




Analysis of deterministic and geostatistical interpolation techniques for mapping meteorological variables at large watershed scales

Mohammad Amin Amini¹ · Ghazale Torkan² · Saeid Eslamian³ · Mohammad Javad Zareian⁴ · Jan Franklin Adamowski⁵ 

Received: 16 February 2018 / Accepted: 15 November 2018 / Published online: 29 November 2018
© Institute of Geophysics, Polish Academy of Sciences & Polish Academy of Sciences 2018, corrected publication 2019

Abstract

The widely scattered pattern of meteorological stations in large watersheds and remote locations, along with a need to estimate meteorological data for point sites or areas where little or no data have been recorded, has encouraged the development and implementation of spatial interpolation techniques. The various interpolation techniques featured in GIS software allow for the extraction of this new information from spatially distinct point data. Since no one interpolation method can be accurate in all regions, each method must be evaluated prior to each geographically distinct application. Many methods have been used for interpolating minimum temperature (T_{\min}), maximum temperature (T_{\max}) and precipitation data; however, only a few methods have been used in the Zayandeh-Rud River basin, Iran, and no comparison of methods has ever been carried out in the area. The accuracies of six spatial interpolation methods [Inverse Distance Weighting, Natural Neighbor (NN), Regularized Spline, Tension Spline, Ordinary Kriging, Universal Kriging] were compared in this study simultaneously, and the best method for mapping monthly precipitation and temperature extremes was determined in a large semi-arid watershed with high temperature and rainfall variation. A cross-validation technique and long-term (1970–2014) average monthly T_{\min} , T_{\max} and precipitation data from meteorological stations within the basin were used to identify the best interpolation method for each variable dataset. For T_{\min} , Kriging (Gaussian) proved to be the most accurate interpolation method (MAE = 1.827 °C), whereas, for T_{\max} and precipitation the NN method performed best (MAE = 1.178 °C and 0.5241 mm, respectively). Accordingly, these variable-optimized interpolation methods were used to define spatial patterns of newly generated climatic maps.

Keywords Precipitation · Sensitivity analysis · Spatial interpolation · Temperature · Zayandeh-Rud River basin

Introduction

It is increasingly important for climatologists to provide better and more accurate information regarding climatic conditions at a given place and time. This may be problematic in watersheds located in semi-arid to arid areas where complex topography and highly spatially differentiated climates make

it difficult to estimate small-scale precipitation and temperature patterns. In addition, such watersheds often have few meteorological monitoring sites and are only at point locations (e.g., climatological or synoptic weather stations) (Skirvin et al. 2003). To address this problem, interpolation techniques are used to produce meteorological data estimates for areas lacking monitoring infrastructure and instruments.

✉ Mohammad Amin Amini
m.amini@modares.ac.ir

Ghazale Torkan
ghazale.torkan@ut.ac.ir

Saeid Eslamian
saeid@cc.iut.ac.ir

Mohammad Javad Zareian
mj_zareian@yahoo.com

Jan Franklin Adamowski
jan.adamowski@mcgill.ca

¹ Department of Water Resources Management, College of Agriculture, Tarbiat Modares University, Tehran, Iran

² Department of Water Resources Management, College of Agriculture, Tehran University, Tehran, Iran

³ Department of Water Engineering, College of Agriculture, Isfahan University of Technology, Isfahan, Iran

⁴ Department of Water Resources Research, Water Research Institute (WRI), Ministry of Energy, Tehran, Iran

⁵ Department of Bioresource Engineering, McGill University, Sainte-Anne-de-Bellevue, QC, Canada

Many interpolation techniques have been tested for spatial interpolation of daily meteorological data (e.g., Dodson and Marks 1997; Thornton et al. 1997; Bolstad et al. 1998; Courault and Monestiez 1999; Shen et al. 2001; Xia et al. 2001; Jarvis and Stuart, 2001; Hasenauer et al. 2003; Garen and Marks 2005).

Choosing among the vast range of interpolation techniques to use in weather data estimation is a complex and sensitive process. The software, Geographic Information System (ArcGIS), is useful in this case as it allows users to compare the different interpolation techniques, including both deterministic [Inverse Distance Weighting (IDW), Spline, Natural Neighbor (NN), Local Polynomial (LP), Global Polynomial (GP), Radius Basis Function (RBF)] and geostatistical [Kriging and Co-Kriging (CK)] (Xiao et al. 2016). Deterministic techniques are directly based on the surrounding data points. Geostatistical theory as applied to interpolation is based on a stochastic model, which allows for optimal estimations at any location in a selected area (Wamwling 2003).

The applicability and accuracy of deterministic and geostatistical interpolation methods have been the focus of numerous studies. Mutua and Kuria (2012) evaluated kriging, CK, GP, and IDW methods for precipitation interpolation based on root mean square error (RMSE) values in Kenya's Nyando River basin and concluded that the Kriging and CK methods to be the most accurate. Delbari et al. (2013) applied different univariate [IDW and Ordinary Kriging (OK)] and multivariate [linear regression, Ordinary Co-Kriging (O-CK), Simple Kriging (SK) with varying local mean and Kriging with an External Drift (KED)] interpolation methods to map monthly and annual rainfall in northeast Iran. Results showed that OK provided the most accurate estimates. De Amorim et al. (2016) studied spatial interpolation methods involving IDW, OK, Multivariate Regression with interpolation of residuals by IDW (MRegIDW) and Multivariate Regression with interpolation of residuals by OK (MRegOK) for the estimation of precipitation distribution in Distrito Federal, Brazil. Their research showed MRegOK had the lowest errors and highest correlation and Nash–Sutcliffe efficiency criteria. Fadavi et al. (2016) investigated several interpolation methods, including IDW, Kriging, Co-K, Kriging-Regression (K-R), multiple regression and Spline in the regionalization of daily T_{\min} data from 30 meteorological stations in 1992 and 54 meteorological stations in 2007 in Isfahan, Iran. They found that multiple regression and K-R were the most accurate methods. Fadavi and Bazarafshan (2016) also compared regional estimation methods for daily T_{\max} including IDW, Kriging, Co-K, K-R, multiple regression and Spline in Isfahan. They found that multiple regression, OK and K-R had better performances than other methods. The potential of deterministic and geostatistical rainfall interpolation was examined under

high rainfall variability and dry spells in Kenya's Central Highlands by Kisaka et al. (2016). The kriging interpolation method emerged as the most appropriate geostatistical interpolation technique and was found to be suitable for spatial rainfall map generation in the study region. Das et al. (2017) concluded that the IDW method was the best spatial interpolation method in West Bengal, India. Xu et al. (2018) used P-BSHADE for temperature spatial interpolation based on sparse historical stations in China. Results were compared with kriging, IDW and combined Spline with Kriging (TPS-KRG), and it was found that the P-BSHADE method had the smallest error.

Interpolation techniques are algorithms used to solve the problem of missing geo-based data. Although these techniques encounter some limitations, including estimating errors, they are still useful. These techniques are used to predict unrecorded climatic data at each station in a catchment (e.g., semi-arid catchments, etc.) containing sufficient weather stations and climatic variable records such as precipitation and temperature. These predicted data are then used to study water resources management in the catchment area through hydrologic models.

The uncertainties associated with interpolation techniques are particularly important in arid and semi-arid watersheds where unstable weather is affected by seasonal variation, a large difference between maximum and minimum temperatures and a vast diversity of topography. Generally, finding techniques that have the best conformity to a region's specific conditions, and which can be accurately used in computing interpolated data to match with regional characteristics, is difficult. In addition, a method that performs well at one site may perform poorly at another and therefore its effectiveness needs to be examined on a site-by-site basis.

