ISSN 1533-9211

DOI 10.17605/OSF.IO/8Z4TD

COMPARATIVE ANALYSIS OF DIFFERENT TECHNIQUES FOR SPATIAL INTERPOLATION OF RAINFALL DATASETS IN DUHOK GOVERNORATE

MARWAN BASHEER ISMAIL GOVAY

Soil and Water Dept., College of Agricultural Engineering Sciences, University of Duhok, Iraqi Kurdistan Region. Email: marwan.basheer@uod.ac

TARIQ HAMA KARIM

Department of Survey and Geomatics Engineering, Faculty of Engineering, Tishk International University-Erbil, Kurdistan Region, Iraq. Email: solavtariq@yahoo.com

Abstract

The spatial distribution of rainfall plays a key role in water management and hydrological modeling. This study employed rainfall datasets of different scales recorded at 22 meteorological stations distributed over Duhok governorate with a time span from 1998 to 2020. Four spatial interpolation methods were employed to transform the discrete values into a continuous spatial pattern over Duhok Governorate, Kurdistan Region, and Iraq. The interpolation methods evaluated were: Inverse Distance Weighted (IDW), Ordinary Kriging (OK), Universal Kriging (UK) and Spline. The reliability of the prediction techniques was examined via leave-one-out cross-validation besides giving ranks using TOPSIS algorithm. Several performance indicators were employed as criteria to evaluate the applied interpolation techniques. The results indicated that the IDW offered the best performance in most cases, while the spline method offered the poorest performance. In general, differences between these first three methods (IDW, OK and UK) are low but substantially different from that the spline method. Overall, the mean absolute percentage of error of estimation was 16% or less and the degree of prediction accuracy tended to increase with a decrease in time scale. Further, the interpolation methods provide similar spatial distributions of rainfall.

Keywords: Spatial distribution, interpolation methods, rainfall estimation, comparative analysis

1. INTRODUCTION

The availability of reliable rainfall data is essential for the design and management of water resources systems and for most hydrological analyses (Di Piazza et al., 2011). Unfortunately, measurement of this variable often suffers from systematic and random errors and gaps (Vieux, 2001). Among these deficiencies, the missing data is probably the most important one (Di Piazza et al., 2011). It is commendable to mention that rain gauge networks offers only point estimates, and under most cases, their distribution is uneven and their number is quite limited (Rata et al., 2020; Mirás-Avalos et al., 2007). Thus, there is the need to estimate rainfall values at unrecorded locations using data of the surrounding sites to get a continuous surface by spatial interpolation of the recorded values (Goovaerts, 2000). The main problems with precipitation data are its spare distribution in space and its discontinuity in time, causing difficulty in the development of consistent climatology in a given area. Under such situations, the spatial interpolation techniques become a good substitute for developing continuous spatial distribution based on the available recorded data. For instance, sometimes, it is difficult to

ISSN 1533-9211
DOI 10.17605/OSF.IO/8Z4TD

cover remote places like mountainous areas besides the high cost of the study in such areas (Antal et al., 2021). Both Rainfall and stream flow may contain missing values which ascribed to a variety of causes such as, instrumental failures, bad weather or human error during data recording (Suhaila et al., 2008). Estimation of missing values becomes first priority in the data preparation process (Ismail et al., 2017) Interpolation methods are technique that can be applied to determine unknown values from data recorded at known surrounding locations. This will enable the researchers to prepare maps for areas with limited synoptic stations for spatial events (Khorsandi et al., 2012). A host interpolation technique has emerged during many years. They were grouped into two main categories, namely, deterministic and geostatistical (Ly et al. 2011). Examples of deterministic methods are radial basis function (RBF) (known as spline), local and global polynomials, and inverse distance weighted. These methods produce continuous distribution of rainfall, starting from measured points using mathematical rules to determine the degree of smoothing (Antal et al., 2021). On the other hand, geostatistical interpolation methods, like Ordinary and Universal Kriging methods employ statistical techniques for producing spatial distribution (Li and Heap 2014). Berndt and Haberlandt (2018) reported that the accuracy of estimation does not only rely on the interpolation method, but also on many other factors like the configuration of the station network, resolution of the temporal data and the spatial variability of the study variables. Suhaila et al., 2008 elucidated that the inverse distance is regarded as one of the simpler methods. It is assumed that the rainfall values at the target station are affected mainly by the nearby stations and less by the more far stations. Diodato and Ceccarelli (2005) have made a comparative analysis for several interpolation methods including IDW, OK and OCK for rainfall. They concluded that cokriging by elevation as co-variable offered the most reliable results, particularly in the mountainous regions due to geomorphic nature of rainfall. (Khorsandi et al. 2012) highlighted that RBF methods are exact interpolation techniques in which the surface must pass through the measured value. Different shapes can be obtained upon using different functions (thin plate, with tension, completely regularized, multiquadratic and inverse multiquadratic functions). This method is not suitable when there is substantial changes in the surface or if there are large changes in the surface within a small distance. (Ruelland et al., 2008). No interpolation method gives accurate results in different regions and under different conditions, each method has own specific hydrological conditions. This implies that the analysis of spatial events has usually been region-specific and generalization may not be reasonable (Chen and Guo 2017). Accordingly, local studies are essential to decide the most appropriate interpolation method (Zaghiyan et al., 2021; Ananias et al., 2021). Goovaerts (1997) pointed out that geostatistics provides a set of statistical tools for incorporating the spatial correlation of observations in data processing. Further, a number of studies have shown that geostatistics produce better estimates of precipitation than traditional methods (Drogue et al. 2002; Buytaert et al. 2006). Chutsagulprom et al. (2022) demonstrated that each spatial interpolation method has inherent advantages and disadvantages, and its selection should be based on the type of the analyzed data. They used a k-fold cross-validation for evaluating the performance of five interpolation methods found that the artificial neural networks underperformed the remaining methods for estimating monthly rainfall in Thailand. Further, they recommended the inverse exponential weighting for their study.

