how far away the unsampled point is from the measured point, the weight parameter regulates how the measured point values affect the interpolated value (Erdogan, 2009).

IDW cannot generate values greater than the maximum value because the influence on observations declines with increasing distance from a known point (Karwariya *et al*., 2021). The formula for IDW is:

$$\hat{Z} = \left(\sum_{i=1}^{n} \frac{Z_i}{d_i^p}\right) \Big/ \left(\sum_{i=1}^{n} \frac{1}{d_i^p}\right) \quad (1)$$

where $\hat{Z}$ is the estimated value at the interpolated point; n is the number of the surrounding sampled points used in the interpolation calculation; Z_i is the IMERG Products value observations at the i^{th} rain gauge station; d_i is the distance between the target and each of the i^{th} rain gauge station; p is an exponential value set between 0.5 and 3 (Shi *et al*., 2022).

2.2.2. Spline method

Spline interpolation is a smoothing method that fits a curve to the sampled data points using polynomials or other functions. The method is predicated on two fundamental premises:

1) The data points should be passed through by the interpolating function;

2) It should be as smooth as possible.

There are two various types of the spline: tension and regularized. The regularized type of splines, among the available types, boosts the analytical characteristics of splines by incorporating and employing third- and higher-order derivatives (Amini *et al*., 2019; Cao *et al*., 2009). In this paper, the regularized spline type is used. A spline, as opposed to IDW, can estimate surface values above and below the observed values. The formula is as follows (Adhikary *et al*., 2017):

$$Z(x) = \sum_{i=1}^{m} a_i f_i(x) + \sum_{j=1}^{n} b_j w(d_j) \quad (2)$$

where $w(d_j)$ illustrates the radial basis purposes and d_j the distance from the rain gauge station to prediction point x, $f_i(x)$ is a trend function, *a* member of a basis for the space of polynomials of degree $< m$. The coefficients and b_j are obtained by solving the system.

2.2.2. Kriging method

Kriging is one of the most prevalent and widely used geostatistical methods. By giving observed values weights based on the spatial correlation between the observed and unknown locations, this method calculates the unknown values of the variable of interest. It should be noted that kriging uses variogram models to measure the spatial dependence between data points which is one of its distinctive characteristics. Variograms are used to model the spatial autocorrelation between data points and to determine the contributions of neighboring data points to the estimation of unknown values (Fadavi, 2015; Khaddari *et al*., 2022; Moradi, 2020; Zandkarimi and Mokhtari, 2018). There are several methods to calculate values based on Kriging. In this research work, two methods, Ordinary Kriging (OK) (including spherical, circular, exponential, Gaussian, and linear), and Universal Kriging (UK) (including quadratic and linear) have been used.

The interpolation and zoning process of long-term mean annual precipitation in the Neyshabur Basin is done with the utilization of ArcGIS. To do so, 70% of the stations (21 stations) were considered as the train data, and 30% of them (9 stations) were considered the test data. Performing the mentioned interpolation methods, the interpolated value was evaluated using the training data in the location of the test data, using the statistical indicators presented in Table 2. Finally, the results obtained from these methods were compared to each other's.

Table 2. Statistical indices used to assess the performance of interpolation methods.

Index (Symbology)	Equation*	Min, Max, Optimal	Unit
Bias	$Bias = \frac{\sum_{i=1}^{n}(S_i - O_i)}{n}$	-∞, ∞, 0	mm
Correlation Coefficient (CC)	$CC = \frac{\sum_{i=1}^{n}(S_i - \mu_s)\ (O_i - \mu_o)}{\sqrt{\sum_{=1}^{n}(S_i - \mu_s)^2}\ \sqrt{\sum_{i}^{1}(O_i - \mu_o)^2}}$	-1, 1, 1	-
Kling-Gupta Efficiency (KGE)	$KGE = 1 - \sqrt{(CC-1)^2 + (\frac{\mu_s}{\mu_o} - 1)^2 + (\frac{\sigma_s/\mu_s}{\sigma_o/\mu_o} - 1)^2}$	-∞, 1, 1	-
Root Mean Square Error (RMSE)	$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(S_i - O_i)^2}{n}}$	0, ∞, 0	mm