Due to the occurrence of multiple droughts and a frequent lack of sufficient meteorological data in semi-arid areas, a study was deemed necessary to determine which interpolation methods could be used for water resources management. As droughts, common to semi-arid areas, result in large changes in temperature and precipitation patterns in specific locations, previous studies on mapping would not necessarily be appropriate.

Although some methods such as IDW, NN, Regularized Spline (RS), Tension Spline (TS), OK and Universal Kriging (UK) have been individually used around the world, a comparison of these methods has never been performed in semi-arid areas simultaneously to determine the best interpolation method for mapping temperature and precipitation (Karamouz et al. 2007; Zareian et al. 2015; Eslamian et al. 2017).

The accuracy of climatic maps of mean annual minimum temperature (T_{\min}), maximum temperature (T_{\max}), and precipitation prepared using six deterministic and geostatistical spatial interpolation methods was compared. The best

method was then selected for each climatic variable in the Zayandeh-Rud River basin, Iran, a large semi-arid area with high variations in temperature and rainfall. The selection of a site-appropriate method and its relative accuracy were determined by statistical indices (e.g., Root Mean Square Error (RMSE), Mean Absolute Error (MAE), etc.) derived through a cross-validation technique.

Study area

The study was carried out in the 26,917 km² Zayandeh-Rud River basin located in semi-arid central Iran (lat. 31°15′–33°45′, long 50°02′–53° 20′; Fig. 1). The Zayandeh-Rud River is about 350 km long and is the most important river in the basin. It flows from the Zagros Mountains in the

east of Chaharmahal province (3974 m A.M.S.L.) toward the Gav-Khouni Marsh in the east of Isfahan province (1470 m A.M.S.L.) (Zareian et al. 2015).

Basin precipitation varies from about 50 mm year⁻¹ in the eastern plains to about 1500 mm year⁻¹ in the western mountainous region and has an average of 140 mm year⁻¹. The average monthly temperature for the basin ranges from 3 °C (in January) to 30 °C (in July) and the average annual temperature is 14.5 °C. The Zayandeh-Rud River had an average flow of about 850 × 10⁶ m³ year⁻¹ between 1970 and 2009 (Madani and Marino 2009). Further characteristics of this region and the distribution of selected weather stations within the basin are shown in Table 1 and Fig. 1, respectively.

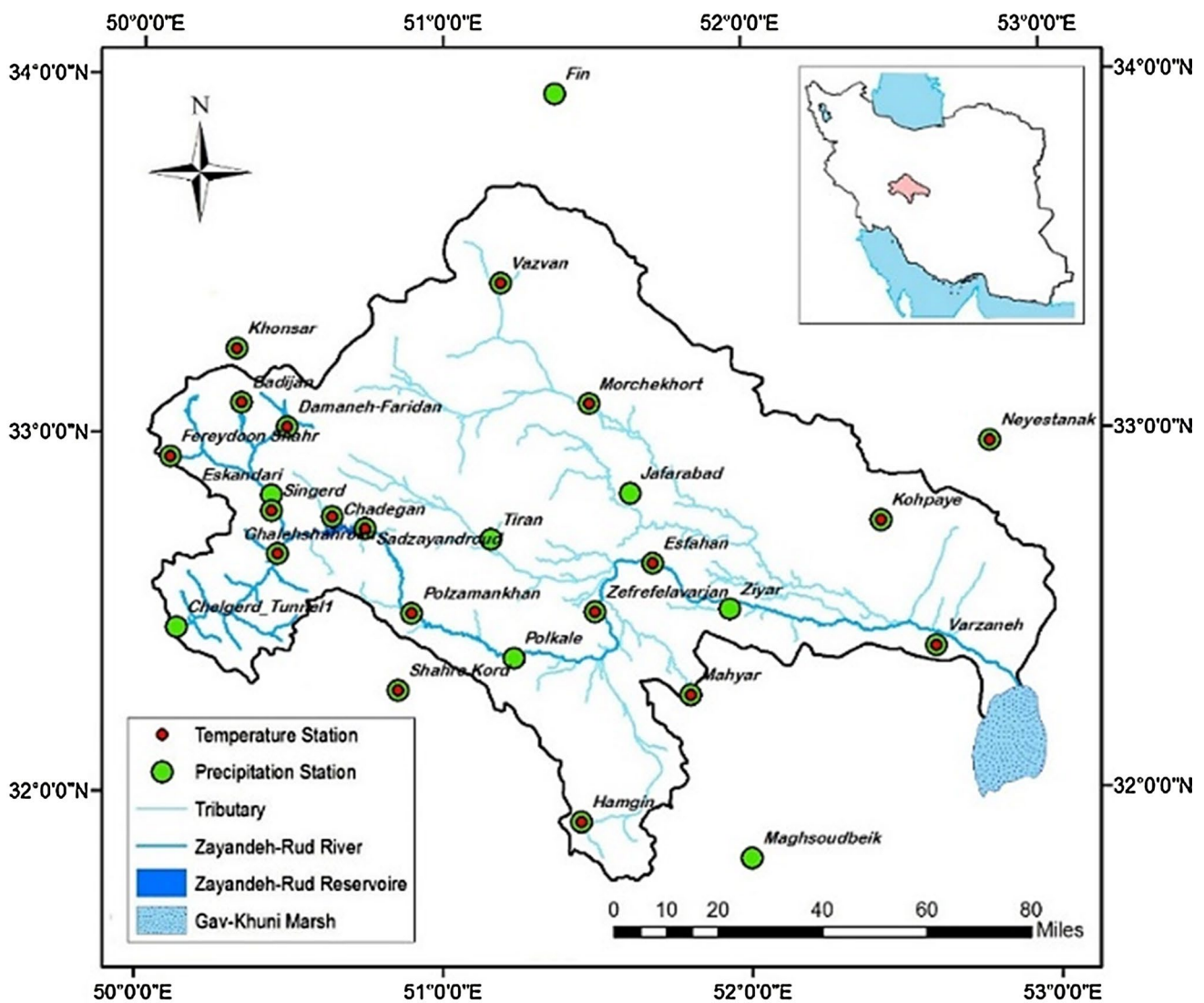


Fig. 1 The location of the Zayandeh-Rud River basin and distribution of selected weather stations

Table 1 Characteristics of the selected weather stations

No.	Station name	Longitude	Latitude	Elevation	Station type
1	Badijan	50° 19'	33° 04'	2300	Climatology
2	Chadegan	50° 37'	32° 46'	2100	Climatology
3	Chelgerd	50° 07'	32° 27'	2372	Rain gauge
4	Damaneh-Faridan	50° 28'	33° 01'	2300	Climatology
5	Esfahan	51° 41'	32° 38'	1550	Synoptic
6	Eskandari	50° 25'	32° 49'	2141	Rain gauge
7	Fereydoon-Shahr	50° 06'	32° 55'	2490	Rain gauge
8	Fin	51° 22'	33° 56'	1050	Rain gauge
9	Ghalehshahrokh	50° 27'	32° 39'	2081	Evaporation gauge
10	Hamgin	51° 27'	31° 55'	2150	Climatology
11	Jafarabad	51° 37'	32° 49'	1582	Rain gauge
12	Khonsar	50° 19'	33° 13'	2300	Climatology
13	Koohpayeh	52° 26'	32° 45'	1876	Evaporation gauge
14	Maghsoudbeik	50° 00'	31° 48'	1976	Evaporation gauge
15	Mahyar	51° 48'	32° 16'	1660	Evaporation gauge
16	Morchekhort	51° 28'	33° 04'	1669	Evaporation gauge
17	Neyestanak	52° 48'	32° 58'	1910	Evaporation gauge
18	Polkaleh	51° 13'	32° 22'	1758	Rain gauge
19	Polzamankhan	50° 24'	32° 30'	1857	Evaporation gauge
20	Sadezayandroud	50° 44'	32° 43'	2130	Evaporation gauge
21	Shahrekord	50° 51'	32° 16'	2049	Synoptic
22	Singerd	50° 25'	32° 46'	2100	Climatology
23	Tiran	51° 09'	32° 42'	1840	Evaporation gauge
24	Varzaneh	52° 37'	32° 24'	1250	Synoptic
25	Vazvan	51° 11'	33° 25'	1981	Evaporation gauge
26	Zefrefelavarjan	51° 30'	32° 30'	1622	Evaporation gauge
27	Ziyar	51° 56'	32° 30'	1530	Rain gauge

Methods

As the Zayandeh-Rud basin has a semi-arid climate and is a complicated and strategic region from the point of view of its water resources, continuous monitoring of those resources is key to sustainable water management. Thus, data obtained by the regionalized meteorological techniques could be very useful.

