CrossMark

ORIGINAL PAPER

Statistical and geostatistical analysis related to geographical parameters for spatial and temporal representation of rainfall in semi-arid environments: the case of Algeria

Hakim Bachir[1,2] · Ahcene Semar[3] · Azzedine Mazari[4]

Abstract The economic challenges related to the fields of agriculture and industry led us to adopt the best suited method to represent the rain on the spatial and temporal plan especially in areas characterized by heterogeneous rainfall distribution additionally to drought periods. The methods of analysis and estimation of rainfall, using a number of tools (statistics, geostatistics and digital mapping), provide the opportunity to represent the average inter-yearly rainfall fields in the eastern high plateaus region of Algeria. In this study, an approach was proposed for yearly rainfall characterization. Data series for the period 1986–2007 were collected from 65 rain-gauging stations. This approach is based on two combined methods (geostatistic and multiple linear regression) including direct relationship between rainfall and geographical parameters (longitude, latitude and altitude). Statistical analysis indicates that the annual rainfall values ranges from 127 to 752.2 mm and that their distribution is Platykurtic. Results show that yearly rainfall structure obeys mainly a north/south gradient, and latitude is the most influential geographical parameter with a coefficient of 261.25 contrary to the longitude (17.06) and altitude (0.04) which have a non-significant effect on precipitation. In addition, other factors such as vegetation, temperature and air mass movement affect negatively the rainfall structure. Moreover, the map of rainfall indicates that the rain bands ranging from 300 to 400 mm dominate 58 % of the total study area whereas rain bands greater than 400 mm occupy 37 % of the total study area.

Keywords Average yearly rainfall · Multiple linear regression · Geographical parameters · High plateaus of Algeria · Geostatistic · GIS

✉ Hakim Bachir
akm7.62@hotmail.fr; hakm7.62@gmail.com

1 Hydraulics Department, Superior National School of Agronomics (ENSA-ES1603), Avenue Pasteur, Hassan Badi El Harrach, Algiers, Algeria

2 Research Division in Bioclimatology and agricultural hydraulic, Algerian National Institute for Agronomic Research (INRAA), CRP, Mahdi Boualem, Baraki, Algiers, Algeria

3 Laboratory of Applied Geology, Department of Soil Science, Superior National School of Agronomics (ENSA), Algiers, Algeria

4 Research Division of Agri-food Technology, Algerian National Institute for Agronomic Research (INRAA), CRP, Mahdi Boualem, Baraki, Algiers, Algeria

Introduction

Rainfall is as a key factor in hydrological regime (Valipour 2016; Sangati and Borga 2009). It is the result of a thermal and atmospheric change (Hamlaoui-Moulai et al. 2013), as it constitutes the main input in watershed systems (Palecki et al. 2005). Rainfall as a random variable is a constraint that may limit the policy actions both on the macro-economic and micro-economic scales, especially for countries suffering from lack of water resources (Valipour 2015a, Valipour 2015b; Valipour et al. 2015). Indeed, this variable may cause a significant risk when it comes to establish strategies without considering it (Valipour 2015c, 2012; Touchan et al. 2011; Delli et al. 2002) especially in areas with frequent droughts. Rainfall variable fluctuate in space and time. Its amount and distribution are affected by many factors including geographical, such as longitude, latitude, altitude; distance from sea; seasonality, like air humidity movement, temperature, atmospheric pressure (Slimani et al. 2007) and topography (Subyani 2004;

Springer

Zekai and Zeyad 2000; Chua and Bras 1982; Rosenberg 1969). Many studies have been conducted to correlate rainfall with the above-mentioned factors based on mathematical modelling (multiple regression, interpolations, etc.) as an alternative to improve rainfall estimation using data provided from available rain gauges (Marks et al. 2013; Lloyd 2005; Naoum and Tsanis 2004; Brunsdon et al. 2001; Goovaerts 2000; Johnson and Hanson 1995; Hevesi et al. 1992). Multiple linear regression (MLR) is a well-known method of mathematics modelling the direct linear relationship between a dependent variable and one or more independent variables (Sheather 2009; Walpole et al. 2007). It has been widely used for the estimation of climate variables (Cook et al. 1994). Independent variables generally include station location and elevation in areas where climate is significantly correlated to topography (Hay et al. 1998; Goodale et al. 1998) and have been applied with a number of topographic variables in order to analyse orographic rainfall. Precipitation could be incorrectly estimated if topographic variables influencing precipitations are not considered (Myoung-Jin et al. 2012). However, maximum cumulative rainfall does not necessarily coincide with the highest altitude (Subyani and Al-Dakheel 2009). Lloyd 2010; Ahrens 2006; and Prise et al. 2000, reported that rainfall can be estimated in space through an interpolation method with simple mathematical models (inverse distance weighting, analysis of surface trend, fluting and Thiessen polygons, etc.). Moreover, rainfall can be estimated through a more complex method relying on kriging (Skirvin et al. 2003). Actually, geostatistical interpolation has become an important tool in climatology as it takes into consideration spatial variability and quantifies the estimation uncertainty. Ashiq et al. 2010; Skirvin et al. 2003; Goovaerts 2000; and Matheron 1965, reported that geostatistical interpolation methods are based upon the structure of a variable's spatial continuity (variable with a random regionalized character). Only the spatial relationship between the sampling points is considered; the other topographical variations are not taken into account. Geostatistical approach provides a set of statistical methods that describe a spatial autocorrelation of the sample data or a natural phenomenon (Slimani et al. 2007; Kyriakidis and Journel 1999; Lam 1983). It is also used for spatialization and mapping of estimated data points from the values of targeted variable in non-sampled locations (Piccini et al. 2012). Variogram is the structure function used to model the variability of a phenomenon. It measures the variability between pairs of variables and it is expressed as a function of distance between points (Delhomme 1976). GIS can serve geostatistics to help with georegistration of data, for easier spatial explanation.

Actually, one of the main preoccupations of researchers is to provide decision-makers with tools to better adapt the future strategies of sustainable development and natural resource management. The importance of precipitation analysis in

different parts of Algeria has been addressed previous studies. However, rainfall representations were developed without considering rainfall trend, region's climate or other factors influencing this random variable in except few recent studies in which rainfall characterization was developed using new methods (Meddi and Toumi 2015: Ouallouche and Ameur 2014: Benabdesselam and Amarchi 2013: Smadhi and Zella 2012). The novelty of the current work is that its output has a direct and important impact on sustainable development of cereal cultivations in the region on one hand. On the other hand, to the best of the authors' knowledge, it is the first study that combines two approaches in order to better estimate precipitation at the spatio-temporal plan and to determine the rainfall gradient structure in the study area. The first approach based on multiple linear regression between rainfall and geographical feature (altitude, latitude and longitude) to get the weighting of each parameter and the second approach based on geostatistical calculation including explanatory information to represent the spatial and temporal distribution of average annual rainfall in the eastern high plateaus region of Algeria. This combined method was set in order to optimize rainfall representation based on data treatment given by conventional measures provided by a poor rain-gauging network. The spatial and temporal distribution of rainfall in the study area was elaborated through digital mapping where isohyet shape clarifies the rainfall behaviour in the region. The choice of this area is justified by its strategic and economic importance for sustainable development. It is historically considered as a potential producing area of cereals. Therefore, it is worthwhile to study the hydro-climatic distribution on this area in order to optimize stakeholder's interventions and the arising results are of crucial importance for agricultural development and environmental issues.