* where S_i is the value of interpolated for the ith station, O_i is the value of long-term mean annual IMERG products for the ith station, n is the number of rain gauge stations, μ_s is the average value of interpolated for n stations, μ_o is the average value of long-term mean annual IMERG products for n stations, σ_s is the interpolated standard deviation value for n stations, and σ_o is the standard deviation value of long-term mean annual IMERG products for n stations.

3. Results and Discussion

The spatial distribution of the Neyshabur basin rainfall was developed in GIS software based on the 11 interpolation methods as shown in Figure 2. Different rainfall spatial distribution patterns were obtained depending on which of the 11 interpolation methods was used. Generally, the Northeast-Southwest rainfall gradient exists in all the maps: high rainfall in the Northeast (more than 350 mm) and low rainfall in the Southwest (less than 170 mm) of the Neyshabur basin.

Four indices including Bias, CC, KGE, and RMSE were used to evaluate interpolation methods. The results are shown in Figure 3 to Figure 6. As can be seen from Figure 3, the lowest value of bias corresponds to Ordinary Kriging (Gaussian), Spline, and Universal Kriging (Quadratic) methods. Bias is defined as the average difference between predicted estimates and IMERG products. Positive and negative values of bias indicate, respectively, overestimation and underestimation of rainfall amount.

According to Figure 4, in general, all the interpolation methods have a high correlation of 93% with IMERG products. The UK method with a quadratic drift and the OK method with linear semi-variogram have the highest and lowest correlation with IMERG products.

KGE is a composite index that includes the concepts of correlation, deviation, and dispersion and expresses the efficiency of the model in estimating observational data. KGE index values are shown in Figure 5. Except for two IDW with power 1 and UK (linear) methods, the rest of the methods have KGE values above 0.7, which shows the proper efficiency of these methods in interpolating precipitation in the Neyshabur basin.