Data and meteorological stations

In the present study, average monthly precipitation and average monthly minimum and maximum temperature data (1970–2014) were recorded by 18 and 27 meteorological stations, respectively. The accuracy of several interpolation methods was compared via a cross-validation technique.

Spatial interpolation techniques

Within the ArcGIS environment, spatial distribution maps allowing for the analysis of spatial variation and zoning in

precipitation, T_{\min} , and T_{\max} values were created through geostatistical (kriging) and deterministic (IDW, NN, and Spline) methods.

Inverse Distance Weighting (IDW) method

The IDW technique assumes the value at an unsampled point can be approximated by a weighted average of values at measured points within a local neighborhood surrounding the unsampled location. The local neighborhood radius can be defined by the range of a fixed number of points or a fixed distance surrounding the unsampled point. The value of any point can be calculated as shown by Burrough and McDonnell (1998).

This method places a greater weight on points closer (vs. farther) to unsampled points and accordingly these have a greater effect on interpolation calculations. The weight parameter controls the effect of measured point values on the interpolated value based on their distance from the unsampled point (Erdoğan 2009).

Natural neighbor (NN) method

The NN method is the most general and robust method of interpolation. Based on a local weighted average approach, the NN method uses a Voronoi diagram to determine the relative contribution of measured points. Weights are defined as a ratio of the area “stolen” from unsampled points in the Voronoi and Delaunay diagrams by adding an interpolated point divided by the area assigned to the new point. The general relationship of interpolation by the natural neighborhood method is defined by Sibson (1981) and Watson and Phillip (1987).

Kriging method

The existence of a spatial structure, where measured points close to each other are more similar than those that are far apart (spatial autocorrelation), is a prerequisite for the application of geostatistics. These use semi-variograms as a descriptive tool to specify the spatial pattern of a feature (Goovaerts 1999). The semi-variance is a value of spatial dependence between measured points calculated on the basis of the distance between them (Tatalovich et al. 2006). The semi-variogram model that best fits the data is used to generate optimum interpolation weights (Burrough and McDonnell 2000a). Therefore, kriging is a geostatistical interpolation method that considers the spatial correlation between measured points to predict attribute values at unsampled locations using information related to other attributes (Goovaerts 1999).

After analysis of the experimental semi-variogram, a compatible model that is fitted by a parametric smooth model implying stationarity, isotropy and the best-fitting parametric model, is used to build up the auto-covariance matrix of the regression residuals (in UK). In this case, OK and SK models can be defined as described by Attorre et al. (2007).

When the measured points are modeled as the sum of constant trend and intrinsic stationary error, a more diverse batch of models is obtained. In the UK model (Ver Hoef 1993), such a trend can be modeled as a linear function in explanatory variables (i.e., climatic, geographical and topographical covariates).

While kriging is well known to be the best linear unbiased (spatial) predictor (BLUP), its applications can be limited by the existence of problems of nonstationarity in real-world data-sets. Accordingly, some authors (e.g., Agnew and Palutikof 2000; Ninyerola et al. 2000; Antonić et al. 2001), instead of using the UK with a trend function modeled via a set of covariates, have proposed a simpler approach based on RK, i.e., kriging after de-trending, such that the trend function and estimated residuals are modeled separately.

Ordinary Kriging (OK) and Universal Kriging (UK) methods

OK and UK are two common univariate kriging methods and two typically general linear regression models. The UK method assumes that spatial variation in measured point values includes a drift or structural component that can be arrived at through the variation of locations, allowing the locational information to then be incorporated into the kriging process. Unlike the OK method, where the weights of data points are determined by a stationary random function model, the weights under the UK method are taken from a nonstationary random function model (Burrough and McDonnell 2000). The OK method requires a constant but unknown mean, whereas the UK method assumes a spatially varying mean, which has proven helpful when it is necessary to account for trends observed in exploratory data analysis (Krivoruchko and Gotway 2004).

Spline method

Spline interpolation consists of the approximation of a function by means of a series of polynomials over adjacent intervals, with continuous derivatives at the end point of the intervals. The method is based on two basic assumptions: (i) The interpolating function should pass through the data points; (ii) it should be as smooth as possible.

Two types of spline techniques (TS and RS) were tested in the present study. The TS method defines a function that passes through the input data points concurrently and minimizes the curving energy function (Franke 1982; Mitas and Mitasova 1988).

Among the types of spline that can be used (see Merwade et al. 2006), the RS method improves the analytical properties of splines by adding and applying third-order and higher order derivatives.

Sensitivity analysis

Sensitivity analysis was used to determine how independent parameter values would affect particular dependent parameters under a given set of assumptions. Their usage depended on one or more input parameters within specific boundaries, such as the search neighborhood, and power and weight of smoothness that changes the prediction of a variable obtained by interpolation methods. Sensitive parameters were determined separately for each interpolation method in order to predict variables accurately.

Evaluation and comparison criteria

Cross-validation

Cross-validation is a common method for validating the accuracy of interpolation techniques (Voltz and Webster 1990). In this technique, information regarding one point is removed temporarily, the removed information is estimated from the remaining data points, and the difference between the actual and estimated values is calculated. This operation is repeated for the remainder of the points (Davis 1987).

Comparative analysis of interpolation techniques

Six different interpolation methods were validated independently for monthly T_{\min} , T_{\max} , and precipitation using cross-validation. Validation was performed for each month from 1970 to 2014, and the accuracy of the interpolation method was evaluated on the basis of Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Bias Error (MBE) and the coefficient of determination (R^2).

$$RMSE = \sqrt{\frac{\sum_{i=1}^{i=N} (x_i^{intp} - x_i^{meas})^2}{n}} \tag{1}$$

$$MAE = \frac{\sum_{i=1}^{i=N} |x_i^{intp} - x_i^{meas}|}{n} \tag{2}$$

$$MBE = \frac{\sum_{i=1}^{i=N} (x_i^{intp} - x_i^{meas})}{n} \tag{3}$$

$$R^2 = \left[\frac{\sum_{i=1}^{i=N} (x_i^{intp} - \overline{x^{intp}})(x_i^{meas} - \overline{x^{meas}})}{n(s_{x^{intp}})(s_{x^{meas}})} \right]^2 \tag{4}$$

where N is the number of evaluated data points, x_i^{intp} is the i th interpolated value of the variable studied, x_i^{meas} is the i th measured value of the variable studied, $\overline{x^{intp}}$ is the average of interpolated values of the variable studied, and $\overline{x^{meas}}$ is the average of measured values of the variable studied.