Study area description

Algeria is located in North Africa, with a total area of 2,381, 741 km^2; it is bounded on the north by the south shore of the Mediterranean Sea and to the south by the Sahel region. It is situated in the transition zone between the Mediterranean sub-humid and humid climate in the north, and arid Saharan climate in the south (Queney Queney 1937, 1943). The northern part is characterized by a cold-rainy winter and a hot-dry summer. Moreover, the climate of Algeria presents a clear east-west and north-south apparent hydro-climatic gradients. The latter, integrator of strong spatial and temporal variability, is more important as many factors are involved for its enhancement (Seltzer 1948). The distribution of precipitation is very heterogeneous and varies generally according to the relief and

the distance from the sea (Touazi et al. 2004). In addition, the yearly rains in the north part of Algeria were distributed according to a normal root distribution (Chaumont and Paquin 1971). However, northeast of Algeria is subjected to very irregular spatio-temporal variations (Meddi and Toumi 2013).

The study area is located in the eastern high plateau region of Algeria, known for its predominantly semi-arid bioclimatic affiliation. It stands between 4.2° to 8.3° latitude north and 35.00° to 36.6° longitude east, extending over an area of 33, 610 km² with a perimeter of 1872 km (Fig. 1).

Several chains of mountains naturally limit the study area. To the north, the Atlas chain is including the mountains of Constantine and Sidi-Dris with the highest points reaching 1285 and 1363 m, respectively. In the northwest is the Djurdjura Mount with a culminating point reaching 2308 m (Lala Khedidja crest). To the west is the Bibans with the relatively high points (Takoucht Mont, 1900 m, and Megress Mont, 1737 m), the highest point reaching 2000 m at Babor Mont. To the east are the mountains of Tébessa (Doukhane and Bou-Roumane Monts) with the highest points reaching 2249 and 2250 m, respectively. To the south, the Saharan Atlas chain includes the Aurès Mountains, Mahmal and Zellatou Monts with culminating points approaching 1550 m. Moreover, the study area is dominated by plains with very low slopes (0–5 %). The main plains are those of Bordj

Bou Arreridj, Setif, Oued El Othmania and El Khroub at Constantine. Furthermore, other plains occupy the area, plains of Mila in the north and plains of Touffana and Batna in the south (Bahlouli et al. 2008).

Methods and materials

Rainfall data base

The prospective work is necessary to identify data sources and to select reliable and relevant data to the objective of our study. The data processing of annual rainfall series, covering a period of 22 years (1986–2007), were collected from professional and auxiliary meteorological stations belonging to the National Meteorological Office (NMO) and from rainfall stations belonging to the National Agency for Water Resources (NAWR). All measured precipitations were in liquid form and the supplier institutions homogenized the rainfall data series. A survey was conducted to identify 95 rain-gauging stations across the study area, but it was found that some stations contained incomplete data and were operating in heterogeneous periods. Therefore, 65 rain-gauging stations were selected, for their complete dataset. For a practical presentation
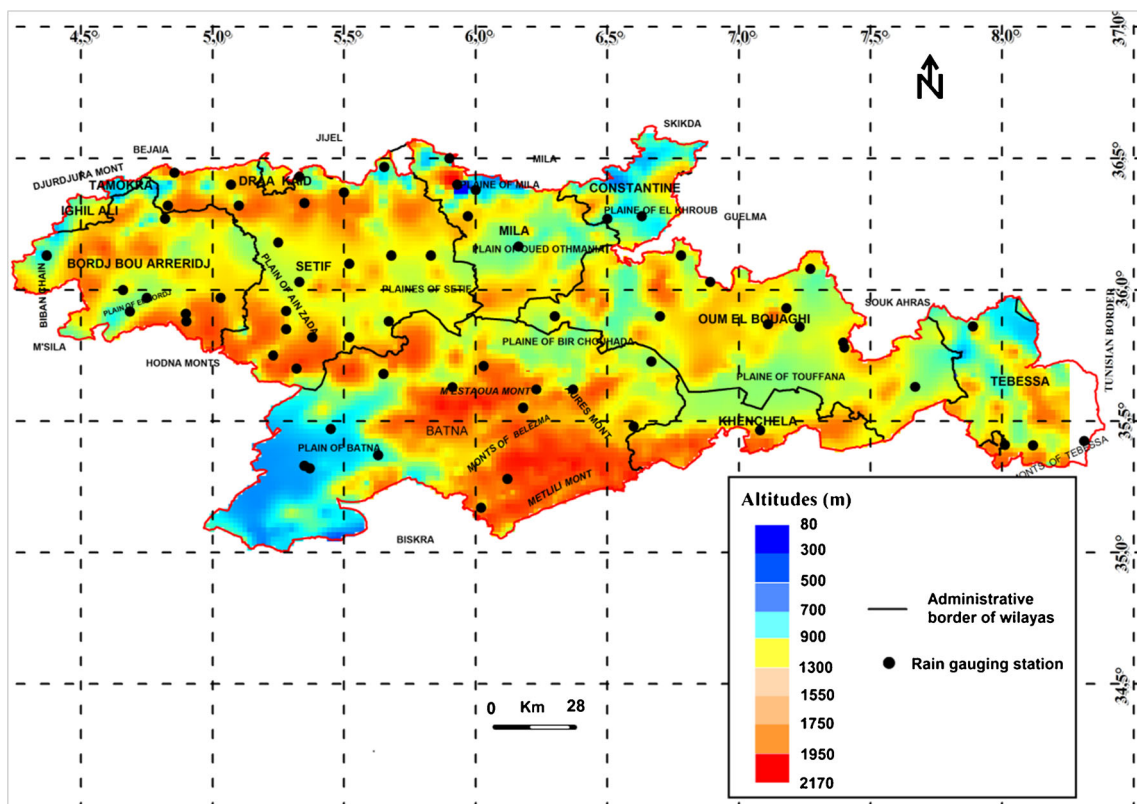

Fig. 1 Rain-gauging stations displayed over digital elevation model (DEM) of the study area

of graphics and an ease of interpretation, a code number was attributed to each rain-gauging station (Table 1).

MLR and cross validation tests

In order to estimate the local gradient of dependent variable "rainfall", Statistica 6.0 software was used to perform a MLR analysis with latitude, longitude and altitude as independent variables. Regression analysis is given by the model equation defined in Eq. (1):

$$P = \beta_1 X + \beta_2 Y + \beta_3 Z + \beta_0 + \varepsilon \tag{1}$$

where P represents rainfall as dependent variable. β_1, β_2 and β_3 are the multiple regression coefficients of the respective independent variables X, Y and Z, where X represents latitude, Y longitude and Z altitude. β_0 represents the intercept and ε the errors. Once the regression model has been constructed, a cross-validation technique was applied to confirm the goodness of fit of the statistical model, including the analysis of multiple R, R^2, R^2-adjusted analysis (Cook 1977) and the statistical tests F test, t test, root mean squared error (RMSE) and standardized mean squared error (SMSE) (Bostana et al. 2012; Walpole et al. 2007; Vargas-Guzman et al. 2000; Dingman et al. 1988; Snee 1986).