ISSN 1533-9211
DOI 10.17605/OSF.IO/8Z4TD

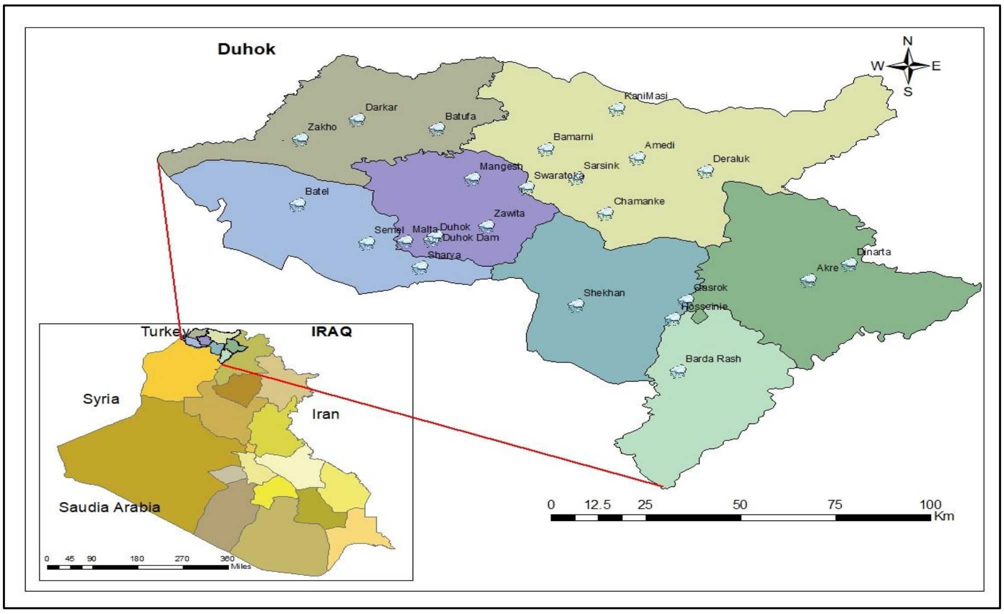
The most reliable interpolator differs from region to region (Charles et al., 2011) depending on the nature of the data and topography of the terrain as a consequence, this study was conducted to compare the performance of different univariate interpolation techniques and to identify the best interpolation method for rainfall estimation in the area under study.

2. MATERIALS AND METHODS

2.1. Description of Study area.

The study area is located within the Duhok governorate, it is bounded approximately by parallels N 36°18′42.63″ and N 37° 20 33′.55″ and meridians E 42° 20′25.36″ and E 44°17′40.5″ (Fig.1). The altitude ranges between 243 and 2550.9 m a.m.s.l. Turkey borders this area to the north, Tigris River to the west, Mosul to the south. Hills and high mountains dominate the north and northeast, while the south and the southwest is mainly plains. The hill and mountains accounts for more than 50% of the total area. Overall, the climate of this area is of Mediterranean type, cold, rainy in winter, dry, and mild in summer. The coldest and hottest months of the year are January and August respectively. Based Koppen system of climate classification, the upper part of investigated site is classified as being of type DSa, indicating a cool wet climate in the winter and dry season in the summer. On the other hand, most of area in the middle and lower parts of the study area is classified as a temperate, dry summer, hot summer (Csa) based on the aforementioned scheme.

Fig 1: Location map showing the distribution of rainfall stations over the study area



2.2. Data Acquisition and Processing

The data employed in the current study are information that can be applied for spatial interpolation. They included mean monthly and mean annual rainfall time series recorded from

ISSN 1533-9211
DOI 10.17605/OSF.IO/8Z4TD

ground-based rain gauges of 22 meteorological stations distributed over Duhok governorate with a time span from 1998 to 2020**.** During data collection, the stations that have available data for at least 20 years were selected. Only 22 stations remained after this type of screening. Table 1. The rainfall distribution is characterized by large spatial variations. Average annual rainfall varies from as low as 251.7 mm near lower boundary to more than 1001.2 mm around Iraqi Turkish border, with most rainfall occurring in December through April. This region also experiences dry season from May until end of September in most years. Table 1 presents the summary statistics of the rainfall data. The datasets for this study were provided by the Ministry of agriculture and Water Resources and the Directorate of Duhok Meteorology. Further, the rainfall time series were subjected to homogeneity tests prior to interpolation. The homogeneity test results revealed that the majority of the monthly and annual rainfall series were labeled as useful.

Table 1: Descriptive statistic for annual rainfall recorded at the study stations during the period from 1998 to 2020

Rainfall Station	Minimum (mm)	Maximum (mm)	Mean (mm)	Std. deviation	CV	Skewness	Kurtosis
Duhok	251.74	1001.20	558.43	197.25	35.32	0.690	-0.058
Zakho	299.45	955.80	582.19	169.07	29.04	0.204	-0.303
Zawita	312.64	1284.70	783.90	259.25	33.07	0.184	-0.860
Akre	328.80	1044.90	660.80	192.18	29.08	0.216	-0.229
Mangeshk	228.86	1660.70	701.31	297.78	42.46	1.490	3.939
Bamarne	326.11	1259.00	788.42	248.61	31.53	0.092	-0.508
Amedi	418.90	1282.20	807.92	234.62	29.04	0.473	-0.506
Semel	190.09	841.10	453.73	160.87	35.46	0.616	0.048
Malta	245.24	1037.40	499.19	180.74	36.21	1.170	2.154
Batufa	303.04	1695.50	705.57	282.80	40.08	1.716	5.559
KaniMase	214.88	1397.50	767.40	291.92	38.04	-0.073	0.152
Batel	254.46	890.40	439.89	150.29	34.17	1.313	2.245
Darkar	266.47	901.20	518.95	175.38	33.80	0.519	-0.593
Deraluk	296.58	1189.54	763.46	235.45	30.84	-0.048	-0.522
Sarsing	199.65	1390.90	843.85	283.46	33.59	-0.195	0.136
Bardarash	197.88	889.50	409.66	156.45	38.19	1.560	3.059
Qasrok	243.64	998.60	520.94	187.67	36.03	0.901	0.773
Swaratoka	328	1018	675.20	180.43	26.72	0.157	0.196
Hosseinie	258	985	539.59	181.39	33.62	0.964	1.038
Dinarta	373	1351	846.16	265.38	31.36	0.176	-0.674
Chamanke	266.70	1580.00	906.03	355.59	39.25	0.250	-0.463
Shekhan	320	1061	604.61	221.80	36.68	0.943	0.123