Results and discussion

Sensitivity analysis

The optimum values of parameters for all spatial interpolation methods were obtained by implementing sensitivity analysis. These parameters were selected using the RMSE index, except for the NN method, which does not require user-specified parameters (e.g., search neighborhood, power (P), and weight of smoothness) to participate in interpolations (Figs. 2, 3 and 4). The search neighborhood

Fig. 2 Variations in RMSE for T_{\min} with **a** number of points (N) for IDW, **b** power (P) for IDW, **c** number of points for RS, **d** weight of third-order derivatives for RS, **e** number of points (n) for TS, **f** weight of first-order derivatives for TS, **g** number of points for OK, **h** number of points for UK

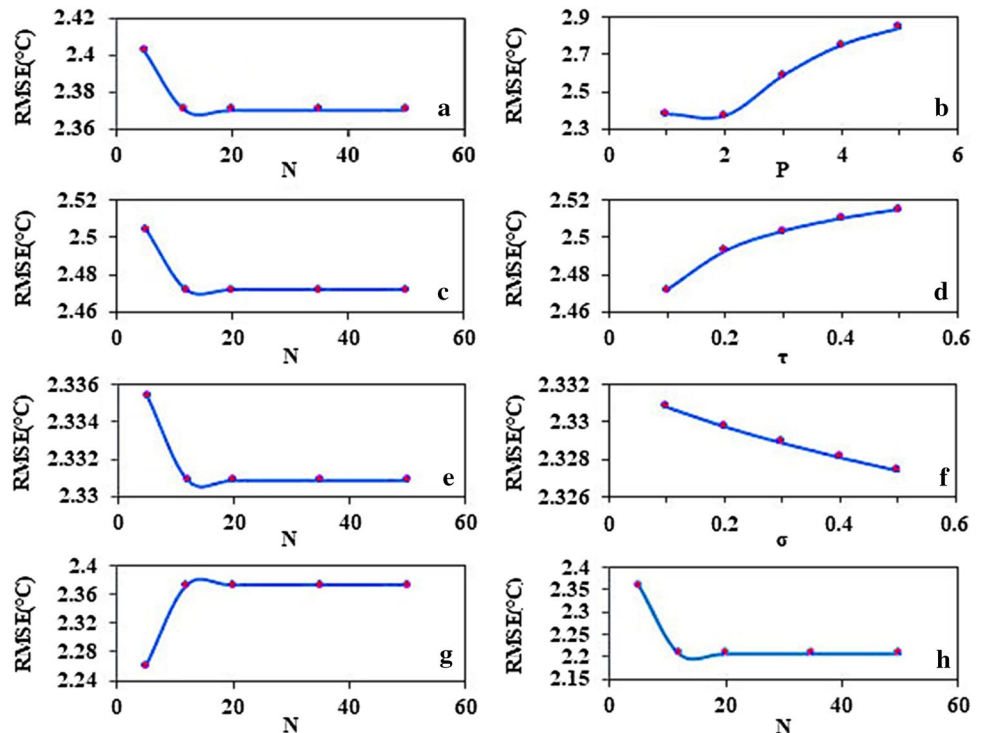


Fig. 3 Variations in RMSE for T_{max} with **a** number of points (N) for IDW, **b** power (P) for IDW, **c** number of points for RS, **d** weight of third-order derivatives for RS, **e** number of points (N) for TS, **f** weight of first-order derivatives for TS, **g** number of points for OK, **h** number of points for UK

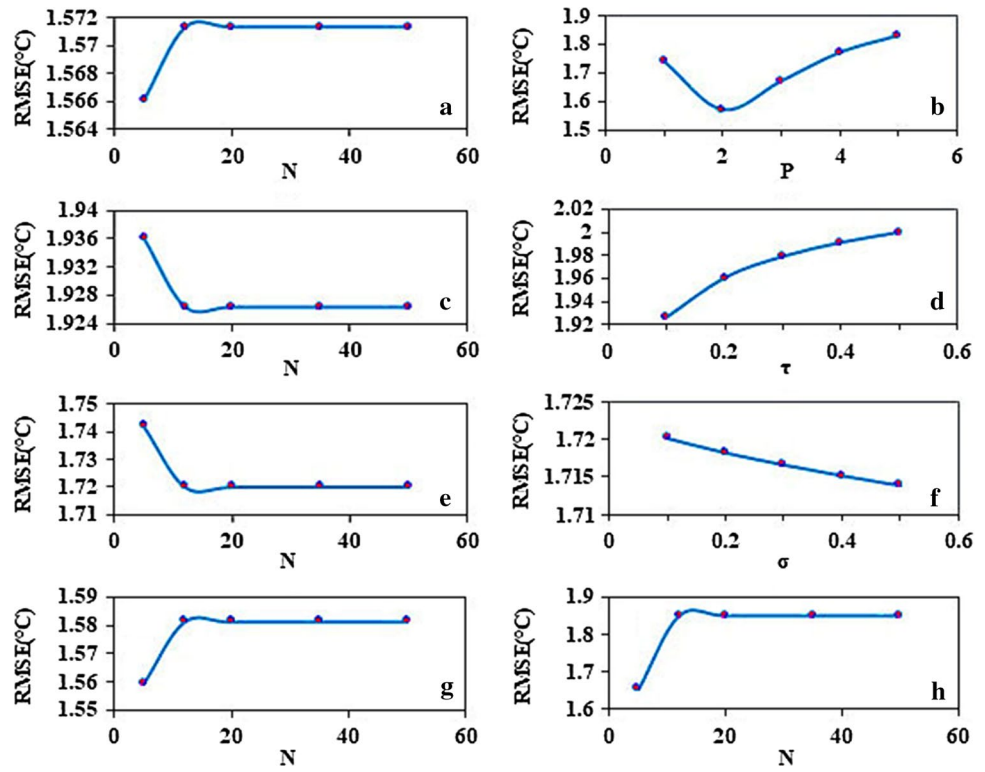
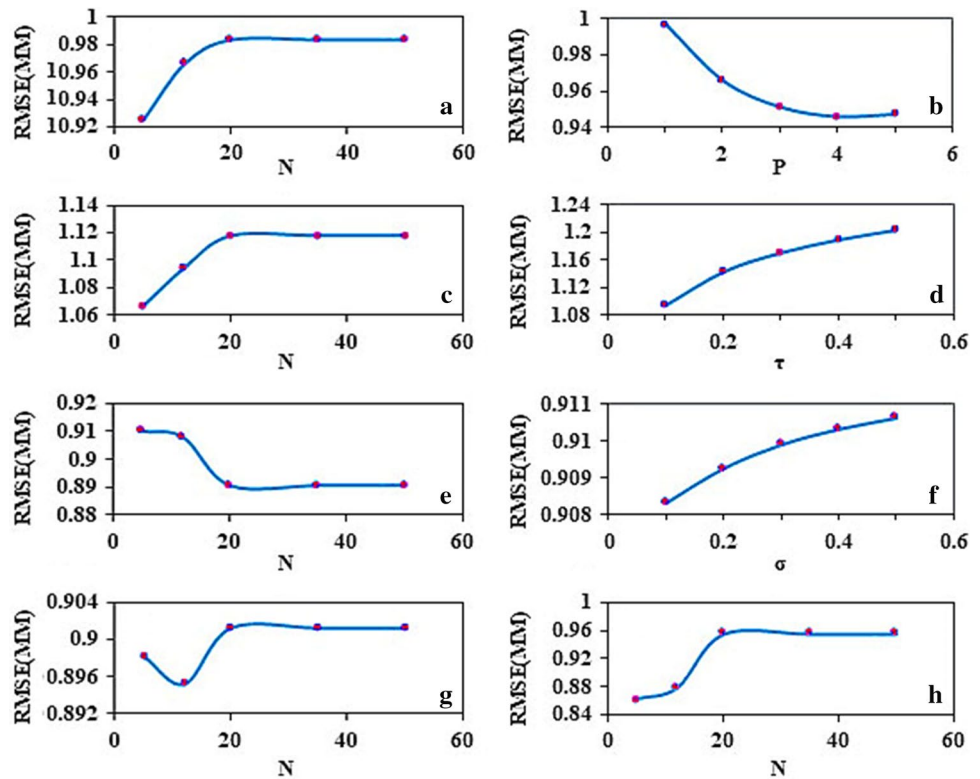