The multiple R is the positive square root of R^2 (the coefficient of multiple determinations). R^2 is a measure of the proportion of variability explained by the fitted model. It is calculated as follows:

$$R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST} \tag{2}$$

where:

$$\left.\begin{array}{ll} SSE = \sum_{i=1}^{n} \varepsilon_i^2 & \text{Sum of squares error} \\[2mm] SST = \sum_{i=1}^{n} \left(P_i - \overline{P}\right)^2 & \text{Total sum of squares} \\[2mm] SSR = \sum_{i=1}^{n} \left(\overset{\wedge}{P_i} - \overline{P}\right)^2 & \text{Regression sum of squares} \end{array}\right\} \tag{3}$$

where n is sample size (number of observations), P^i observed value of rainfall and P_i predicted rainfall value.

R^2adj is estimated by dividing SSE and SST by their respective degrees of freedom as follows:

$$R_{\text{adj}}^2 = 1 - \frac{SSE/(n-k-1)}{SST/(n-1)} \tag{4}$$

where n is the number of observations, k is the number of variables, and $(n - k - 1)$ and $(n - 1)$ represent the degrees of freedom of SSE and SST, respectively.

F-test (Fisher) is a statistical test in which the test statistic has an F-distribution under the null hypothesis. It is calculated as follows:

$$F = \frac{SSR/k}{SSE/(n-k-1)} \tag{5}$$

where k and $(n - k - 1)$ are respective degrees of freedom of SSR and SSE.

The t test most often used in multiple regression tests the significance of individual coefficients. It is calculated as follows:

$$t = \frac{\beta_j - \beta_{j0}}{s\sqrt{\sigma}} \tag{6}$$

with $s^2 = \text{SSE}/(n - k - 1)$ and B_j coefficient ($j = 0,1,2,\ldots k$).

These tests often contribute to what is termed variable screening, where the analyst attempts to reach the most useful model (Walpole et al. 2007).

The appropriateness of the chosen model is tested using the cross-validation technique based on standardized mean squared error (SMSE) and root mean squared error (RMSE) which are used together to diagnose the error variation in a set of forecasts (Bostana et al. 2012; Lloyd 2005; Miniscloux et al. 2001; Vargas-Guzman et al. 2000; Dingman et al. 1988). The RMSE is used as a measure of error's magnitude. It is calculated as follow:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left\{z(x_i) - \overline{z(x_i)}\right\}^2} \tag{7}$$

SMSE is the average ratio of the squared prediction error at validation points and the corresponding prediction error variance.

$$\text{SMSE} = \frac{1}{n}\sum_{i=1}^{n}\frac{\left\{z(x_i) - \overline{z(x_i)}\right\}^2}{\sigma_{\text{PE}}^2(x_i)} \tag{8}$$

with $\sigma_{\text{PE}}^2(x_i)$ as the prediction error variance at location x_i.

Spatial pattern analysis and geostatistical approach

A digital elevation model (DEM) with a cell size of 30 m (Fig. 1) was generated from the highest elevation point values and equidistance digital maps provided by the Algerian National Institute of Maps and Remote Sensing (NIMRS). DEMs were developed using "vertical mapper 3.0" extension of mapinfo 7.5 software. The influence of geographic parameters affecting rainfall was identified through multiple linear regression analysis. Geostatistical calculation for variogram adjustment was performed using "VarioWIN 2.2" software

Table 1　Geographical coordinate of rain-gauging stations

| Station code	Wilaya (province)	Station name	Latitude (degree)	Longitude (degree)	Altitude (m)
1	Bordj Bou Arreridj	El Mehir	36,13	4,37	550
2	Bordj Bou Arreridj	B.B.A	36,00	4,66	928
3	Bordj Bou Arreridj	Bouaziz	36,27	4,82	850
4	Bordj Bou Arreridj	El Hamadia	35,97	4,75	850
5	Bordj Bou Arreridj	Ghafsitiane	35,93	4,67	900
6	Bordj Bou Arreridj	Ras El Oued	35,97	5,03	1000
7	Bordj Bou Arreridj	Bordj Ghdir ANRH	35,91	4,90	1054
8	Bordj Bou Arreridj	Bordj Ghdir ONM	35,88	4,90	1050
9	Batna	Barika	36,33	5,35	460
10	Batna	O.S.Slimane	35,47	5,45	622
11	Batna	Bitam	35,32	5,37	474
12	Batna	Ras El Ayoun	35,68	5,65	880
13	Batna	Seggana	35,37	5,63	673
14	Batna	Merouana	35,63	5,91	1000
15	Batna	Batna A	35,71	6,03	468
16	Batna	Tifelfel	35,62	6,23	800
17	Batna	Attouta	35,55	6,18	471
18	Batna	Bouzina	35,28	6,12	1350
19	Batna	Menaa	35,17	6,02	1000
20	Batna	El Madher	35,62	6,37	913
21	Batna	Boulhilet	35,73	6,67	985
22	Batna	Toufana	35,48	6,60	1040
23	Constantine	Ain Smara	36,27	6,50	680
24	Constantine	Constantine	36,28	6,63	419
25	Khenchela	Khenchela	35,47	7,08	983
26	Mila	Ferdjioua	36,40	5,93	580
27	Mila	Beni Guecha	36,38	6,00	550
28	Mila	Ain Trik	36,28	5,97	1100
29	Mila	Chelghoum Laid	36,16	6,16	768
30	Oum el Bouaghi	Bir Chouhada	35,90	6,30	800
31	Oum el Bouaghi	Sigus	36,13	6,78	820
32	Oum el Bouaghi	Ain Kercha	35,90	6,70	780
33	Oum el Bouaghi	Ain El bordj	36,03	6,89	780
34	Oum el Bouaghi	O.E.B	35,87	7,11	889
35	Oum el Bouaghi	Ksar Sbahi	36,08	7,27	850
36	Oum el Bouaghi	Ain Babouche	35,93	7,18	893
37	Oum el Bouaghi	Medfoun	35,86	7,23	900
38	Oum el Bouaghi	Ain El Beida ANRH	35,80	7,40	1004
39	Oum el Bouaghi	Ain El Beida ONM	35,78	7,40	1000
40	Oum el Bouaghi	Meskiana	35,63	7,67	970
41	Setif	Ourtilen	36,45	4,86	1020
42	Setif	Guenzet	36,32	4,83	1100
43	Setif	Maoklane	36,40	5,07	950
44	Setif	Bougaa	36,32	5,10	886
45	Setif	Setif	36,18	5,25	1030
46	Setif	Tizi Nbechar	36,43	5,33	900
47	Setif	Guellal	36,03	5,33	908
48	Setif	Ain Oulmane	35,92	5,28	960
49	Setif	Saleh Bey	35,85	5,28	978
50	Setif	Ain El Kebira	36,37	5,50	1020
51	Setif	Ras El Ma	36,10	5,52	980
52	Setif	Beni Aziz	36,47	5,65	750
53	Setif	El Eulma	36,13	5,68	964
54	Setif	Beida Bordj	35,88	5,67	895
55	Setif	Bouandas	36,50	5,90	580
56	Setif	Bir El Arch	36,13	5,83	950
57	Setif	Rasfa	35,75	5,23	1100
58	Setif	O.Teben	35,82	5,38	1030
59	Setif	Boutaleb	35,70	5,32	960
60	Setif	Meghres	35,33	5,35	450
61	Setif	Ain Azel	35,82	5,52	960
62	Tebessa	El Aouinet	35,86	7,89	654
63	Tebessa	Tebessa	35,41	8,01	820
64	Tebessa	Tabessa ANRH	35,41	8,12	850
65	Tebessa	Tebessa pciv	35,42	8,31	853
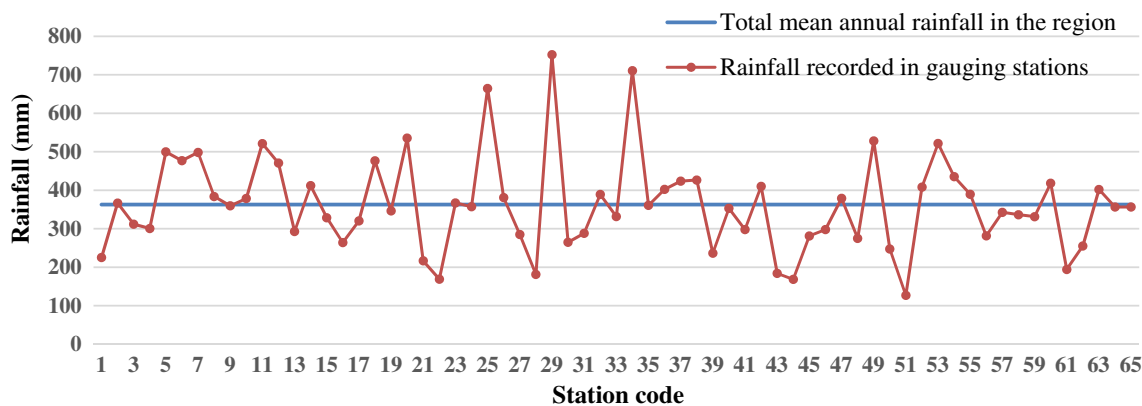
Fig. 2 Yearly rainfall recorded in rain-gauging stations across the study area for the period of 1986 to 2007