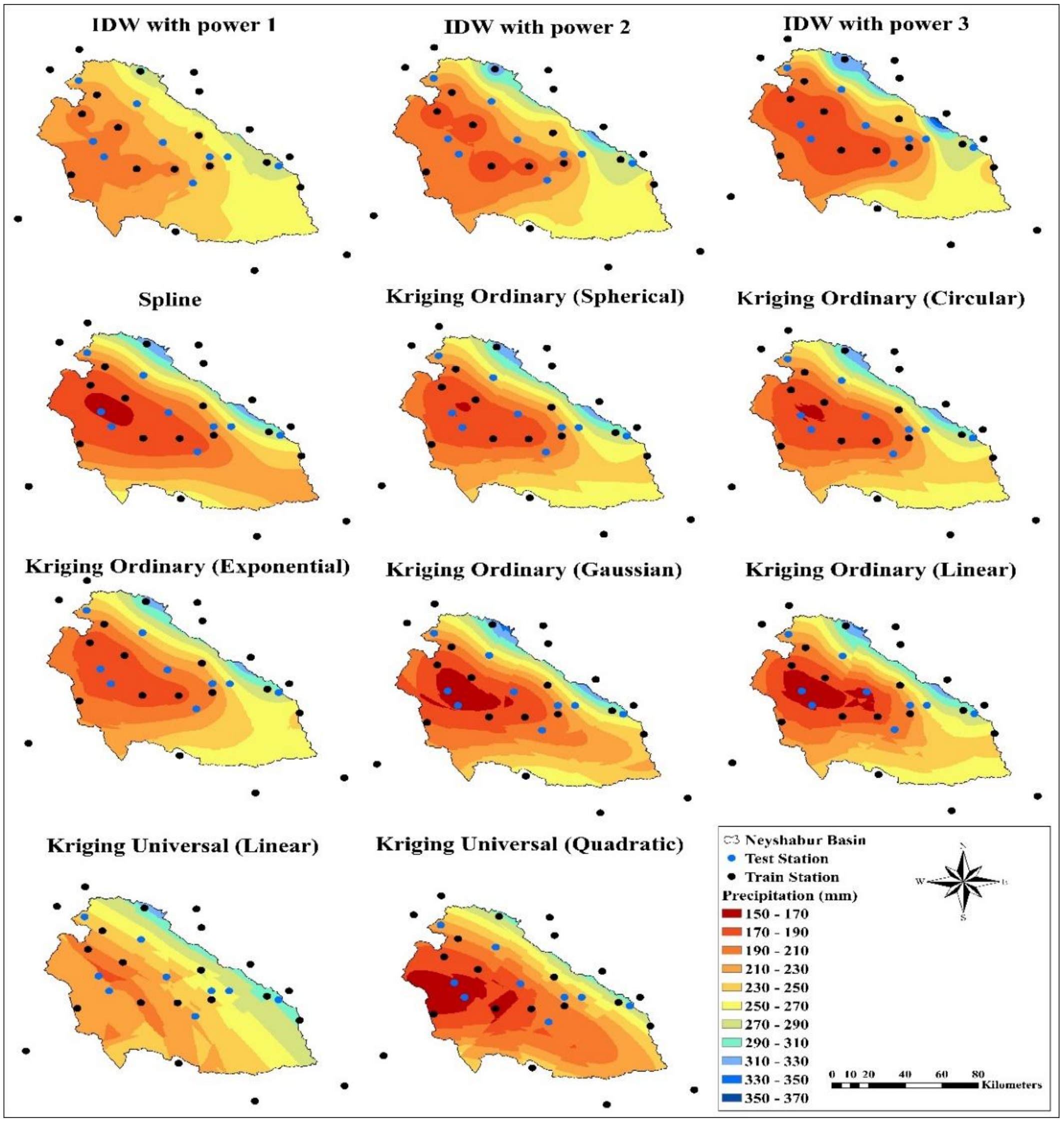

Figure 2. Precipitation zoning maps obtained by the different interpolation methods.

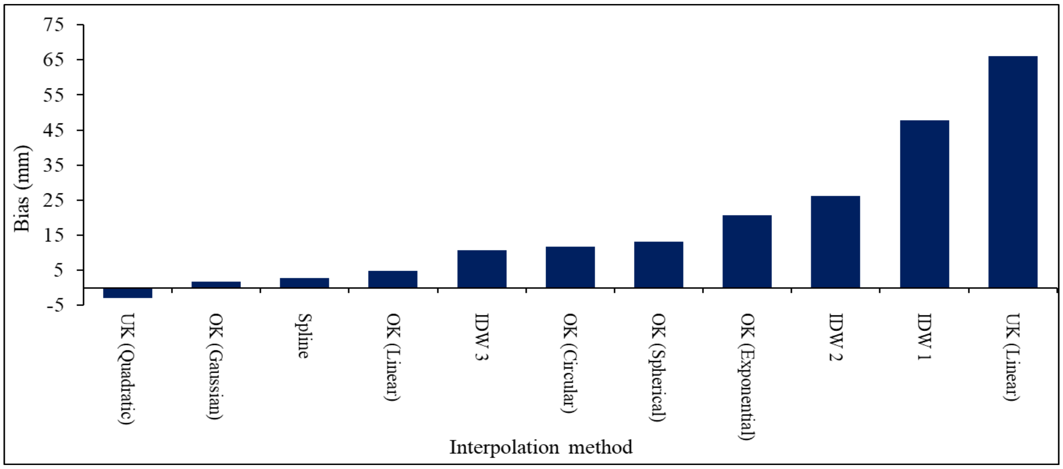

Figure 3. Evaluation of interpolation methods accuracy using BIAS index.