2.3. Interpolation Schemes.

Spatial interpolation methods were employed to transform the discrete values into a continuous spatial pattern over Duhok Governorate, Kurdistan Region, and Iraq. The input rainfall data have been interpolated in ESRI ArcMap by different algorithms (conventional and geostatistical approaches) with default input parameters (Mitas and Mitasova, 1999). For

The Seybold REPORT
ISSN 1533-9211
DOI 10.17605/OSF.IO/8Z4TD

instance, in case of IDW the variable radius type was selected and in case of Kriging method spherical semivariogram model was chosen. The methods encompassed, inverse Distance

Rainfall stations	X-coordinate(utm)	Y-coordinate(utm)	Altitude	Av. annual rainfall (mm)
Duhok	321698.6	4081853	569	558.43
Zakho	294590.2	4113373	440	582.19
Zawita	333402.6	4085824	890	783.90
Akre	401265.5	4068143	636	660.80
Mangeshk	330727.6	4100522	964	701.31
Bamarne	346296.9	4109220	1220	788.42
Amedi	365557.4	4106293	1190	807.92
Semel	307984.6	4081230	456	453.73
Malta	316117	4081400	507	499.19
Batufa	323559.9	4116009	860	705.57
KaniMasi	361483.9	4121505	1280	767.40
Batel	293588.8	4093488	506	439.89
Darkar	306855.1	4119256	652	518.95
Deraluk	379826	4101976	639	763.46
Sarsink	352514.6	4100287	1048	843.85
Bardarash	373190	4040529	383	409.66
Qasrok	375132.1	4062241	415	520.94
Swaratoka	342105	4097392	1188	675.20
Hosseinie	372297	4056566	363	539.59
Dinarta	409770.9	4072952	780	846.16
Chamanke	358575	4089240	862	906.03
Shekhan	351947.3	4061185	454	604.61

Weighted (IDW), Ordinary Kriging (OK), Universal Kriging and spline. The regression model was also developed to predict annual rainfall from elevation and geographical coordinates using Microsoft Excel program version 2016. (Table 2) but it get the poorest result and overlooked.

Table 2: Geographical coordinates for the study stations along with some information on rainfall data

2.4. Cross Validation

The reliability of the prediction techniques was examined via leave-one-out cross-validation ((LOOCV) (Chilès and Delfner 2012). This type of cross validation has been accomplished by removing a measured value temporarily. The value of rainfall was then predicted for this point (station) from the rest of the data. This process was repeated for all the measured data. Thereafter, the measured value of rainfall at each station was compared with predicted value. The interpolation results were evaluated using cross-validation techniques and the degree of mismatch between the measured and interpolated values were materialized using a host of indicators as sown in the incoming subsection.

ISSN 1533-9211

DOI 10.17605/OSF.IO/8Z4TD

2.5. Efficiency Criteria for performance assessment

The following statistical indices were selected to evaluate adequately the models performance (Anees et al., 2017; **Murphy and Epstein, 1989**; Mello et al., 2013; Krause et al., 2005, Akaike, 1974):

1. $$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}\left|O_i - P_i\right|$$

2. $$\mathrm{MAPE} = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{O_i \quad P_i}{O_i}\right| \times 100$$

3. $$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}(P_i - O_i)^2}{N}}$$

4. $$\mathrm{CV} = 100 * \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{n}(O_i - P_i)^2}}{\bar{O}}$$

5. $$\mathrm{CRM} = \frac{\sum_{i=1}^{N} O_i - \sum_{i=1}^{N} P_i}{\sum_{i=1}^{n} O_i}$$

6. $$\mathrm{CRE} = \frac{\sum_{i=1}^{N}(P_i - O_i)^2}{\sum_{i=1}^{N}(O_i - \bar{O})}$$

7. $$S - \mathrm{Index} = 1 - \frac{\sum_{i=1}^{N}(O_i - \bar{O})^2}{\sum_{i=1}^{N}\left(\left|O_i - \bar{O}\right| + \left|P_i - \bar{P}\right|\right)^2}$$

Where: N = number of data points; O_i = Observed value; P_i = Model predicted value; MAE = Mean absolute error; MAPE = Mean absolute percent error; RMSE = Root mean square error; CV = Coefficient of variation; CRM = Coefficient of residual mass; CRE = compound relative error and O = Mean of observed values To further examine the accuracy of estimation by different interpolation techniques, Spearman correlation coefficient (ρ) was calculated for the measured and estimated values according to:

$$\rho = \frac{1 - 6\sum_{i=1}^{n} d_i^{\,2}}{n(n2 - 1)}$$

Where n= number of observations and d_i is the ranking difference set.

Additionally, to graphically figure out the performance of the interpolation techniques in estimating rainfall, the absolute error of distribution of the estimation of each techniques was drawn in form of Box –Whisker plot. Further, to evaluate the fitness of each technique, the observed values were plotted versus the estimated in relation to line 1:1.