(Pannatier 1996) and inter-yearly rainfall mapping was achieved using "Surfer 8.0" and "Mapinfo 7.5" software.

Geostatistical interpolation method of rainfall is based on the spatial continuity structure of this variable. It provides a set of statistical methods that describe a spatial autocorrelation of the sample data of rainfall. In its overall conception, the kriged or predicted value $Z(x_0)$ is a linear combination of observations at N_{nb} neighbour stations (Bargaoui and Chebbi 2009). Kriging is applied to estimate the values of the rainfall unsampled locations using the points around it. The kriging estimation is expressed as follows:

$$Z(x_0) = \sum_{i=1}^{N_{nb}} \lambda_i Z(x_i) \tag{9}$$

with $Z(x_0)$ as the estimator of the mean Z on x_0 and $Z(x_i)$ the known value Z at the point x_i. N_{nb} are a number of data points used for estimation and λ_i are kriging weights which are estimated as a solution of the kriging system. The weightings involved in the linear combination are obtained by solving the minimization problem whose equations depend on the theoretical variogram and the geometric configuration of rainfall data point's knowledge (Arnaud and Emery 2000). The equation of semi-variogram is expressed as follows:

$$\gamma(h) = \frac{1}{2m} \sum_{i}^{m} \{Z(Xi) - Z(Xi + h)\}^2 \tag{10}$$

where h is the distance between X_i and X_j, and m is the number of pairs which are separated by the distance h. The obtained variogram is characterized by the nugget effect, the range and the Sill. The experimental variogram is adjusted on theoretical models based on the value of the Indicative Goodness of Fit (IGF) which is a basic criterion for the selection of the adjusted variogram model (Pannatier 1996). An IGF value close to zero indicates a good fit of the model.

Results and discussion

Considering the standards for meteorological station setting, whose approximate coverage tends from 20 to 40 km^2 in plains and 2 to 10 km^2 in mountainous area (Bertrand-Krajewski et al. 2000), the density of rain-gauging stations in the study area is low. Fundamental statistics tests for yearly rainfall data were applied to characterize the location and variability of a data set. As can be seen in Fig. 2, the average annual rainfall values are fluctuating considerably from one gauge station to another varying between an interval of 127 mm in Ain Kercha (code 32) and 752.2 mm in Beni Aziz (code 52). The average annual rainfall of the study area is 362.5 mm with a coefficient of variation of 0.33 and a standard deviation of 122.33 mm (Table 2).

A further skewness, Kolmogorov-Smirnov and kurtosis tests were applied to characterize rainfall frequency and its distribution shape. The computed Kolmogorov-Smirnov test value (K-S = 0.1) at null hypothesis was less than the corresponding critical value of significance ($p = 0.614$). Thus, the

Table 2 Descriptive statistics and normality test of used data

Statistics	Rainfall	Longitude	Latitude	Altitude
Minimum	127.06	4.37	35.17	419
Maximum	752.13	8.31	36.5	1350
Mean	362.44	5986	35.91	855.88
First quartile	284.28	5.31	35.66	777
Third quartile	413.57	6.64	36.16	988.75
Standard error	15.17	0.11	0042	24.06
Average deviation	89.80	0.77	0278	150.97
Standard deviation	122.33	0.95	0.34	194.02
Coefficient of variation	0.33	0.15	0.0095	0.22
Skew	0.85	0.64	−0.16	−0.62
Kurtosis	1.52	−0.34	−0.82	0.14
K-Smirnov	0.1	0118	0.08	0.15
p value	0.614	0.071	0.375	0.002

Table 3 Coefficients of the MLR application and validation test

Coefficient synthesis

	β	t test	p level
Origin	−9159	−7.082	0.000000
Longitude	17.06	1.354	0.180
Latitude	261.25	7.454	0.000000
Altitude	0.04	0.758	0.450
	Cross validation of MLR		
Multiple R	0.703	ε	89.9
R^2	0.595	t test	−7.126
$R^2_{adjusted}$	0.515	p level	0.00000
MSE	0.023	df	3.61
RMSE	0.15	F-test	19.97

hypothesis regarding the distributional shape is not rejected as the K-S value is smaller than the critical value of significance. Moreover, in Table 2, the skew value is equal to 0.85, indicating that the distribution is moderately skewed with an asymmetric tail extending toward positive values. The calculated kurtosis value (1.52) is lower than 3, while a normal distribution has a kurtosis of 3 (Bulmer 1979), indicating that the distribution is Platykurtic, which means that the probability for extreme values is less than a normal distribution and the

values are wider spread around the mean (Walpole et al. 2007; Bulmer 1979).

