Spatial analysis of annual rainfall using ordinary kriging techniques and lognormal kriging in the Cheliff watershed. Algeria.

Rata M.[1,*], Douaoui A.[2] ,Larid M.[3] and Daouik A[4]

[1] research laboratory PRAVDURN, Department of Biology, Djillali Bounaama University of Khemis Miliana, Ain Defla, Algeria. 44225

[2] research laboratory PRAVDURN, Departement of Science Technology, University Center of Tipaza, Alegria 42000

[3] University abdel hamid Ibn Badis of Mostaghanem, Algeria 27000.

[4] INRA Rabat, Maroc

[*] m.rata@univ-dbkm.dz

ABSTRACT

Spatial analysis of annual rainfall using ordinary kriging techniques is well known to be robust when the data have a normal distribution. But our rainfall data never fully meet these hypotheses. Our study area is centered in the Cheliff watershed, located in northwestern Algeria; It is known by heterogeneous regions in annual rainfall. Data from 58 rainfall stations were also used to interpolate and predict errors by the two kriging techniques. A comparative study of two interpolation methods is presented. Lognormal kriging is very sensitive to extreme values more or less remote from the arithmetic mean (non-normal distribution). The objective of this study is to map annual rainfall and to minimize interpolation errors by the more robust technique. The difference between the accuracy of ordinary and lognormal kriging is well represented on the error estimation map, so the second technique is more robust than the first one in our watershed.

Key words: Cheliff watershed, rainfall, kriging and mapping

Materials and methods

Choice of stations and study period

The spatial and temporal representativeness of the rainfall stations on the study area has a major influence on the reliability of a final map. The rainfall data collected directly from the National Agency for Hydraulic Resources (A.N.R.H.). The heterogeneity of the series of observations thus creates a great problem; more the dissymmetry in the distribution of observation posts and the almost complete absence of data from stations in the high areas during the years 1990 to 2000. Or Algeria experienced a great security problem, particularly in southwestern Watershed. For it. We have limited our study to the northern region where there is a sufficient number of stations and well distributed. The 58 posts selected and spread over 40 years have a series of measures sufficient to carry out this study, that is to say a minimum of 30 years as recommended by the World Meteorological Office (O.M.M.). Some stations have gaps that will be filled using the linear regression method on a monthly scale with base stations (complete and correct measurement).

Comparative analysis of the results of two applied methods

The objective of the comparative analysis is to Establishing rainfall map with a minimum of errors by ordinary kriging (estimated by a linear combination of observations: $\hat{Z}(s) = \sum_i \lambda_i Z(s_i)$. Matheron ,G. ;1962. Due to the asymmetry of data distribution; We have transformed this climatic data into log with the application of box cox technique. Then calculate the original variogram using the algorithm of Guiblin *et al*; 1995. According to Guiblin *et al*. (1995) the variogram of the variable z in the original scale (formula 1),where $y_L(h)$ is the variogram in the logarithmic scale, $\check{Z}_s$ is the sample mean of the variable in the original scale, β is a constant added to the variable in order to avoid zero values for which the logarithm is not defined (1 in the present case), s^2 is the variance of the samples in the original scale,S^2_L, the variance of the samples in the logarithmic scale, and α2 is computed as formula 2 .

$$\gamma_L(h) = ((\beta + \bar{Z}_2)^2 + S^2)(1 - e^{-\alpha^2 \frac{\gamma_t(h)}{\gamma_t^2}}) \quad \text{........formula 1;}$$

$$\alpha^2 = \log(1 + \frac{S^2}{(\beta + \bar{Z}_s)^2}) \quad \text{.........formula 2}$$