3. RESULTS AND DISCUSSION

3.1. Description of the Datasets

The applied data selected for testing through cross-validation schemes encompassed monthly, seasonal and annual rainfall distributed across the Duhok governorate. The aggregated data

The Seybold REPORT
ISSN 1533-9211
DOI 10.17605/OSF.IO/8Z4TD

were based on daily rainfall data of a time span from 1998 to 2020 from 22 stations. The statistics given in Table 1 include indices for measures of tendency, dispersion and symmetry for average annual rainfall at different stations. The mean annual rainfall over the area under study ranged from a minimum of **409.66** at Bardarash to a maximum of **906.03** mm at Chamanke. The coefficient of variability (CV) ranged from 26.72% to 42.46%. It was also observed that the coefficient of variability for temporal variation fell in the high class of variability (30% < CV < 40%) (Alahacoon and Edirisinghe, 2021). Additionally, the coefficient of variation for spatial variability was 22.65 % indicating that the annual rainfall over the study area is characterized by high temporal and moderate spatial variations. Moreover, the presented data showed that the annual rainfall did not deviate substantially from normal distribution in term of skewness and skewness. Virgilio et al. (2007) and PazGonzales et al (2000) have shown that the data with a range of -1 to +1 skewness were regarded as normally distributed data.

3.2. Replicating Rainfall Magnitudes by Interpolation Methods

Table 3 exhibits the average measured annual rainfall for the meteorological stations within the study area and the interpolated values by different schemes, namely, IDW, OK, UK and spline.

Table 3: Average actual and interpolated rainfall values for average annual rainfall. Recorded and the existing stations

Stations	Annual rainfall (mm)	Estimated rainfall (mm)			
		IDW	OK	UK	Spline
Duhok	558.43	557.61	588.89	562.33	592.77
Zakho	582.19	548.17	516.11	479.84	453.24
Zawita	783.90	646.91	634.51	628.01	627.71
Akre	660.80	769.01	728.75	683.04	626.89
Mangeshk	701.31	698.72	681.88	660.55	600.99
Bamarni	788.42	766.21	751.19	739.38	747.21
Amedi	807.92	787.91	824.27	772.65	872.44
Semel	453.73	542.01	484.42	491.25	454.07
Malta	499.19	551.73	513.07	533.25	488.08
Batufa	705.57	642.58	631.69	657.88	636.39
KaniMasi	767.40	783.5	747.01	831.82	783.55
Batel	439.89	566.52	535.57	436.47	499.99
Darkar	518.95	617.96	640.07	621.32	659.37
Deraluk	763.46	781.71	791.98	810.58	760.91
Sarsing	843.85	770.66	795.94	751.28	755.07
Bardarash	409.66	560.22	560.39	483.78	595.21
Qasrok	520.94	573.92	591.79	615.64	592.80
Swaratoka	675.20	772.03	796.92	723.60	815.33
Hosseinie	539.59	546.69	562.31	559.04	482.21
Dinarta	846.16	657.69	667.78	645.19	806.53
Chamanke	906.03	761.08	755.81	695.55	830.84
Shekhan	604.61	621.35	634.93	545.42	785.59

The ANOVA test, Table 4 although different methods yielded different values revealed that there were no significant differences among the interpolation methods at ($P \leq 0.05$).

Table 4: Analysis of variance for testing the difference between the interpolation methods

Source of variation	Total sum squares	d.f	Mean sum squares	F	P-value	F-critical
Methods	10206.46	3	3402.16	0.274	0.841	2.713
Error	1027200	84	12228.58			
Total	1037406	87				

Visual examination of the results indicates that the IDW, OK and UK are the best interpolation for estimating annual rainfall in the region under study. Calculation of coefficient residual mass

ISSN 1533-9211
DOI 10.17605/OSF.IO/8Z4TD

(CRM) revealed that UK method underestimated the annual rainfall, while the remaining methods overestimated the annual rainfall. This result is in line with finding of Ibrahim and Nasser (2015), who observed that The IDW method overestimates the interpolated values while Kriging method underestimates the interpolated height values. It is commendable to mention that like the annual rainfall, the rainfall at other time scales (monthly and seasonally) exhibited similar characteristics and trends and for the sake of brevity were not shown here.

3.2. Evaluation of Interpolation Methods

To examine the efficiency of the four interpolation methods in estimating rainfall at different time scales based on leave-one-out cross validation (LOOCV), five performance criteria, namely, MAE, MAPE, RMSE, D and NSE were adopted. The results of the analysis are exhibited in Table 5. Close examination of Table 5 revealed that with no exception, the spline method had the lowest precision (lowest D and NSE) and the highest MAE, MAPE and RMSE. Unlike the spline method, the IDW has the highest precision in estimating rainfall for most of the time scales. The IDW was followed by either Universal Kriging or Ordinary Kriging. It was also discerned from the results presented in Table 5 that with two exceptions the mean absolute percentage of error from all the models was below 16% or less. Overall, the methods can be ranked in the following order of preference: IDW > OK>UK > Spline.