A rain-specific MLR model is developed for all 65 rain-gauging stations, using cumulated yearly rainfall variables of each station. From the interactive model regression analysis, the coefficient values found were as follows: $\beta_1 = 261.25$, $\beta_2 = 17.06$, $\beta_3 = 0.04$ and the intercept $\beta_0 = −9159$. By inputting the obtained results in Eq. (1), the equation of annual rainfall obtained from regression analysis becomes

$$P(\text{mm}) = 261.25X + 17.06Y + 0.04Z − 9159 + \varepsilon \qquad (11)$$

with X, Y and Z being, respectively, latitude, longitude and altitude. ε represents error of estimation and it is equal to 89.92. Table 3 displays the regression coefficients values and the validation tests of the MLR model. According to Table 3, the goodness of fit of rainfall MLR model has relatively low multiple R, R^2 and adjusted R^2 ($R^2 = 0.595$, multiple $R = 0.703$ and the adjusted $R^2 = 0.515$). F-test was used to check the overall significance of the developed MLR model. The advantage of the F-test over R^2 is that the F-test takes into account the degrees of freedom, which depend on the sample size and the number of predictors in the model (F (3.61) = 19.41, p value = 0.00000). In addition, t test, RMSE and MSE were used to evaluate the performance of the MLR equation ($t = −7.126$, p level = 0.00000, RMSE = 0.15, MSE = 0.023) indicating that indeed the
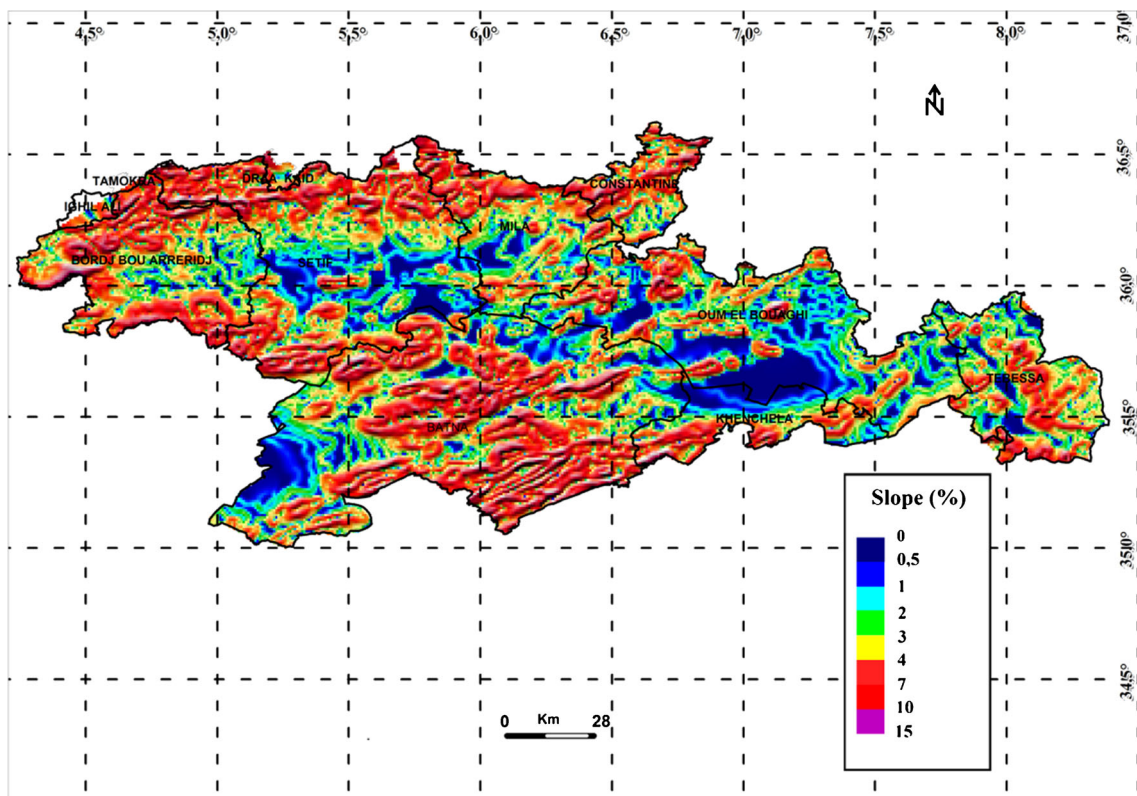

Fig. 3 Slope range distribution in the study area

Table 4   Slope range distribution of the study area

Slope range (%)	Surface (km²)	Percentage of slope surface range versus total surface
<5	29738.13	88.48
5–10	3697.10	11
10–15	151.24	0.45
>15	23.53	0.07

MLR model is significant at the 0.05 confidence level. Moreover, X ($t = 7.454$, $p = 0.000000$) and intercept β_0 ($t = -7.454$, $p = 0.000000$) are highly significantly explanatory variables, indicating higher t values and very low p value at the 95 % significance level. However, Y ($t = 1.354$, $p = 0.180$) and Z ($t = 0.758$, $p = 0.450$) are statistically insignificant in the MLR estimation.

Equation (11) shows the impact of each geographic parameter on rainfall. The weight of latitude, with a coefficient of 261.25, is the most important among the weights of the other parameters. With a coefficient of 17.06, longitude affects slightly rainfall. Moreover, altitude with a coefficient of 0.04 has insignificant effect on precipitation. This may be due to the fact that the relief is not sufficiently contrasted in the study area. To better explain this situation, a slope map of the region was developed (Fig. 3).

Around 29,738 km² consists mainly of height plateaus with low slopes (<5 %), representing 88.5 % of the total study area and 11 % covers areas with a slope ranged between 5 to 10 % (Table 4).

The majority of rain-gauging stations (45 stations, see Table 1) are located between 800 and 1100 m elevations. To check whether precipitation is correlated to altitude, the correlation coefficient was calculated according to Eq. 12 expressed as follows:

$$P = 324.93 + 0.056\,Z \tag{12}$$

where $R = 0.099$, P: rainfall (mm), and Z: altitude (m).

The calculated R confirm the insignificant correlation ($R = 0.099$) between rainfall and altitude (Fig. 4).

According to the MLR model, 59.5 % of the observed variability of rainfall is attributed to the geographical parameters ($R^2 = 0.595$), leaving 40.5 % residual variance attributable to unmeasured complications. Multiple regression in this case gives a more accurate description of the regional distribution of precipitation; it appears that the intercept (−9159) plays an important negative impact on the annual rainfall. This result indicates that the rainfall is influenced by other environmental factors such as distance from sea, air humidity movement, land cover vegetation, temperature and atmospheric pressure (Subyani 2004; Zekai and Zeyad 2000; Chua and Bras 1982; Rosenberg 1969). Rainfall spatial variability is clearly related to a north/south gradient. Considering the position of the study area, which is situated on the southside of the Tellian Atlas chain, the northern part of this area can take much advantage of orographic rains contrarily to the central and south part (Miniscloux et al. 2001). High temperatures and dry air masses coming from the Sahara generate warm front that is playing the role of barrier against the cold front
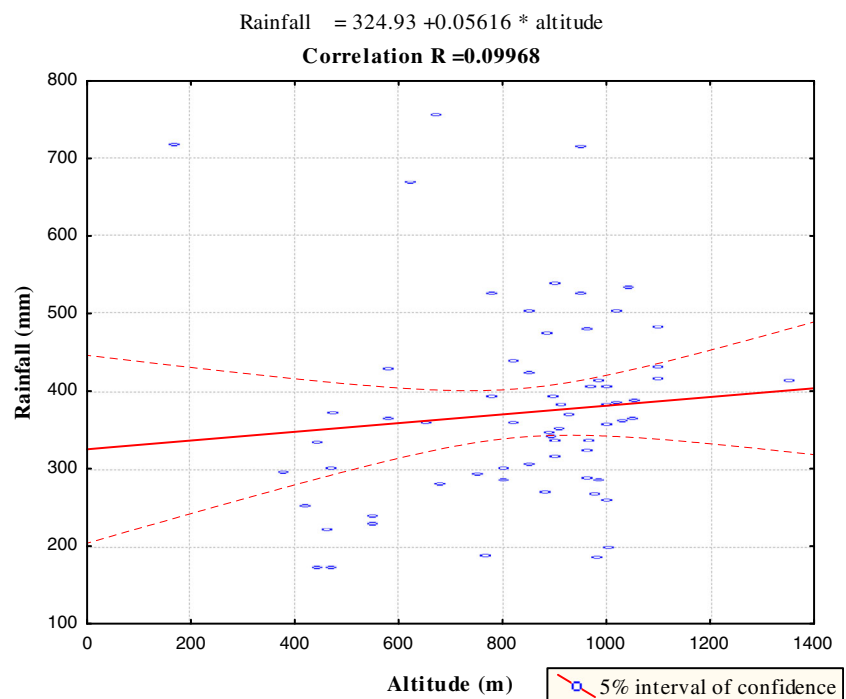
Fig. 4 Correlation between rainfall and altitude