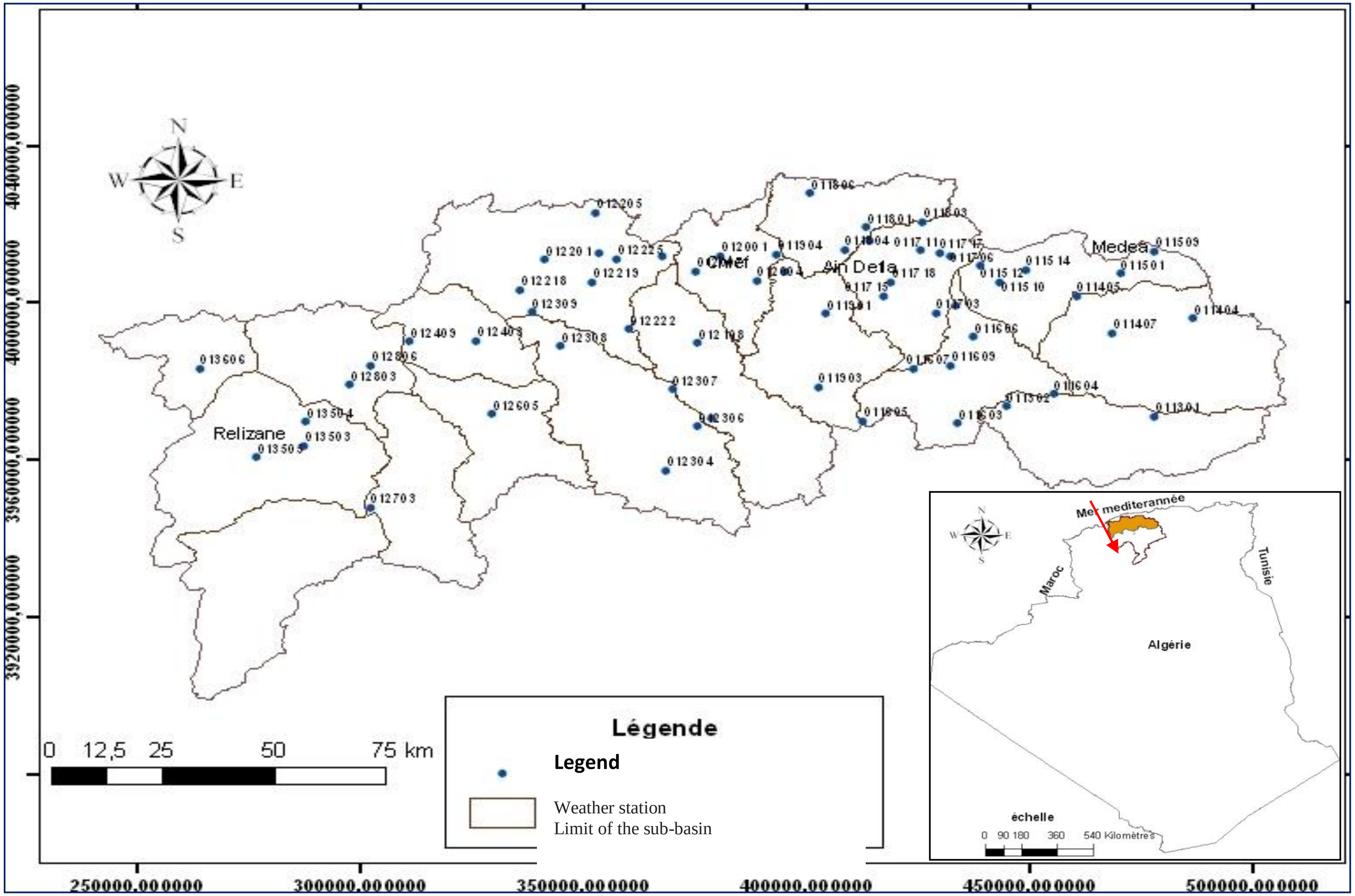

Fig. 1. Geographical location of rainfall stations

Tab 01: Statistical parameters

Statistical parameters	Rainfall (mm)	Ln Rainfall (mm)
Minimum	251,000	5,526
Median	369,710	5,913
Maximum	644,800	6,469
Average	395,010	5,957
CV (standard deviation / mean)	0,210	0,035
Sample variance	7393,940	0,044
Sample standard deviation	85,980	0,209

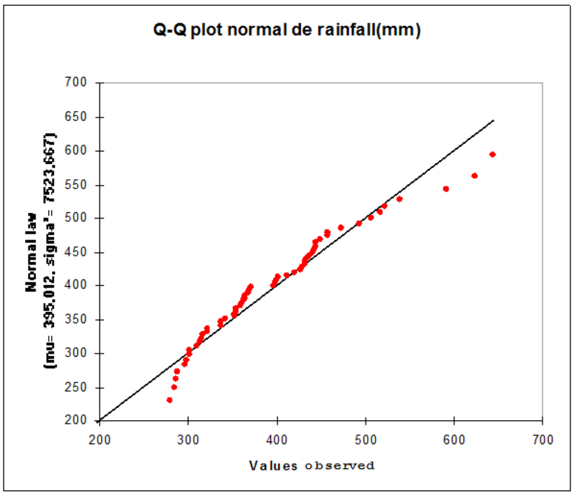

Fig. 2. Shapiro-Wilk Test Curve

Tab 02: Test of Shapiro-Wilk

W (observed value)	0,944
Unilateral p-value	0,010
Alpha	0,05

Results

Statistical parameters

The table (Tab 01) below presents the statistical results calculated on the rainfall data. The coefficient of variation shows that the standard deviation is of the order of magnitude of the mean but indicates variability not significant enough. On the other hand, the coefficient of variation of data transformed into logarithmic is lower, which shows that these data follow a normal distribution.