Fig. 4 Variations in RMSE for precipitation with **a** number of points (N) for IDW, **b** power (P) for IDW, **c** number of points for RS, **d** weight of third-order derivatives for RS, **e** number of points (N) for TS, **f** weight of first-order derivatives for TS, **g** number of points for OK, **h** number of points for UK



was determined by the number of points (N) for all methods. The selection of the optimal parameter in interpolation by each of the methods, except for NN, can have a significant effect on the estimation error. It is important to note that these parameters are valid only for the interpolation of data points in a specific area and may be different for data points from other regions. Sensitivity analysis results are presented for the T_{\min} , T_{\max} , and precipitation variables in Table 2.

For T_{\min} , the RMSE decreased as the value of N increased for all methods except OK (Fig. 2). For the IDW method, as P increased, the error rate first decreased and then increased significantly. For the RS and TS methods, the error rate increased and then decreased as the value of the weight parameter increased.

For T_{\max} , the two methods of the spline family showed that increasing N decreased the error rate; otherwise an increase in N led to an increase in the error rate (Fig. 3). For the IDW method, increasing the value of P led to an initial decline in estimation error, followed by an increase. Increasing the weight parameter under the RS and TS methods led to a respective increase and decrease in the error rate.

For precipitation, increasing N reduced the estimation error rate for the TS method, while for the OK method the same increase in N first decreased and then increased the estimation error (Fig. 4). Increasing the value of P when using the IDW method led to a slow decrease in the estimation error. Moreover, for the spline model family, an increase in the weight parameter led to an increase in the error rate.

Accuracy of interpolations

The interpolators applied for T_{\min} , T_{\max} and precipitation were compared by cross-validation based on R^2 , MAE,

Table 2 Accuracy of spatial interpolation methods with optimal interpolation parameters for T_{\min} , T_{\max} , and precipitation as assessed by RMSE

Interpolation method		RMSE			
Name	Optimal parameters	Minimum	Maximum	Average	SD
<i>Variable interpolated</i>					
T_{\min} (°C)					
IDW	$N=12, P=2$	2.3707	2.8421	2.4818	0.1802
NN	–	–	–	–	–
RS	$N=12, \tau=0.1$	2.4716	2.5151	2.4883	0.0186
TS	$N=12, \sigma=0.5$	2.3275	2.3354	2.3304	0.0022
OK	$N=5$	2.2607	2.3726	2.3502	0.0501
UK	$N=12$	2.2079	2.3607	2.2385	0.0683
T_{\max} (°C)					
IDW	$N=5, P=2$	1.5661	1.8309	1.6437	0.1020
NN	–	–	–	–	–
RS	$N=12, \tau=0.1$	1.9264	1.9995	1.9497	0.0299
TS	$N=12, \sigma=0.5$	1.7139	1.7419	1.7207	0.0078
OK	$N=5$	1.5593	1.5814	1.5769	0.0099
UK	$N=5$	1.6562	1.8497	1.8110	0.0866
Precipitation (mm)					
IDW	$N=5, P=4$	0.9252	0.9965	0.9649	0.0223
NN	–	–	–	–	–
RS	$N=5, \tau=0.1$	1.0667	1.2040	1.1315	0.0447
TS	$N=20, \sigma=0.1$	0.8905	0.9106	0.9039	0.0092
OK	$N=12$	0.8952	0.9012	0.8994	0.0027
UK	$N=5$	0.8616	0.9553	0.9212	0.0471

N =number of points, P =power, σ =weight factor of tension spline, τ =weight factor of regularized spline. The natural neighbor (NN) method does not include user-defined parameters

MBE and RMSE. The errors for these techniques were estimated by applying an optimal power function. The ranking of all interpolation techniques is shown for T_{\min} , T_{\max} , and precipitation in Tables 3, 4 and 5, respectively.

Table 3 Evaluation of interpolation method accuracy using cross-validation for monthly mean T_{\min} from 1970 to 2014

Method	Parameter	Overall rank	R^2	#	MAE	#	MBE	#	RMSE	#
IDW	$N=12, P=2$	11	0.934	2	1.942	10	1.558	11	2.370	10
NN	–	10	0.932	3	1.902	9	1.477	10	2.356	9
RS	$N=12, \tau=0.1$	8	0.897	11	1.978	11	1.101	1	2.471	11
TS	$N=12, \sigma=0.5$	4	0.917	10	1.870	8	1.172	4	2.327	8
OK (Spherical)	$N=5$	7	0.929	8	1.833	6	1.313	8	2.260	4
OK (Circular)	$N=5$	6	0.930	7	1.832	5	1.302	7	2.260	5
OK (Exponential)	$N=5$	3	0.930	5	1.805	1	1.234	5	2.232	2
OK (Gaussian)	$N=5$	1	0.929	9	1.827	3	1.113	2	2.254	3
OK (Linear)	$N=5$	5	0.930	6	1.829	4	1.268	6	2.264	6
UK (Linear Drift)	$N=12$	2	0.941	1	1.821	2	1.168	3	2.207	1
UK (Quadratic Drift)	$N=12$	9	0.931	4	1.861	7	1.358	9	2.292	7

N =number of points, P =power, σ =weight factor of tension spline, τ =weight factor of regularized spline. The natural neighbor (NN) method does not include user-defined parameters

Table 4 Evaluation of interpolation method accuracy using cross-validation for monthly mean T_{max} from 1970 to 2014

Method	Parameter	Overall rank	R^2	#	MAE #	MBE #	RMSE #
IDW	$N=5, P=2$	9	0.970	6	1.289	8	1.955
NN	–	1	0.972	1	1.178	1	1.521
RS	$N=12, \tau=0.1$	11	0.955	11	1.478	11	1.926
TS	$N=12, \sigma=0.5$	8	0.964	10	1.342	10	1.714
OK (Spherical)	$N=5$	7	0.971	5	1.240	5	1.559
OK (Circular)	$N=5$	5	0.971	4	1.240	4	1.559
OK (Exponential)	$N=5$	6	0.971	3	1.237	3	1.554
OK (Gaussian)	$N=5$	2	0.969	8	1.256	6	1.626
OK (Linear)	$N=5$	4	0.971	2	1.236	2	1.555
UK (Linear Drift)	$N=5$	3	0.968	9	1.286	7	1.656
UK (Quadratic Drift)	$N=12$	10	0.970	7	1.334	9	1.681

N =number of points, P =power, σ =weight factor of tension spline, τ =weight factor of regularized spline. The natural neighbor (NN) method does not include user-defined parameters

Table 5 Evaluation of interpolation method accuracy using cross-validation for monthly mean precipitation from 1970 to 2014