Rainfall = 324.93 +0.05616 * altitude

Correlation R =0.09968

Table 5 Theoretical variogram parameters

Variogram	Theoretical variogram (model)				
Parameters	Model	Nugget effect	Sill/power	Range/slope (km)	IGF (best fit found)
Annual rainfall according to the north/south direction (0°)	Power	1280	8200	171	$1.44e^{-3}$
Annual rainfall according to the east/west direction (90°)	Gaussian	4500	17500	100	$4.98e^{-2}$

coming from the north. In addition, the lack of vegetation in the study area, which is dominated by annual crops like wheat and barley, contributes to reduce rainfall amounts.

The achievement of directional variograms from rainfall data, according to the north/south and east/west directions, permitted to identify the structure of the rainfall phenomenon. Results revealed that directional variograms fit well with the theoretical power and Gaussian variograms respectively in both this two directions. Table 5 lists the variogram parameters and the cross-variogram test.

Figures 5 and 6 illustrate the structure of variograms which fit a power model with a best fit found value equal to $1.44e^{-3}$ and a Gaussian model with best fit found (IGF) equal to $4.98e^{-2}$, respectively. The linear structure of the power variogram indicates the existence of a drift. Thus, rainfall varies faster toward the north/south direction. In addition, variograms show the existence of increasing range of fluctuations limited to the distances of 171 and 100 km, respectively (Table 5). The nugget effect denotes the existence of rapid fluctuations undetectable by the climate network set in place. The Gaussian variogram with higher nugget effect indicates the existence of a microclimate in the region (Slimani et al. 2007).

Ordinary kriging interpolation was applied for the precipitation data in order to perform isohyet mapping. Figure 7 shows the vertical rainfall distribution and the north/south contrast effect in the study area. The isohyets are generally arranged in parallel to the northern orographic barriers. The horizontal organization of mount chains delimiting the northern part of the study area lead to upstream interception of clime disturbances coming from the north. Thus, the rain bands greater than 500 mm occupy the northern part of the area representing 12.41 % where 10 % of the total study area is affected by the rain band ranging from 500 to 550 mm. In the central part of the study area, isohyets follow a clear north/south gradient caused by hot and dry air masses coming from the south emphasized during the months of sirocco. The rain bands covering isohyets 300 to 400 mm dominate 58 % of the study area, followed by the rain band 400 to 500 mm (24.31 %). In the far east region of the study area, the rain bands adopt an east/west gradient; this could be due to humid airflow masses coming from the Mediterranean Sea through Tunisia, thus promoting rainfall to the limits of Tunisian border. Also, circular rain pockets that appear in the central part of the region substantiate the presence of a microclimate as demonstrated by the higher value of nugget effect detected (Table 5).

Rainfall bands ranging from 200 to 350 mm cover 39.25 % of the area should be dedicated exclusively to the cultivation of barley with supplemental irrigation for areas with annual rainfall amount less than 300 mm. Rain bands ranging from 350 to 500 mm occupy 48.32 % of the study area is suitable for the cultivation of durum wheat with supplemental irrigation for areas with annual rainfall amount less than 450 mm. Rain bands greater than 500 mm is more favourable for the cultivation of bread wheat (*Triticum aestivum*).

Fig. 5 Adjusted theoretical variogram of mean yearly rainfall according to the north/south direction
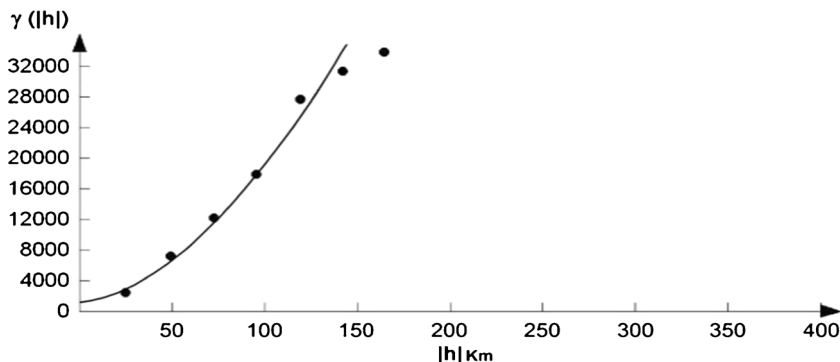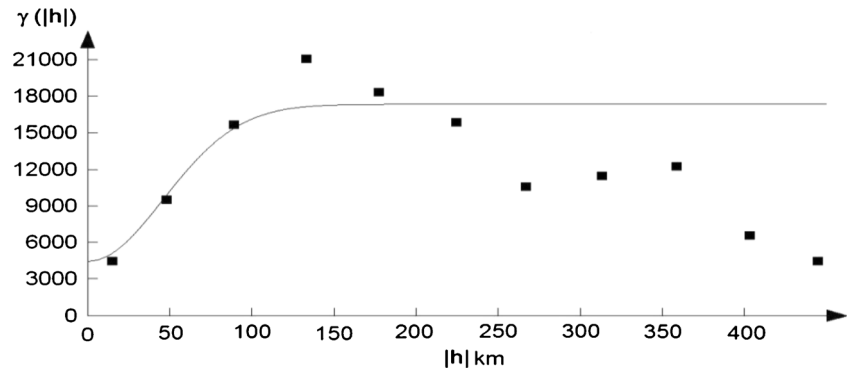
Fig. 6 Adjusted theoretical variogram of mean yearly rainfall according to the east/west direction

Conclusion

The present hydro-meteorological study estimates the precipitation potential of Algeria's eastern high plateau region. The results show the importance of considering the influence of geographical parameters on the spatial rainfall distribution.

Rainfall data provided from rain-gauging station may be inadequate to define and delimit vulnerable areas, and traditional mapping is of little help when the uncertainty associated with the estimated values at unsampled locations is required to support decision-making. In this study, MLR associated to geostatistical methods give a useful tool to generate rainfall map and to assess the uncertainty of the predicted values as an alternative to a detailed investigation. The geostatistical analysis indicates the existence of a drift, supporting the MLR results concerning the existence of a strong north/south rainfall gradient in the study area. The large extent of the study area and its economic importance underline the importance of strengthening and expanding the current climate network. The complexity of rainfall estimation in the region suggests the requirements to consider other variables, such as temperature, air masses movement and land cover vegetation, to better understand the influence of these parameters on rainfall. Rain bands greater than 350 mm are considered as potentially suitable for cereal crops, occupying about 61 % of the total study area.