Table 5: The performance indicators for estimating rainfall at different time scales using different interpolation methods

Data	Criteria/Methods	IDW	OK	UK	Spline
Annual	MAE	60.75	63.51	66.25	74.96
	MAPE	9.65	9.98	10.35	12.32
	RMSE	79.94	83.04	82.16	92.97
	D	0.885	0.879	0.883	0.866
	NSE	0.648	0.664	0.654	0.556
Autumn rainfall	MAE	15.07	14.83	15.43	17.09
	MAPE	14.58	14.12	14.77	16.91
	RMSE	20.74	20.19	20.76	23.37
	D	0.762	0.807	0.790	0.696
	NSE	0.437	0.43	0.399	0.001
Winter rainfall	MAE	55.44	55.62	55.63	56.11
	MAPE	16.12	16.07	16.09	15.86
	RMSE	81.9	77.87	77.97	80.19
	D	0.624	0.700	0.699	0.705
	NSE	0.267	0.338	0.336	-0.081
Spring rainfall	MAE	30.96	31.79	33.76	43.39
	MAPE	15.34	15.31	16.12	23.21
	RMSE	43.59	42.4	44.42	58.96
	D	0.818	0.860	0.845	0.603
	NSE	0.553	0.574	0.509	-0.194
January	MAE	15.05	16.24	16.59	18.44
	MAPE	11.55	12.67	13.02	14.41
	RMSE	21.04	21.88	22.14	23.37
	D	0.768	0.836	0.834	0.827
	NSE	0.526	0.568	0.559	0.445
April	MAE	8.3	9.48	9.33	12.95
	MAPE	12.22	14.08	13.79	19.91
	RMSE	10.46	12.97	12.92	17.76
	D	0.905	0.895	0.896	0.738
	NSE	0.744	0.658	0.663	0.003
October	MAE	3.31	3.79	3.98	4.47
	MAPE	9.33	11.01	11.55	13.12
	RMSE	4.47	5.25	5.48	5.93
	D	0.915	0.929	0.922	0.905
	NSE	0.771	0.752	0.725	0.658

ISSN 1533-9211

DOI 10.17605/OSF.IO/8Z4TD

The mean absolute error (mae) value of the ok, uk and spline methods for annual rainfall increased by 4.5, 9.1 and 23.4% respectively compared to that of idw.

In similar stAudies, the idw scheme performed well when the data were regularly spaced (isaaks and srivastava 1989). It was also observed that among the univariate interpolation methods the idw offered better performance for rainfall distribution (fung et al., 2022). In contrast, firdaus, and talib (2015) used five interpolation methods for rainfall data from malaysia. Based to their results the best method was the kriging method, while, the inverse distance weighting perform worst.

Regarding the rainfall times series at seasonal time scale, it was noticed from Table 5 that that IDW method offered slightly larger MAE, MAPE and RMSE and slightly lower D and NSE than OK and UK methods. This reflects that these three methods exhibited comparable results. In general, differences between these first three methods are low but substantially different from that the spline method.

This implies that the latter method introduced the poorest results. There is also indication of decreasing the difference between the performances of different interpolation methods with a decrease in time scale. On the contrary, Liu et al., 2020 reported that the differences in performance between the spatial interpolation methods decreased with increasing time scales.

Moreover, the displayed results in Table 5 revealed that the IDW was superior to the other three methods for interpolating monthly rainfall for specific months, namely. January, April and October as representatives for the rainy season seasons of winter, spring and April respectively.

3.3. Description of the Error of Estimation by Box and Whisker Plot

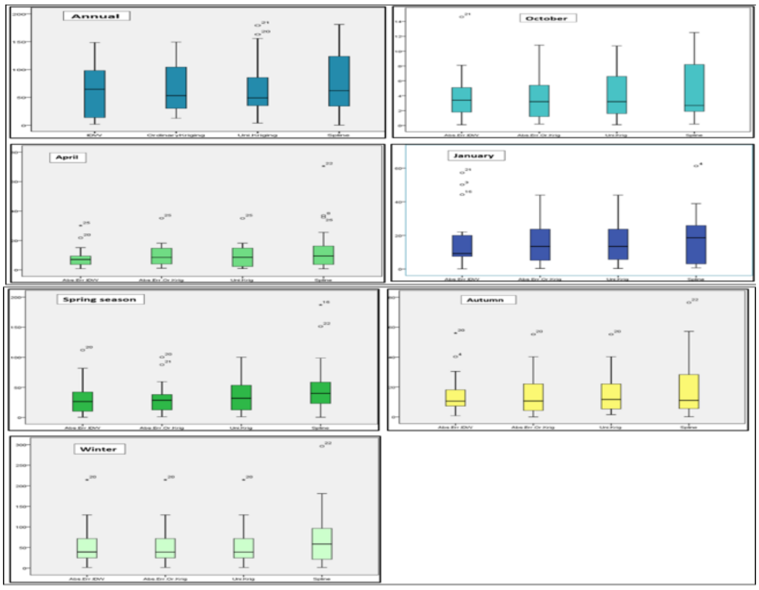
To get deeper insight into the data analysis, A box plot (known as box and whisker plot) was also plotted to visually show the distribution of mean absolute error and its skewness through displaying the data quartiles (or percentiles) and averages (Fig. 2).

As is visible from the box plots that the number of data points located outside of the boxes hardly exceed three for the all the data sets and interpolation methods signifying that the estimated depth of rainfall at some points differ significantly from the measured values (Yang and Xing, 2021). The inner fence is typically defined as 1.5 times the interquartile range (IQR) in each direction and a ‘far’ outlier or extreme case is typically defined as 3 times the IQR in either direction (Everitt and Skrondal, 2010).

ISSN 1533-9211

DOI 10.17605/OSF.IO/8Z4TD

Fig 2: Description of the Error of Estimation by different interpolation methods by Box and Whisker Plot



Additionally, it is evident from Fig. 2. That overall, the spline method has the longest box, highlighting that this method is the poorest estimator for rainfall. On the other hand, it can be observed that IDW has the shortest box for most of the datasets and followed by either Ordinary or Universal Kriging, particularly for January, April, October or autumn season. Also the results, revealed that the Universal and Ordinary Kriging offered the best performance under annual rainfall and spring season rainfall time series respectively.

As the median is not located in the middle of the box of most of the methods, and the whiskers are not about the same on both sides of the boxes, then the distribution of most of them deviated slightly from symmetric or normal distribution. Close inspection of Fig 2. Also revealed that if the median line of a given box plot does not lie outside of the other boxes indicating no significant differences among the methods.