At the level of significance Alpha = 0.050 we can reject the null hypothesis according to which the sample follows a normal law. In other words, non-normality is significant (Tab 02).

The Shapiro-Wilk test (Royston P; 1982) curve shows that the rainfall data do not follow a distribution Normal (Fig. 2), and at the level of significance Alpha = 0.050 we can reject the null hypothesis according to which the sample follows a normal law. In other words, non-normality is significant (Tab. 2).

When a non linear distribution was found. A form of transformation was sought by applying the most well-known transformation technique is the Box-Cox transformation, which was proposed by Box and Cox. (Box GEP and Cox DR; 1964). It has been found that the logarithmic transformation is ideal for studying this data

.Spatial analysis of rainfall

The spatial estimation of the annual rainfall averages was done by ordinary kriging, which is an interpolation technique developed in the 1950s by the South African geologist Krige. Matheron , G ;1962. The use of our data as part of a Geographic Information System (GIS) enabled us to establish two rainfall maps. In our case, ordinary kriging and ordinary kriging are applied as an interpolation method, and the ArcGis 9.2 software has been used as a GIS.

The variogram established shows continuity in the spatial structure of the rainfall at the scale of the study area (fig. 3). This variogram is characterized by an exponential model and with these parameters: Nugget: 75,72 mm, Range : 44501,3 meters and Sill : 6171,85mm^2. The nugget effect which is very high shows the existence of randomness at the scale of small distances.

On the other hand, the experimental variogram of data calculated after logarithmic transformation (fig. 4) characterized by the same exponential model but with a nugget effect smaller than the first variogram and characterized by these parameters: Nugget : 60mm, Range : 28545meters and Sill : 4071mm^2

Goodness of model fit

The degree of precision of the model was quantified using five statistical indicators. Vicente-Serrano *et al*; 2003 : the coefficient of Determination (R2), where R is the coefficient of correlation between the predicted values and the observed values; (3) the mean square error (RMSE); (4) the absolute mean error (MAE); (5) the mean relative percentage error (MRE%) and (6)the Lin coefficient. The numerical formulas are as follows:

$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(z_i^* - z_i)^2}$ .....3 ; $MAE = \frac{1}{n}\sum_{i=1}^{n}|z_i^* - z_i|$.......4 ; $MRE\% = \frac{1}{n}\sum_{i=1}^{n}\frac{(z_i^* - z_i)}{z_i}100$....5

$CC_{Lin}=2s_{xy}/s^2_{x}+s^2_{y}+(\bar{x}-\bar{y})^2$....6

Z^* : estimated value   Z: Observed value

Note: for the estimated values, the cross-validation method was used to determine these values.

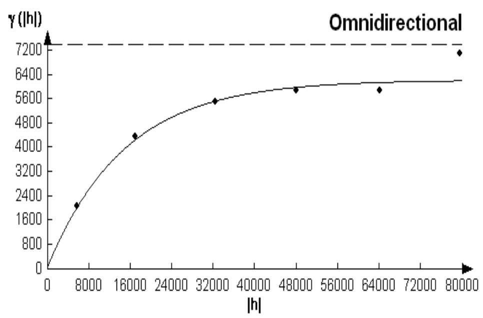

Fig. 3. Experimental variogram of rainfall (original data)

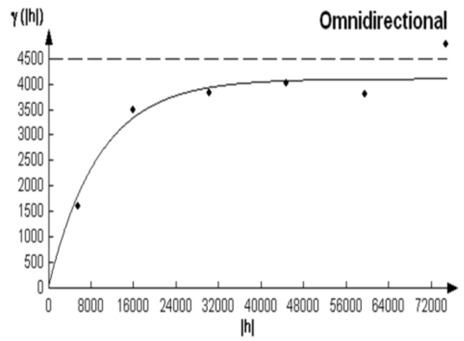

Fig.4Experimental variogram of rainfall (data calculated after transformation) goodness of model fit

Tab 03: Statistics of two interpolation methods (Log_KO and KO)

variable	Method	pearson correlation coefficient	RMSE	MRE (%)	ME	Lin coefficient of concordance
average annual rainfall	OK	0,693	62,34	0,00377	6,31	0,65
	log_OK	0,696	62,08	0,00369	6,48	0,648