Method	Parameter	Overall rank	R^2	#	MAE #	MBE #	RMSE #
IDW	$N=5, P=4$	10	0.496	2	0.567	9	0.938
NN	–	1	0.506	1	0.241	1	0.425
RS	$N=5, \tau=0.1$	11	0.360	11	0.647	11	1.067
TS	$N=20, \sigma=0.1$	4	0.462	8	0.536	4	0.891
OK (Spherical)	$N=12$	8	0.483	6	0.548	7	0.898
OK (Circular)	$N=12$	6	0.485	4	0.548	6	0.897
OK (Exponential)	$N=12$	5	0.488	3	0.546	5	0.895
OK (Gaussian)	$N=12$	9	0.449	9	0.568	10	0.943
OK (Linear)	$N=12$	7	0.482	7	0.548	8	0.897
UK (Linear Drift)	$N=5$	3	0.484	5	0.520	3	0.862
UK (Quadratic Drift)	$N=12$	2	0.423	10	0.500	2	0.840

N =number of points, P =power, σ =weight factor of tension spline, τ =weight factor of regularized spline. The natural neighbor (NN) method does not include user-defined parameters

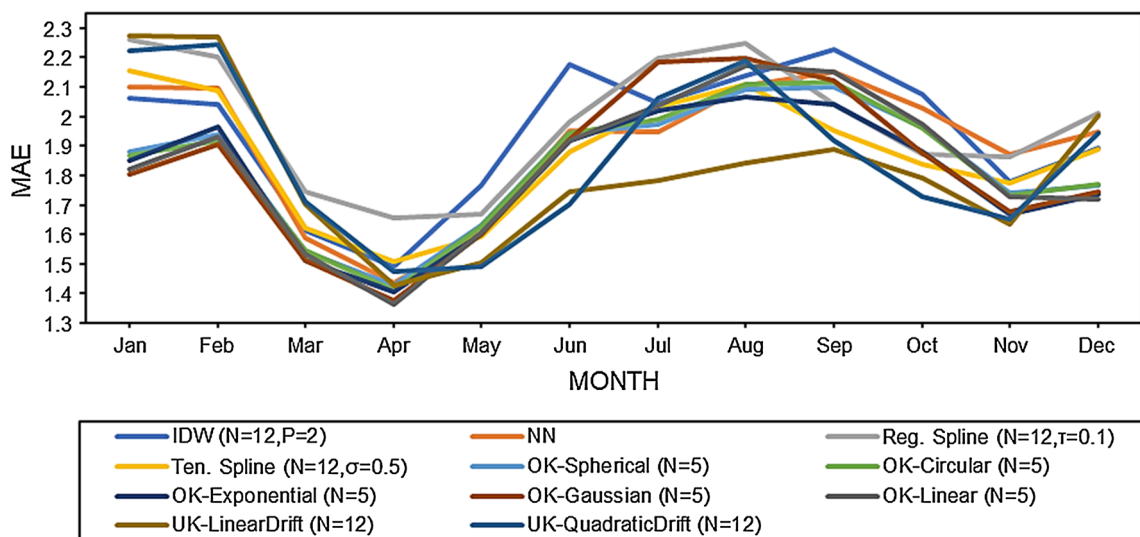


Fig. 5 Variation of monthly MAE for T_{min} using different interpolation methods

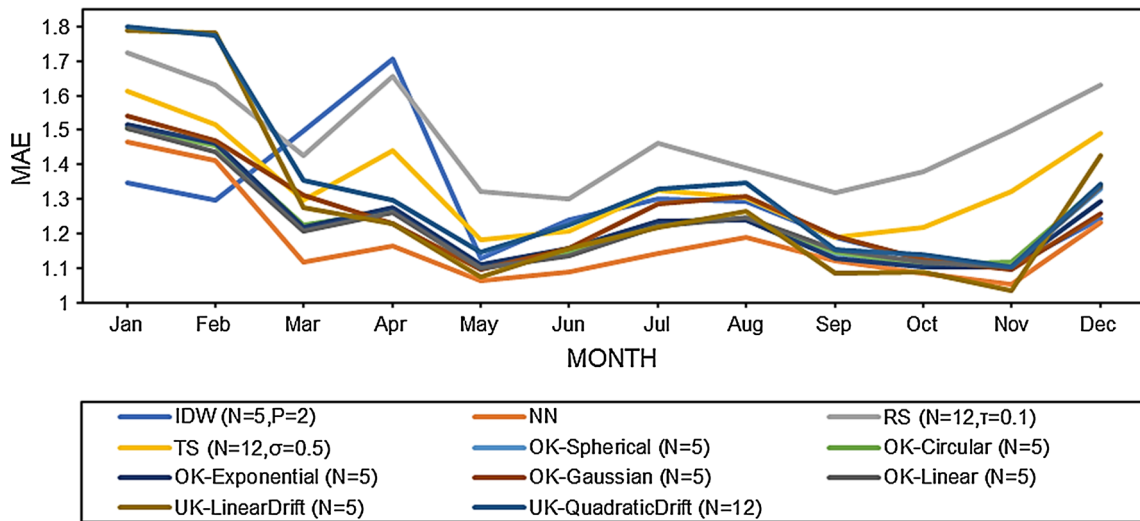


Fig. 6 Variation of monthly MAE for T_{max} using different interpolation methods

For monthly average T_{min} data, the OK method implementing a Gaussian semi-variogram (MAE = 1.827 °C) and the UK method with a Linear Drift semi-variogram (MAE = 1.821 °C), were the most accurate methods over 45 years (1970–2014) of data, achieving the 1st and 2nd ranks, respectively (Table 3). The interpolation techniques that showed the weakest results were NN (MAE = 1.902 °C) and IDW (MAE = 1.942 °C), which ranked 10th and 11th among all methods. For the spline family of methods, tension type interpolation functions outperformed regularized functions. Daily differences between the maximum and minimum temperatures are strongly influenced by factors such as frequent changes in

daily weather, rapid changes in surface temperature and the lack of air humidity, which are common in summer and winter compared to other seasons. Therefore, all interpolation methods showed the highest estimation errors for T_{min} in the summer and winter months and the lowest errors in the spring (Fig. 5).

In contrast to the monthly average T_{min} data, the NN method (MAE = 1.178) was the most accurate interpolation method for T_{max} (Table 4; Fig. 6). In a manner similar to T_{min} , the OK method with Gaussian semi-variogram (MAE = 1.256) and the UK method with Linear Drift semi-variogram (MAE = 1.286) were also among the more accurate methods. In general, kriging methods provided

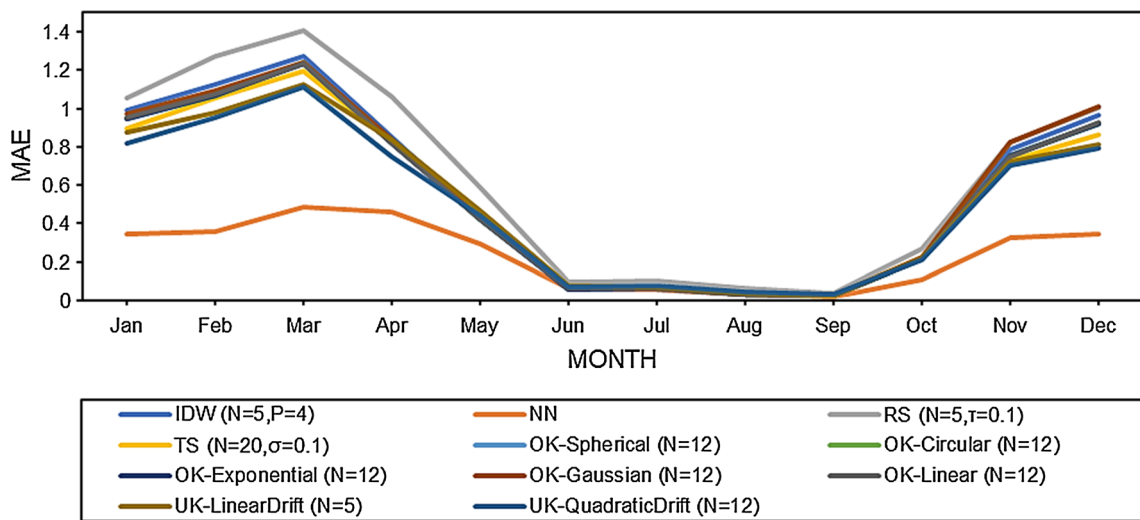


Fig. 7 Variation of monthly MAE for precipitation using different interpolation methods