3.4. Ranking of interpolation methods

As seen in Table 6, the interpolation methods (as four alternatives) were ranked for estimating different datasets in Duhok governorate. The Technique for Order Preference by Similarity to Ideal Solution or TOPSIS technique was used for ranking. The ranking process was based on MAE. MAPE, RMSE, agreement index (d) and Nash-Sutcliffe Efficiency coefficient (NSE) and on calculation of weights for the above criteria by Entropy method. Under each interpolation method the Euclidian distance between positive and negative ideal solutions was calculated to determine the comparatively proximity to the ideal solution. The values displayed in the last column of Table. 6 are the ranks of the values presented the column preceding the last column, which were used to evaluate the performance of the four interpolation methods. It is apparent from this analysis that IDW offered the highest performance under annual rainfall,

ISSN 1533-9211

DOI 10.17605/OSF.IO/8Z4TD

and the autumn season and the monthly rainfall of April and October followed by Ordinary Kriging (OK).

Table 6: Ranking the performance of four interpolation methods for determining the optimal method for estimating rainfall using entropy-weighted TOPSIS in Duhok governorate

Dataset	Interpolation method	d_i^b	d_i^w	$d_i^b + d_i^w$	$\frac{d_i^w}{d_i^b + d_i^w}$	Ranking
Annual rainfall	IDW	0.0013	0.0294	0.0307	0.9582	1
	OK	0.0042	0.0258	0.0301	0.8588	2
	UK	0.0084	0.0218	0.0302	0.7211	3
	Spline	0.0297	0.0000	0.0297	0.0000	4
Seasonal rainfall during winter	IDW	1.7746	6.6734	8.4480	0.7899	3
	OK	0.0002	8.4480	8.4482	1.0000	1
	UK	0.0500	8.3980	8.4480	0.9941	2
	Spline	8.4480	0.0032	8.4512	0.0004	4
Seasonal rainfall during Spring	IDW	0.3323	8.7487	9.0811	0.9634	2
	OK	0.0046	9.0809	9.0855	0.9995	1
	UK	1.0284	8.0526	9.0810	0.8868	3
	Spline	9.0809	0.0000	9.0809	0.0000	4
Seasonal rainfall during Autumn	IDW	0.0027	7.9783	7.9810	0.9997	1
	OK	0.1278	7.8505	7.9783	0.9840	2
	UK	0.6938	7.2845	7.9783	0.9130	3
	Spline	7.9783	0.0000	7.9783	0.0000	4
Monthly rainfall during January	IDW	0.5349	6.6991	7.2340	0.9261	3
	OK	0.0069	7.2339	7.2408	0.9991	1
	UK	0.1150	7.1193	7.2343	0.9841	2
	Spline	7.2340	0.0010	7.2350	0.0001	4
Monthly rainfall during April	IDW	0.0000	8.3539	8.3539	1.0000	1
	OK	0.9692	7.3865	8.3557	0.8840	3
	UK	0.9127	7.4428	8.3556	0.8908	2
	Spline	8.3539	0.0000	8.3539	0.0000	4
Monthly rainfall during October	IDW	0.0000	7.1910	7.1910	1.0000	1
	OK	0.1798	7.0136	7.1934	0.9750	2
	UK	0.4309	6.7617	7.1926	0.9401	3
	Spline	7.1910	0.0000	7.1910	0.0000	4

To further figure out or verify the results of the interpolation methods, the Spearmen rank correlation analysis as a non-parametric test was also conducted and the rank correlation coefficients were displayed in Table 7. With no exception, all the correlation coefficients were significant at ($P \leq 0.050$). The spear correlation coefficient ranged from as low as 0.527 for spine method under seasonal rainfall during spring to as high as 0.891 for IDW method under monthly rainfall for October. Based on the value of this parameter, the spline method offered the poorest performance and the IDW exhibited the highest performance for most of the analyzed datasets. It is commendable to mention that the results of this analysis support the obtained results from the remaining analysis during this study.

ISSN 1533-9211

DOI 10.17605/OSF.IO/8Z4TD

Table 7: Assessment of different interpolation methods for estimating rainfall in Duhok governorates based on Spearmen's rank correlation coefficient.

Dataset	Interpolation method	Spearman's coefficient	P-value
Annual rainfall	IDW	**0.815**	**0.0000**
	OK	**0.798**	**0.0000**
	UK	**0.806**	**0.0000**
	Spline	**0.758**	**0.0000**
Seasonal rainfall during winter	IDW	**0.686**	**0.0004**
	OK	**0.697**	**0.0003**
	UK	**0.697**	**0.0003**
	Spline	**0.743**	**0.0000**
Seasonal rainfall during Spring	IDW	**0.765**	**0.0000**
	OK	**0.839**	**0.0000**
	UK	**0.804**	**0.0000**
	Spline	**0.527**	**0.012**
Seasonal rainfall during Autumn	IDW	**0.733**	**0.0004**
	OK	**0.738**	**0.0003**
	UK	**0.704**	**0.0003**
	Spline	**0.607**	**0.000**
Monthly rainfall during January	IDW	**0.825**	**0.000**
	OK	**0.791**	**0.000**
	UK	**0.771**	**0.000**
	Spline	**0.703**	**0.0002**
Monthly rainfall during April	IDW	**0.848**	**0.000**
	OK	**0.841**	**0.000**
	UK	**0.851**	**0.000**
	Spline	**0.677**	**0.001**
Monthly rainfall during October	IDW	**0.891**	**0.000**
	OK	**0.872**	**0.000**
	UK	**0.844**	**0.000**
	Spline	**0.843**	**0.000**