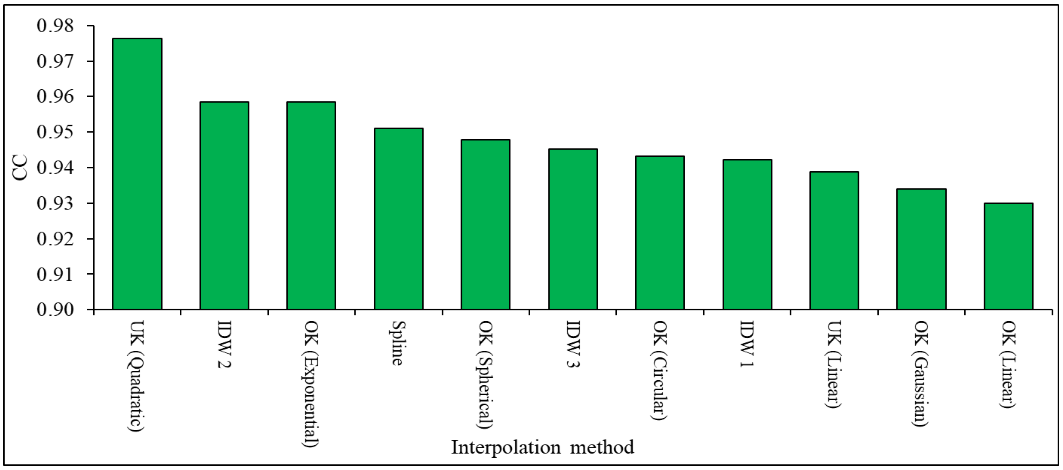

Figure 4. Evaluation of interpolation methods accuracy using CC index.