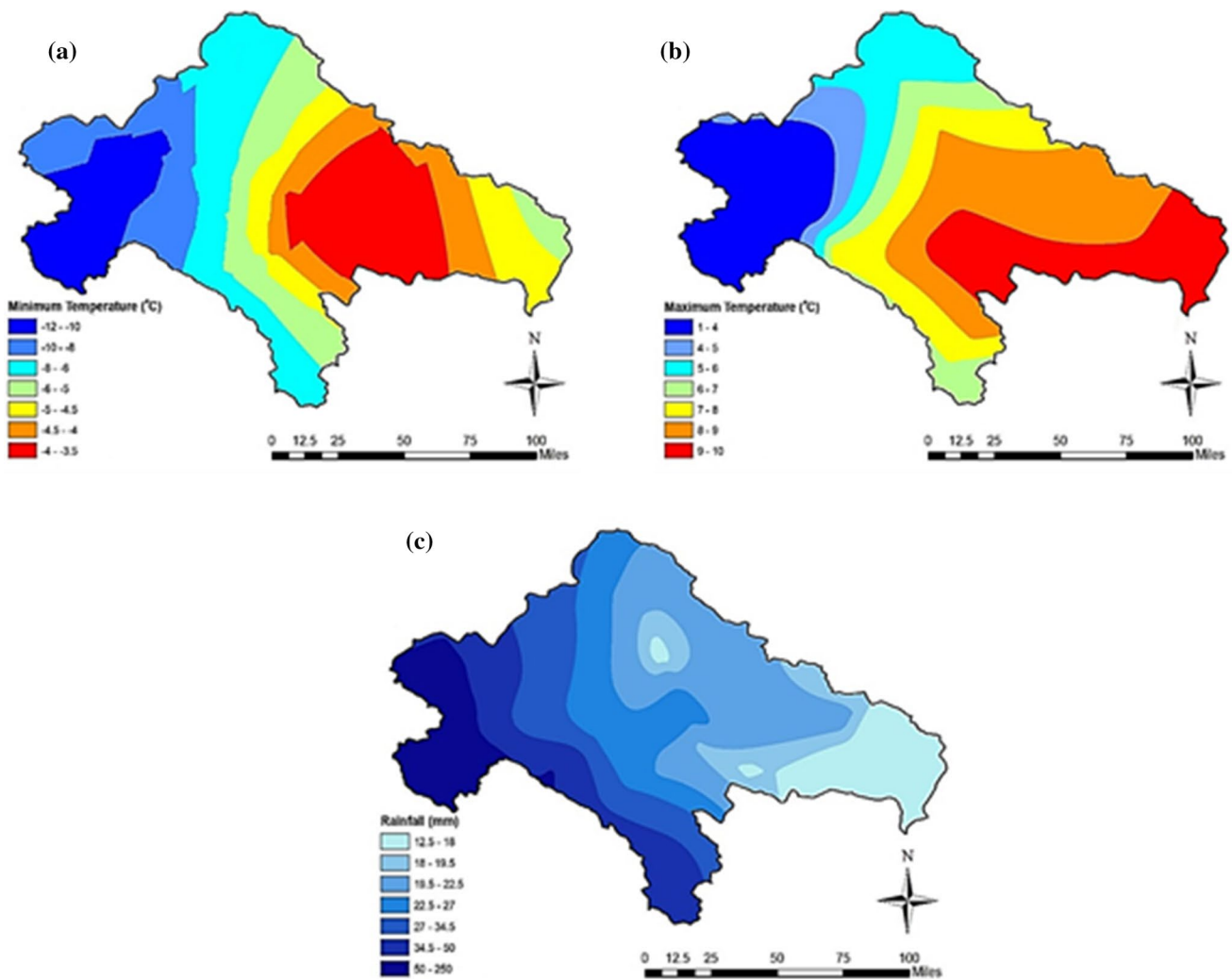


Fig. 8 Optimal monthly mean zoning using climatic data (in January) in the study area: **a** T_{\min} , **b** T_{\max} and **c** precipitation

acceptable results, except for UK with a Quadratic Drift semi-variogram, which, along with SR were among the weaker methods. All interpolation methods presented a higher estimation error in winter and lower errors in spring and autumn (Fig. 6).

The results of interpolations of mean monthly precipitation showed that the IDW and RS methods were relatively weak compared to other interpolation techniques (Table 5). All three evaluation criteria ranked them 10th and 11th in accuracy. Among kriging interpolation techniques, OK with Exponential semi-variogram (MAE=0.546) and UK with Linear with Quadratic Drift semi-variogram (MAE=0.5) showed the greatest accuracy. Overall, the NN method with the least error value (MAE=0.241) gave the best estimates among all methods followed by UK with Quadratic Drift (MAE=0.5) and UK with Linear Drift semi-variogram (MAE=0.52). It is evident that in all months of the year,

the NN method had the lowest MAE among the methods (Fig. 7).

Most of the basin's precipitation occurs during spring and autumn, from October to June. This is related to the seasonal variation in temperature and precipitation, a result of clouds generated in Europe and northern Africa, which bring heavy rains. On the other hand, in the summer, the basin receives little precipitation except for high elevation areas in the west where the mountains are located.

Accordingly, higher estimation errors occurred in the winter and lower errors in the spring and autumn, while in the summer months all interpolation methods studied showed very small estimation errors (Fig. 7).

Mapping of climatic variables

The spatial distribution patterns of optimally interpolated monthly T_{\min} , T_{\max} , and precipitation are shown for the month of January (Fig. 8). Plots for T_{\min} and T_{\max} (Fig. 8a, b) illustrated a temperature gradient from west to east with both variables lowest in the west and highest in the central and eastern portions of the study area. Precipitation increased gradually from the west and southwest to the eastern part of the basin (Fig. 8c). The results obtained in this study confirmed the results of past studies in this area such as Fadavi and Bazarafshan (2016) and Fadavi et al. (2016).

Conclusions

The present research was carried out to determine the most accurate spatial monthly patterns for three meteorological variables (T_{\min} , T_{\max} and precipitation) using six spatial interpolation techniques including IDW, NN, RS, TS, OK and UK in the Zayandeh-Rud River basin of Iran using 45 years of data (1970–2014). Among the different spatial interpolation methods, the IDW method was found to be the most sensitive to interpolation variables including T_{\min} , T_{\max} , and precipitation. Therefore, choosing correct values of parameters (N and P) is a critical step in the overall process of IDW interpolation.

The best method for spatial pattern analysis of the three climatic variables was identified by applying different error criteria. Among all methods studied, OK with Gaussian semi-variogram was identified as an appropriate technique for spatial analysis and interpolation of T_{\min} and NN was found to be the most accurate technique for both T_{\max} and precipitation.