3.5. Spatial Distribution of Annual Rainfall across the Study Area

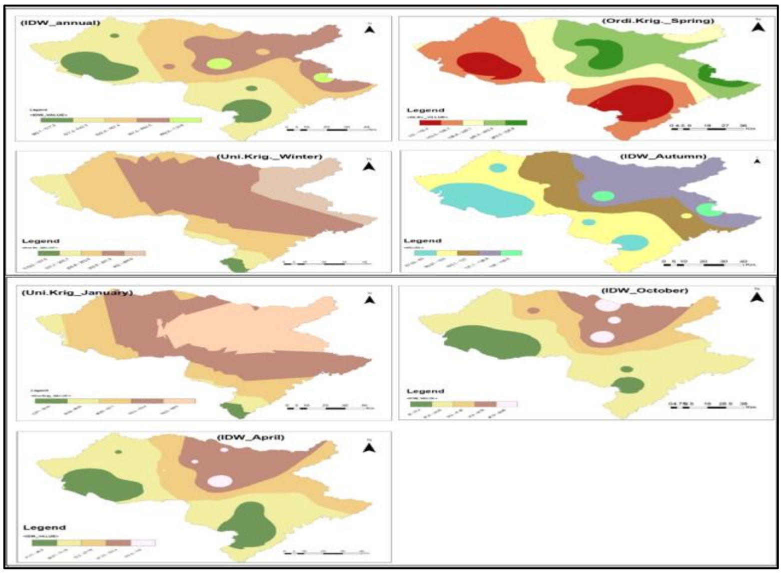
Upon selecting, the best interpolation methods applied to rainfall at different time scales, continuous rainfall surfaces were generated for each time scale within GIS environment and spatial variability maps were shown in Fig. 3. As shown in Fig. 3, it is evident that the rainfall at a given time scale tend to increase from the southern part to the mountainous area in the north and northeast.

Similar increasing trend can be observed as one moves from the western part to the eastern part. Overall, the interpolation methods provide similar spatial distributions of rainfall. Borges et al. (2016) observed that in general, dense observation network provides similar spatial distribution. The generation of interpolated map is crucial for natural resource management and for analyses of climate changes and their bad consequences.

ISSN 1533-9211

DOI 10.17605/OSF.IO/8Z4TD

Fig 3: Spatial patterns in Duhok based on four interpolation methods (IDW, OK, UK and Spline) for mean annual; seasonal and monthly rainfall



CONCLUSIONS

It was noticed that the ordinary Kriging, universal Kriging and the inverse distance weighted methods offered nearly comparable results, with no significant differences among them for estimating rainfall at different time scales. Overall, the inverse distance weighted method outperformed other interpolators for most of the time scales, while the spline interpolator offered the poorest performance. Further, for the same interpolation method, the accuracy of estimation of rainfall tended to increase with a decrease in time scale.

References

1. Akaike, H. (1974). A new look at the statistical model identification. IEEE transactions on automatic control, 19(6), pp.716-723.
2. Alahacoon, N. and Edirisinghe, M. (2021). Spatial variability of rainfall trends in Sri Lanka from 1989 to 2019 as an indication of climate change. ISPRS International Journal of Geo-Information, 10(2), p.84.
3. Ananias, D.R.S., Liska, G.R., Beijo, L.A., Liska, G.J.R. and de Menezes, F.S. (2021). The assessment of annual rainfall field by applying different interpolation methods in the state of Rio Grande do Sul, Brazil. SN Applied Sciences, 3(7), pp.1-17.
4. Anees-ur-Rehman, M., Saraniemi, S., Ulkuniemi, P. and Hurmelinna-laukkanen, P. (2017). The strategic hybrid orientation and brand performance of B2B SMEs. Journal of Small Business and Enterprise Development.
5. Antal, A., Guerreiro, P.M. and Cheval, S. (2021). Comparison of spatial interpolation methods for estimating the precipitation distribution in Portugal. Theoretical and Applied Climatology, 145(3), pp.1193-1206.

ISSN 1533-9211

DOI 10.17605/OSF.IO/8Z4TD

6. Berndt, C. and Haberlandt, U. (2018). Spatial interpolation of climate variables in Northern Germany—Influence of temporal resolution and network density. Journal of Hydrology: Regional Studies, 15, pp.184-202.

7. Borges, P.D.A., Franke, J., da Anunciação, Y.M.T., Weiss, H. and Bernhofer, C. (2016). Comparison of spatial interpolation methods for the estimation of precipitation distribution in Distrito Federal, Brazil. Theoretical and applied climatology, 123(1), pp.335-348.

8. Buytaert, W., Celleri, R., Willems, P., Bièvre, B.D., Wyseure, G. (2006). Spatial and temporal rainfall variability in mountainous areas: A case study from the south Ecuadorian Andes. Journal of Hydrology, 329(3): 413-421.

9. Charles, C, Ly, S.,. and Degre, A.(2011). Geostatistical interpolation of daily rainfall at catchment scale: the use of several variogram models in the Ourthe and Ambleve catchments, Belgium. Hydrology and Earth System Sciences, 15(7), pp.2259-2274.

10. Chen, S. and Guo, J.(2017). Spatial interpolation techniques: their applications in regionalizing climate-change series and associated accuracy evaluation in Northeast China. Geomatics, Natural Hazards and Risk, 8(2), pp.689-705.

11. Chilès JP, Delfiner P. (2012) Geostatistics: Modeling spatial uncertainty. Wiley Series in Probability and Statistics. https://doi.org/ 10.1002/9781118136188.

12. Chutsagulprom, N., Chaisee, K., Wongsaijai, B., Inkeaw, P. and Oonariya, C. (2022). Spatial interpolation methods for estimating monthly rainfall distribution in Thailand. Theoretical and Applied Climatology, 148(1), pp.317-328.