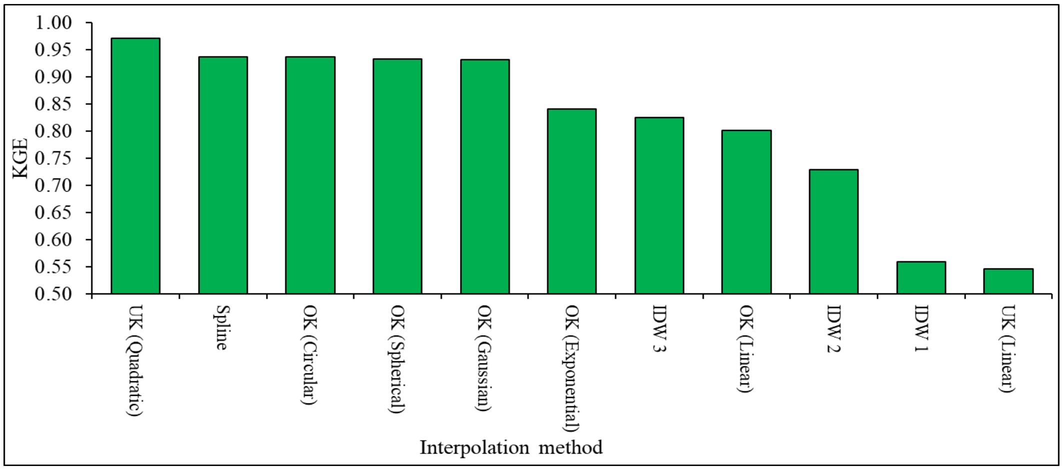

Figure 5. Evaluation of interpolation methods accuracy using KGE index

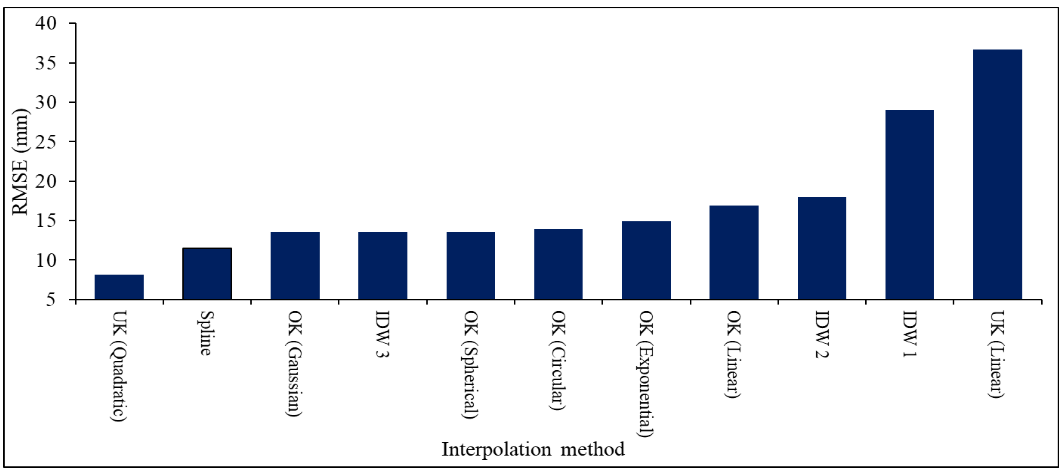

Figure 6. Evaluation of interpolation methods accuracy using RMSE index.

The lower the RMSE value, the lower the model error in estimating and interpolating precipitation values. Based on this, the lowest error value is related to the UK (quadratic) method with a value of 8.1 mm (Figure 6).

4. Conclusions

The interpolation toolbox in ArcGIS is the foundation for this study, which aimed to assess the performance of various spatial interpolation methods including IDW, spline, OK, and UK using various techniques with IMERG products over 17 years in Iran's Neyshabur basin. Depending on interpolation methods, different spatial precipitation patterns were obtained (Figure 2). According to the bias index, all interpolation methods had overestimations except the UK method. The spline method and various OK method techniques indicate acceptable performance in estimating the precipitation values of IMERG products. On the other hand, IDW with power 1 and UK (linear) methods had the worst results in all indices. The majority of the 11 methods were effective at predicting IMERG products values based on the results of the Bias, CC, KGE, and RMSE indices. However, the UK method with quadratic drift is the best interpolation technique for zoning rainfall in the Neyshabur basin. Similar to Shadeed *et al*. (2022) the UK method has a higher correlation coefficient than the other methods which makes this method the most optimal one for mapping rainfall distribution. It is important to note that these results are only applicable to rainfall interpolation in the Neyshabur basin.

While there are many methods available, the choice of interpolation method ultimately depends on the spatial scale, density of stations, and variability of the data.

Data Availability

The data used to support the findings of this study is available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this paper.

Acknowledgement

We gratefully acknowledge the support and funding provided by the Shahrood University of Technology to conduct this research work.

References

Adhikary, S.K., Muttil, N., and Yilmaz, A.G., 2017. Cokriging for enhanced spatial interpolation of rainfall in two Australian catchments. Hydrological processes, 31(12): 2143-2161.

Amini, M.A., Torkan, G., Eslamian, S., Zareian, M.J., and Adamowski, J.F., 2019. Analysis of deterministic and geostatistical interpolation techniques for mapping meteorological variables at large watershed scales. Acta Geophysica, 67(1): 191-203.

Berger, J.O., De Oliveira, V., and Sansó, B., 2001. Objective Bayesian analysis of spatially correlated data. Journal of the American Statistical Association, 96(456): 1361-1374.

Cao, W., Hu, J., and Yu, X., 2009. A study on temperature interpolation methods based on GIS, 2009 17th International Conference on Geoinformatics. IEEE, pp. 1-5.

Erdogan, S., 2009. A comparision of interpolation methods for producing digital elevation models at the field scale. Earth surface processes and landforms, 34(3): 366-376.

Esmaeili, F. and Vafakhah, M., 2022. A Comparative Performance Evaluation of Interpolation Methods for Estimating Annual and Monthly Rainfall in the Namak Lake Watershed, Iran. Extension and Development of Watershed Management, 10(36): 1-10.