References

- Agnew MD, Palutikof JP (2000) GIS-based construction of baseline climatologies for the Mediterranean using terrain variables. *Clim Res* 14(2):115–127
- Antonić O, Križan J, Marki A, Bukovec D (2001) Spatio-temporal interpolation of climatic variables over large region of complex terrain using neural networks. *Ecol Model* 138(1–3):255–263
- Attorre F, Alfo M, De Sanctis M, Francesconi F, Bruno F (2007) Comparison of interpolation methods for mapping climatic and bioclimatic variables at regional scale. *Int J Climatol* 27(13):1825–1843
- Bolstad PV, Swift L, Collins F, Régnière J (1998) Measured and predicted air temperatures at basin to regional scales in the southern Appalachian Mountains. *Agric For Meteorol* 91(3–4):161–176
- Burrough PA, McDonnell RA (1998) Creating continuous surfaces from point data. In: Burrough PA, Goodchild MF, McDonnell RA, Witzer P, Worboys M (eds) *Principles of geographic information systems*. Oxford University Press, Oxford
- Burrough PA, McDonnell RA (2000) *Principles of geographical information systems*. Oxford University Press, New York
- Courault D, Monestiez P (1999) Spatial interpolation of air temperature according to atmospheric circulation patterns in southeast France. *Int J Climatol J R Meteorol Soc* 19(4):365–378
- Das M, Hazra A, Sarkar A, Bhattacharya S, Banik P (2017) Comparison of spatial interpolation methods for estimation of weekly rainfall in West Bengal, India. *MAUSAM* 68(1):41–50
- Davis BM (1987) Uses and abuses of cross-validation in geostatistics. *Math Geol* 19:241–248
- De Amorim Borges P, Franke J, Da Anunciação YMT, Weiss H, Bernhofer C (2016) Comparison of spatial interpolation methods for the estimation of precipitation distribution in Distrito Federal, Brazil. *Theor Appl Climatol* 123(1–2):335–348
- Delbari M, Afrasiab P, Jahani S (2013) Spatial interpolation of monthly and annual rainfall in northeast of Iran. *Meteorol Atmos Phys* 122:103–113
- Dodson R, Marks D (1997) Daily air temperature interpolated at high spatial resolution over a large mountainous region. *Clim Res* 8(1):1–20
- Erdoğan S (2009) A comparison of interpolation methods for producing digital elevation models at the field scale. *Earth Surf Process Landf* 34(3):366–376
- Eslamian S, Safavi HR, Gohari A, Sajjadi M, Raghbi V, Zareian MJ (2017) Climate change impacts on some hydrological variables in the Zayandeh-Rud River Basin, Iran. *Reviving the dying giant*. Springer, Cham, pp 201–217
- Fadavi G, Bazarafshan J (2016) Comparative study of regional estimation methods for daily maximum temperature (a case study of the Isfahan province). *J Soil Water* 29(2):504–516
- Fadavi G, Bazarafshan J, Ghahreman N (2016) Comparison of different regional estimation methods for daily minimum temperature (a case study of Isfahan province). *J Agric Meteorol* 3(2):14–23
- Franke R (1982) Smooth interpolation of scattered data by local thin plate splines. *Comput Math Appl* 8(4):237–281
- Garen DC, Marks D (2005) Spatially distributed energy balance snowmelt modelling in a mountainous river basin: estimation of meteorological inputs and verification of model results. *J Hydrol* 315(1–4):126–153
- Goovaerts P (1999) *Geostatistics in soil science: state-of-the-art and perspectives*. *Geoderma* 89:1–45
- Hasenauer H, Merganicova K, Petritsch R, Pietsch SA, Thornton PE (2003) Validating daily climate interpolations over complex terrain in Austria. *Agric For Meteorol* 119(1–2):87–107
- Jarvis CH, Stuart N (2001) A comparison among strategies for interpolating maximum and minimum daily air temperatures. Part II: the interaction between number of guiding variables and the type of interpolation method. *J Appl Meteorol* 40(6):1075–1084
- Karamouz M, Torabi S, Araghinejad S (2007) Case study of monthly regional rainfall evaluation by spatiotemporal geostatistical method. *J Hydrol Eng* 12(1):97–108
- Kisaka M, Monicah Mucheru-Muna O, Ngetich FK, Mugwe J, Mugendi D, Mairura F, Shisanya C, Makokha GL (2016) Potential of deterministic and geostatistical rainfall interpolation under high rainfall variability and dry spells: case of Kenya's Central Highlands. *Theor Appl Climatol* 124(1–2):349–364
- Krivoruchko K, Gotway CA (2004) Creating exposure maps using kriging. *Public Health GIS News Info* 56:11–16
- Madani K, Marino MA (2009) System dynamics analysis for managing Iran's Zayandeh-Rud River Basin. *Water Resour Manage* 23:2163–2187
- Merwade VM, Maidment DR, Goff JA (2006) Anisotropic considerations while interpolating river channel bathymetry. *J Hydrol* 331(3–4):731–741
- Mitas L, Mitasova H (1988) General variational approach to the interpolation problem. *Comput Math Appl* 16(12):983–992

- Mutua F, Kuria D (2012) A comparison of spatial rainfall estimation techniques: a case study of Nyando River Basin Kenya. *J Agric Sci Technol* 14(2):149–165
- Ninyerola M, Pons X, Roure JM (2000) A methodological approach of climatological modelling of air temperature and precipitation through GIS techniques. *Int J Climatol* 20(14):1823–1841
- Shen SS, Dzikowski P, Li G, Griffith D (2001) Interpolation of 1961–97 daily temperature and precipitation data onto Alberta polygons of ecodistrict and soil landscapes of Canada. *J Appl Meteorol* 40(12):2162–2177
- Sibson R (1981) A brief description of natural neighbor interpolation. In: Barnett V (ed) *Interpreting multivariate data*. Wiley, Chichester, pp 21–36
- Skirvin SM, Marsh SE, McClaran MP, Meko DM (2003) Climate spatial variability and data resolution in a semi-arid watershed, southeastern Arizona. *J Arid Environ* 54(4):667–686
- Tatalovich Z, John Wilson P, Cockburn M (2006) A comparison of Thiessen polygon, kriging, and spline models of potential UV exposure. *Cartogr Geogr Inf Sci* 33(3):217–231
- Thornton PE, Running SW, White MA (1997) Generating surfaces of daily meteorological variables over large regions of complex terrain. *J Hydrol* 190(3–4):214–251
- Ver Hoef JM (1993) Universal kriging for ecological data. In: Goodchild MF, Parks BO, Steyaert LT (eds) *Environmental modelling with GIS*. Oxford University Press, New York, pp 447–453
- Voltz M, Webster R (1990) A comparison of kriging, cubic splines and classification for predicting soil properties from sample information. *J Soil Sci* 41(3):473–490
- Wamwling A (2003) Accuracy of geostatistical prediction of yearly precipitation in Lower Saxony. *J Environmetr* 14(7):699–709
- Watson DF, Phillip GM (1987) Neighborhood based interpolation. *Geobase* 2(2):12–16
- Xia Y, Fabian P, Winterhalter M, Zhao M (2001) Forest climatology: estimation and use of daily climatological data for Bavaria, Germany. *Agric For Meteorol* 106(2):87–103
- Xiao Y, Gu X, Yin S, Shao J, Cui Y, Zhang Q, Niu Y (2016) Geostatistical interpolation model selection based on ArcGIS and spatio-temporal variability analysis of groundwater level in piedmont plains, northwest China. *SpringerPlus* 5(1):425
- Xu C, Wang J, Li Q (2018) A new method for temperature spatial interpolation based on sparse historical stations. *J Climate* 31(5):1757–1770
- Zareian MJ, Eslamian S, Safavi HR (2015) A modified regionalization weighting approach for climate change impact assessment at watershed scale. *Theor Appl Climatol* 122:497–516