Moreover, applying this method will not only provide better rainfall estimation but could be used efficiently for
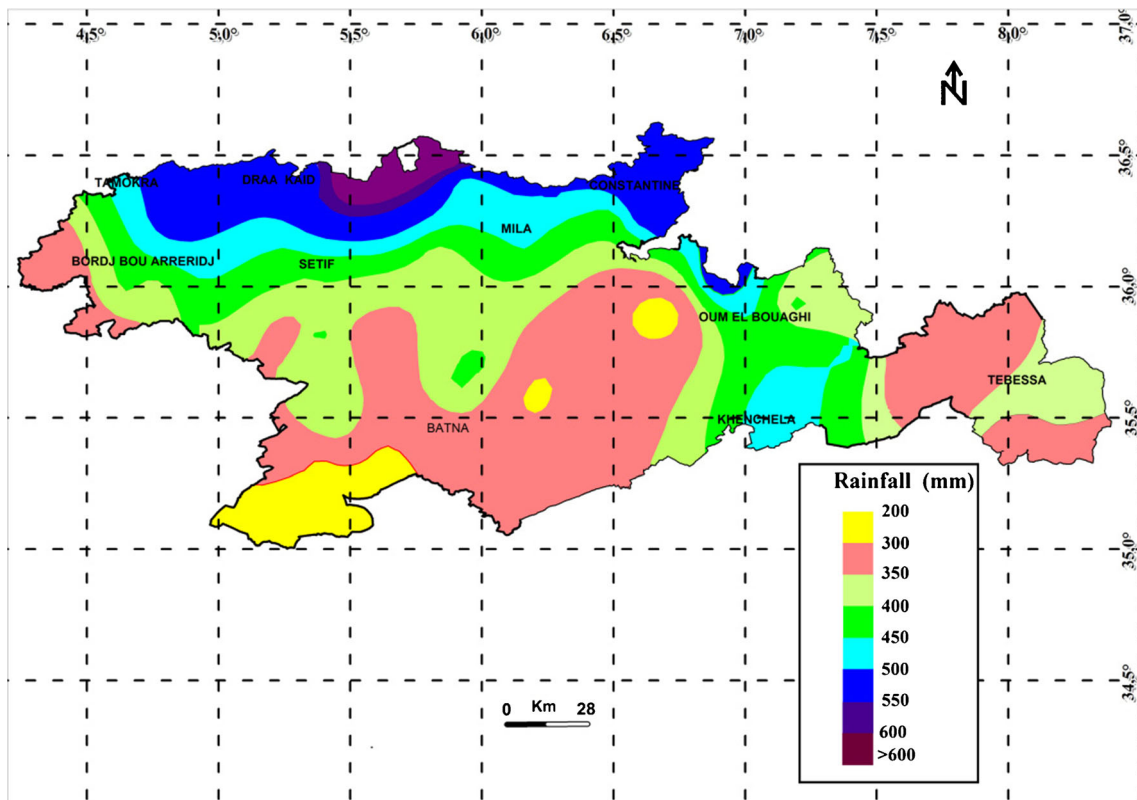

Fig. 7 Spatio-temporal map of rainfall distribution (mm) for the period 1986–2007 in the eastern high plateau region of Algeria

monitoring and development of such areas that are particularly intended for agronomic activities, potentially advocated to the production of winter wheat and barley. The stakeholders must respect this repartition in order to optimize crop yields. Finally, the proposed approach provides a useful tool that could be easily applied to other regions with poor rain-gauging network; it could also be used for other climatic parameters needed for climatology investigation and environmental and agronomic studies.

References

Ahrens B (2006) Distance in spatial interpolation of daily rain gauge data. Hydrol Earth Syst Sci 10: 197–208. www.hydrol-earth-syst-sci.net/10/197/2006/

Arnaud M and Emery X (2000) Estimation et interpolation spatiale, Méthodes déterministes et méthodes géostatistiques. Hermes science Europe

Ashiq M-W, Zhao C, Ni J, Akhtar M (2010) GIS-based high-resolution spatial interpolation of precipitation in mountain–plain areas of upper Pakistan for regional climate change impact studies. Theor Appl Climatol 99:239–253. doi:10.1007/s00704-009-0140-y

Bahlouli F, Bouzerzour H, Benmahammed A (2008) Effets de la vitesse et de la durée du remplissage du grain ainsi que de l'accumulation des assimilats de la tige dans l'élaboration du rendement du blé dur (*Triticum durum* Desf.) dans les conditions de culture des hautes plaines orientales d'Algérie. Biotechnol Agron Soc Environ 1:31–39

Bargaoui Z-K, Chebbi A (2009) Comparison of two kriging interpolation methods applied to spatiotemporal rainfall. J Hydrol 365:56–73. doi:10.1016/j.jhydrol.2008.11.025

Benabdesselam T, Amarchi H (2013) Regional approach for the estimation of extreme daily precipitation on north-east area of Algeria. Int J Water Ress Environ Eng 5(10):573–583. doi:10.5897/IJWREE12.071

Bertrand-Krajewski J-L, Laplace D, Joannis C (2000) Mesures en hydrologie urbaine et assainissement, Paris (France). Tech, ET Doc

Bostana P-A, Heuvelink G-B-M, Akyurekc S-Z (2012) Comparison of regression and kriging techniques for mapping the average annual precipitation of Turkey. Int J Appl Earth Obser Geoinfo 19:115–126. doi:10.1016/j.jag.2012.04.010

Brunsdon C, Mcclatchey J, Unwin D-J (2001) Spatial variations in the average rainfall–altitude relationship in Great Britain: an approach using geographically weighted regression. Int J Climatol 21:455–466. doi:10.1002/joc.614

Bulmer M-G (1979). Principles of statistics. Dover 108–123

Chaumont M. and Paquin C (1971) Carte pluviométrique de l'Algérie du Nord à l'échelle 1/500 000, carte plus notice. Soc Hist Nat. Afrique du Nord, Alger. 25p

Chua S, Bras R-L (1982) Optimal estimators of mean annual rainfall in regions of orographic influences. J Hydrol 57:23–48

Cook R-D (1977) Detection of influential observations in linear regression. Technometrics 19:15–18

Cook E-R, Briffa K-R, Jones P-D (1994) Spatial regression methods in dendroclimatology—a review and comparison of 2 techniques. Int J Climatol 14:379–402

Delhomme J-P (1976) Application de la théorie des variables régionalisées dans les sciences de l'eau. Thèse Doct., Université de Pierre et Marie Curie- Paris. 129p

Delli G, Sarfatti P, Bazzani F (2002) Application of GIS for agro-climatological characterization of northern Algeria to define durum wheat production area. J agri environ inter develop 96:3

Dingman S, Seely D, Reynolds R (1988) Application of kriging to estimate mean annual rainfall in a region of orographic influence. Water Res Bull 24:329–339

Goodale C-L, Aber J-D, Ollinger S-V (1998) Mapping monthly precipitation, temperature, and solar radiation for Ireland with polynomial regression and a digital elevation model. Clim Res 10:35–49

Goovaerts P (2000) Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. J Hydrol 228:11–129

Hamlaoui-Moulai L, Mesbah M, Souag-Gamane D, Medjerab A (2013) Detecting hydro-climatic change using spatiotemporal analysis of rainfall time series in western Algeria. Nat Hazards 65(3):1293–1311. doi:10.1007/s11069-012-0411-2

Hay L, Viger R, Mccabe G (1998) Precipitation interpolation in mountainous regions using multiple linear regressions. Hydrol Water Resour Ecol Headwaters 248:33–38

Hevesi J, Flint A, Istok J (1992) Rainfall estimation in mountainous terrain using multivariate geostatistics: part 1 and 2. J Appl Meteo 31:66–688