Fadavi, G., 2015. Comparison of different regional estimation methods for daily minimum temperature (A case study of Isfahan province). Journal of Agricultural Meteorology, 3(2): -.

Golian, S., Asadi Oskouei, E., and Lopez-Baeza, E., 2018. Assessment of interpolation methods for annual and seasonal precipitation in Mashhad plain. Nivar, 42(100-101): 11-20.

Goovaerts, P., 2000. Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. Journal of hydrology, 228(1-2): 113-129.

Hassim, M. *et al*., 2020. Comparison of rainfall interpolation methods in Langat River Basin, IOP Conference Series: Earth and Environmental Science. IOP Publishing, pp. 012018.

Karwariya, S., Dey, P., Bhogal, N.S., Kanga, S., and Singh, S.K., 2021. A Comparative Study of Interpolation Methods for Mapping Soil Properties: A Case Study of Eastern Part of Madhya Pradesh, India. Recent Technologies for Disaster Management and Risk Reduction: Sustainable Community Resilience & Responses: 431-449.

Keblouti, M., Ouerdachi, L., and Boutaghane, H., 2012. Spatial interpolation of annual precipitation in Annaba-Algeria-comparison and evaluation of methods. Energy Procedia, 18: 468-475.

Khaddari, A. et al., 2022. Evaluation of Precipitation Spatial Interpolation Techniques using GIS for Better Prevention of Extreme Events: Case of the Assaka Watershed (Southern Morocco).

Mendez, M. and Calvo-Valverde, L., 2016. Assessing the performance of several rainfall interpolation methods as evaluated by a conceptual hydrological model. Procedia Engineering, 154: 1050-1057.

Moradi, H., 2020. Evaluation of Different Interpolation Methods in Climatic Zones of Ilam Province. Geography and Human Relationships, 2(4): 1-15.

Rata, M., Douaoui, A., Larid, M., and Douaik, A., 2020. Comparison of geostatistical interpolation methods to map annual rainfall in the Chéliff watershed, Algeria. Theoretical and Applied Climatology, 141: 1009-1024.

Seyyed Nezhad Golkhatmi, N., Sanaeinejad, H., Ghahraman, B., and Rezaee Pazhand, H., 2013. Daily rainfall interpolation of Mashhad Drainage basin. Journal of Climate Research, 1392(15): 17-30.

Shadeed, S., Jayyousi, A., Khader, A., Chwala, C., and Kunstmann, H., 2022. Comparative analysis of interpolation methods for rainfall mapping in the Faria catchment, Palestine. An-Najah University Journal for Research-A (Natural Sciences), 36(1): 1-20.

Shi, J. et al., 2022. Comparison of the Performance of IMERG Products and Interpolation-Based Precipitation Estimates in the Middle Reaches of Yellow River Basin. Water, 14(9): 1503.

Taher, L.S.B., 2020. Evaluation of Geostatistical Interpolation Methods for Rainfall data Estimation in Libya. Albahit J Appl Sci, 1(1): 54-9.

Vieux, B.E., 2001. Distributed hydrologic modeling using GIS. Springer.

Xiao, Y. et al., 2016. Geostatistical interpolation model selection based on ArcGIS and spatio-temporal variability analysis of groundwater level in piedmont plains, northwest China. SpringerPlus, 5(1): 1-15.

Zandkarimi, A. and Mokhtari, D., 2018. Accuracy of Various Interpolation Methods in Estimating Rainfall Values to Select the Most Optimal Algorithm (Case Study: Kurdistan Province). Physical Geography Research Quarterly, 50(2): 323-338.

Zhang, X., Lu, X., and Wang, X., 2016. Comparison of Spatial Interpolation Methods Based on Rain Gauges for Annual Precipitation on the Tibetan Plateau. Polish journal of environmental studies, 25(3).

Zhao, W. et *al*., 2022. Comparison and correction of IDW based wind speed interpolation methods in urbanized Shenzhen. Frontiers of Earth Science: 1-11.