13. Di Piazza, A., Conti, F.L., Noto, L.V., Viola, F. and La Loggia, G. (2011). Comparative analysis of different techniques for spatial interpolation of rainfall data to create a serially complete monthly time series of precipitation for Sicily, Italy. International Journal of Applied Earth Observation and Geoinformation, 13(3), pp.396-408.

14. Di Virgilio, N., Monti, A. and Venturi, G. (2007). Spatial variability of switchgrass (Panicum virgatum L.) yield as related to soil parameters in a small field. Field crops research, 101(2), pp.232-239.

15. Diodato, N. and Ceccarelli, M. (2005). Interpolation processes using multivariate geostatistics for mapping of climatological precipitation mean in the Sannio Mountains (southern Italy). Earth Surface Processes and Landforms: The Journal of the British Geomorphological Research Group, 30(3), pp.259-268.

16. Drogue, G., Humbert, J., Deraisme, J., Mahr, N. and Freslon, N. (2002). A statistical–topographic model using an omnidirectional parameterization of the relief for mapping orographic rainfall. International Journal of Climatology: A Journal of the Royal Meteorological Society, 22(5), pp.599-613.

17. Everitt, B.S., Skrondal, A. (2010). The Cambridge dictionary of statistics, Cambridge University Press.

18. FirdausHum, N.N.M. and Talib, S.A. (2016). Spatial interpolation of monthly precipitation in selangor, Malaysia-Comparison and evaluation of methods. Journal of Applied and Physical Sciences, 2(1), pp.1-9.

19. Fung, K.F., Chew, K.S., Huang, Y.F., Ahmed, A.N., Teo, F.Y., Ng, J.L. and Elshafie, A. (2022). Evaluation of spatial interpolation methods and spatiotemporal modeling of rainfall distribution in Peninsular Malaysia. Ain Shams Engineering Journal, 13(2), p.101571.

20. Goovaerts, P. (1997). Geostatistics for natural resources evaluation. Oxford University Press on Demand.

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

22. Ibrahim, A.M. and Nasser, R.H.A. (2015). Comparison between Inverse Distance Weighted (IDW) and Kriging. International Journal of Science and Research (IJSR). Index Copernicus Value (2015): 78.96.

ISSN 1533-9211

DOI 10.17605/OSF.IO/8Z4TD

23. Isaaks, E.H. and Srivastava, M.R. (1989). Applied geostatistics (No. 551.72 ISA).

24. Ismail, W.W., Zin, W.Z.W. and Ibrahim, W. (2017). Estimation of rainfall and stream flow missing data for Terengganu, Malaysia by using interpolation technique methods. Malays. J. Fundam. Appl. Sci, 13, pp.214-218.

25. Khorsandi, N., Mahdian, M.H., Pazira, E., Nikkami, D. and Chamheidar, H. (2012). Comparison of Different Interpolation Methods for Investigating Spatial Variability of Rainfall Erosivity Index. Polish Journal of Environmental Studies, 21(6).

26. Krause, P., Boyle, D.P. and Bäse, F. (2005). Comparison of different efficiency criteria for hydrological model assessment. Advances in geosciences, 5, pp.89-97.

27. Li, J. and Heap, A.D. (2014). Spatial interpolation methods applied in the environmental sciences: A review. Environmental Modelling & Software, 53, pp.173-189.

28. Liu, D., Zhao, Q., Fu, D., Guo, S., Liu, P. and Zeng, Y. (2020). Comparison of spatial interpolation methods for the estimation of precipitation patterns at different time scales to improve the accuracy of discharge simulations. Hydrology Research, 51(4), pp.583-601.

29. Ly, S., Charles, C. and Degre, A. (2011). Geostatistical interpolation of daily rainfall at catchment scale: the use of several variogram models in the Ourthe and Ambleve catchments, Belgium. Hydrology and Earth System Sciences, 15(7), pp.2259-2274.

30. Mello, C.D., Viola, M.R., Beskow, S. and Norton, L.D. (2013). Multivariate models for annual rainfall erosivity in Brazil. Geoderma, 202, pp.88-102.

31. Mirás-Avalos, J.M., Paz-González, A., Vidal-Vázquez, E. and Sande-Fouz, P. (2007). Mapping monthly rainfall data in Galicia (NW Spain) using inverse distances and geostatistical methods. Advances in Geosciences, 10, pp.51-57.

32. Mitas, L. and Mitasova, H. (1999). Spatial interpolation. Geographical information systems: principles, techniques, management and applications, 1(2).

33. Murphy, A.H. and Epstein, E.S. (1989). Skill scores and correlation coefficients in model verification. Monthly weather review, 117(3), pp.572-582.

34. 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(3), pp.1009-1024.

35. Ruelland, D., Ardoin-Bardin, S., Billen, G. and Servat, E. (2008). Sensitivity of a lumped and semi-distributed hydrological model to several methods of rainfall interpolation on a large basin in West Africa. Journal of Hydrology, 361(1-2), pp.96-117.

36. Suhaila, J., Sayang, M.D. and Jemain, A.A. (2008). Revised spatial weighting methods for estimation of missing rainfall data. Asia-Pacific Journal of Atmospheric Sciences, 44(2), pp.93-104.

37. Vieux, B.E. (2001). Distributed hydrologic modeling using GIS. In Distributed hydrologic modeling using GIS (pp. 1-17). Springer, Dordrecht.

38. Yang, R. and Xing, B. (2021). A Comparison of the Performance of Different Interpolation Methods in Replicating Rainfall Magnitudes under Different Climatic Conditions in Chongqing Province (China). Atmosphere, 12(10), p.1318.

39. Zaghiyan, M.R., Eslamian, S., Gohari, A. and Ebrahimi, M.S. (2021). Temporal correction of irregular observed intervals of groundwater level series using interpolation techniques. Theoretical and Applied Climatology, 145(3), pp.1027-1037.