Johnson G-L, Hanson C-L (1995) Topographic and atmospheric influences on rainfall variability over a mountainous watershed. J Appl Meteo 34:68–74

Kyriakidis P-C, Journel AG (1999) Geostatistical space-time models. Math Geol 6(31):651–684

Lam N (1983) Spatial interpolation methods. American Cartographer 10(2):129–149

Lloyd C-D (2005) Assessing the effect of integrating elevation data into the estimation of monthly precipitation in Great Britain. J Hydol 308:128–150

Lloyd C-D (2010) Multivariate interpolation of monthly precipitation amount in the United Kingdom. geoENV VII. Geostatistics for Environmental Applications, Quantitative Geology and Geostatistics 16, Springer Science + Business Media. doi:10.1007/978-90-481-2322-33

Marks D, Winstral A, Reba M, Pomeroy J, Kumar M (2013) An evaluation of methods for determining during-storm precipitation phase and the rain/snow transition elevation at the surface in a mountain basin. Adv Water Resour 55:98–110. doi:10.1016/j.advwatres.2012.11.012

Matheron G (1965) Les variables régionalisées et leur estimation. Edition Masson & Cie, Paris 306p

Meddi M, Toumi S (2013) Study of the interannual rainfall variability in northern Algeria. L J E E 23:40–59

Meddi M, Toumi S (2015) Spatial variability and cartography of maximum annual daily rainfall under different return period in the north of Algeria. J Mt Sci 12(6):1403–1421

Miniscloux F, Anquetin S, Creutin J-D (2001) Geostatistical analysis of orographic rainbands. J Appl Meteo 40(11):1835–1854

Myoung-Jin U-M, Hyeseon Y, Chang-Sam J, Jun-Haeng H (2012) Factor analysis and multiple regression between topography and precipitation on Jeju Island, Korea. J Hydrol 410:189–203. doi:10.1016/j.jhydrol.2011.09.016

Naoum S, Tsanis I-K (2004) A multiple linear regression GIS module using spatial variables to model orographic rainfall. J Hydroinf 06(1):39–56

Ouallouche F, Ameur S (2014) Rainfall detection over northern Algeria by combining MSG and TRMM data. Appl Water Sci 6:1–10. doi:10.1007/s13201-014-0204-8

Palecki M-A, Angel J-R, Hollinger S-E (2005) Storm precipitation in the United States. Part I: Meteorol Charact J Appl Meteo 44:933–946

Pannatier Y (1996) Variowin—software for spatial data analysis in 2D. Springer Verlag, notice 91p

Piccini C, Marchetti A, Farina R, Francaviglia R (2012) Application of indicator kriging to evaluate the probability of exceeding nitrate contamination thresholds. Int J Environ Res 6(4):853–862

Prise D-T, McKenney D-W, Nalder I-A, Hutchinson M-F, Kesteven J-L (2000) A comparison of two statistical methods for spatial interpolation of Canadian monthly mean climate data. Agri Forest Meteo 101:81–94

Queney P (1937) Le régime pluviométrique de l'Algérie et son évolution depuis 1850. Météorologie 427-440

Queney P (1943) Les fronts atmosphériques permanents et leurs perturbations. (Travaux de l'Institut de Météorologie, fasc. 3. Alger. pp1–6

Rosenberg M (1969) Hydrologie. Choix d'un modèle régional expliquant la répartition des précipitations annuelles dans l'espace en fonction des facteurs climatiques et topographiques CRAC Sci 761–764

Sangati M, Borga M (2009) Influence of rainfall spatial resolution on flash flood modelling. Nat Hazards Earth Syst Sci 9(2):575–574. doi:10.5194/nhess-9-575-2009

Seltzer P (1948) Le climat de l'Algérie. Inst. Météorol. Phys, Globe, Alger 219p

Sheather S-J (2009) A modern approach to regression with R, chapter 5: multiple linear regression. Springer Science –Business Media 125–149

Skirvin S-M, Marsh S-E, McClaran M-P, Meko D-M (2003) Climate spatial variability and data resolution in a semi-arid watershed, southeastern Arizona. J Arid Environ 54:667–686. doi:10.1006/jare.2002.1086

Slimani M, Cudennec C, Feki H (2007) Structure of the rainfall gradient in the Sahara transition in Tunisia: geographical determinants and seasonality. Hydrol Sci J 52(6):1088–1102. doi:10.1623/hysj.52.6.1088

Smadhi D, Zella L (2012) The pluviometrical deficiencies in the pluvial cereal regions in Algeria. Afric Jour of Agric Resea 7(48):6413–6420. doi:10.5897/AJAR12.1838

Snee R-D (1986) An alternative approach to fitting models when re-expression of the response is useful. J Quality Techno 18:211–225

Subyani A-M (2004) Geostatistical study of annual and seasonal mean rainfall patterns in southwest Saudi Arabia. Hydrol Sci J 49(5):803–817. doi:10.1623/hysj.49.5.803.55137

Subyani A-M, Al-Dakheel A-M (2009) Multivariate geostatistical methods of mean annual and seasonal rainfall in southwest Saudi Arabia. Arab J Geosci 2:19–27. doi:10.1007/s12517-008-0015-z

Touazi M, Laborde J-P, Bhiry N (2004) Modelling rainfall-discharge at a mean inter-yearly scale in northern Algeria. J Hydro 296:179–191. doi:10.1016/j. jhydrol.2004.03.030

Touchan R, Anchukaitis KJ, Meko DM, Sabir M, Attalah S, Aloui A (2011) Spatiotemporal drought variability in northwestern Africa over the last nine centuries. Clim Dyn 37:237–252. doi:10.1007/s00382-010-0804-4

Valipour M (2012) Critical areas of Iran for agriculture water management according to the annual rainfall. Eur J Sci Res 84:600–608

Valipour M (2015a) A comprehensive study on irrigation management in Asia and Oceania. Arch Agron Soil Sci 61(9):1247–1271. doi:10.1080/03650340.2014.986471

Valipour M (2015b) Future of agricultural water management in Africa. Arch Agron Soil Sci 61(7):907–927. doi:10.1080/03650340.2014.961433

Valipour M (2015c) Long-term runoff study using SARIMA and ARIMA models in the United States. Meteorol Appl 22(3):592–598. doi:10.1002/met.1491

Valipour M (2016) Optimization of neural networks for precipitation analysis in a humid region to detect drought and wet year alarms. Meteorol Appl 23(1):91–100. doi:10.1002/met.1533

Valipour M, Ziatabar Ahmadi M, Raeini-Sarjaz M, Ali Gholami Sefidkouhi M, Shahnazari A, Fazlola R, Darzi-Naftchali A (2015) Agricultural water management in the world during past half century. Arch Agron Soil Sci 61(5):657–678. doi:10.1080/03650340.2014.944903

Vargas-Guzman J-A, Myers D-E, Warrick A-W (2000) Derivates of spatial variances of growing windows and the variogram. Math Geol 32(7):851–871

Walpole R-E, Myers R-H, Myers S-L and Ye K (2007) Probability and statistics for engineers and scientists. , 8th ed., Pearson Prentice Hall, Pearson Prentice Hall 791p

Zekai S, Zeyad H (2000) Spatial precipitation assessment with elevation by using point cumulative semivariogram technique. Water Resour Manag 14:311–325