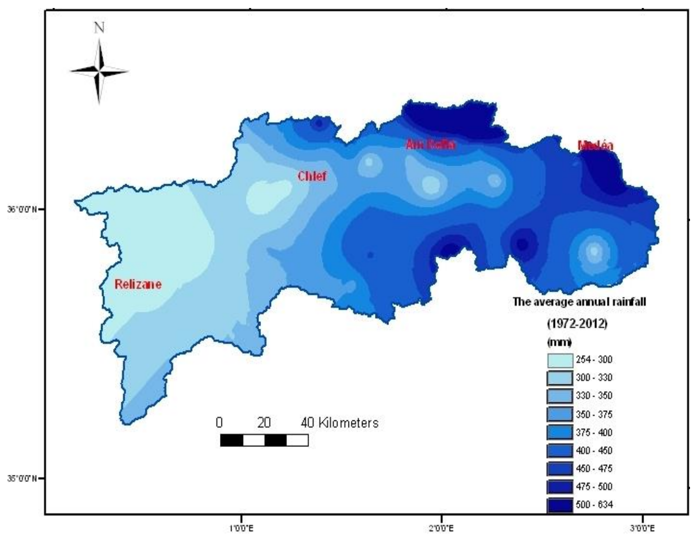

Fig. 5 Map of annual mean rainfall (estimated by KO)

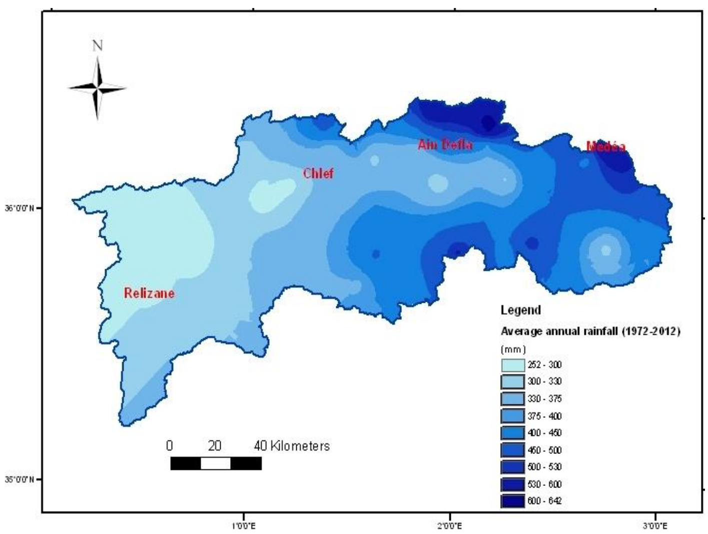

Fig. 6 Map of average annual rainfall (estimated by Log KO)

The results shown in Table 03 indicate that the ordinary kriging Log method is optimal because the RMS (62.08 mm) is lower compared to RMS (62.34 mm) than is determined by ordinary kriging, despite the difference Is lower in our case study, but it has been shown that the logarithmic transformation minimized an estimate error. For this purpose, the Pearson correlation coefficient is higher for the kriging method compared to ordinary kriging. Other clues such as Lin's coefficient of concordance show that the agreement is passable (less than 0.65). Among the results shown in the above, it is clear that the ordinary Kriging Log method works better than ordinary kriging.

Conclusion

In this study, two interpolation methods were tested to obtain the best distribution of annual rainfall patterns for a region centered in the major watershed in Algeria. It is interesting to note that the Log_KO interpolation method is the best when compared to the Ko method, when the climate data does not follow a normal distribution. In addition, it was recorded that the Log Ko method takes into considerations the extremely high data in the spatial estimation. On the other hand the homogeneous region in the plains at the annual average rainfall is less than 350 mm; the spatial estimate is almost identical. So the Log_Ko interpolation method is best when the rainfall data is heterogeneous and does not follow a normal distribution.

Bibliographic reference

- Box GEP, Cox DR (1964) An analysis of transformations. Journal of the Royal Statistical Society,

  Series B 26: 211–252.

- GUIBLIN *et al*1995 - Analyse structurale de données à distribution dissymétrique, exemple du

  hareng écossais. Cahiers de Géostatistique, Compte-rendu des journées de Géostatistique 1.5-16

  juin 1995, Paris, E.N.S.M.P

- Matheron , G. (1962).Traitée de géostatistique appliquée, Tome I .Mémoires du Bureau de

  Recherches Géologiques et Minières, No.14. Editions Technip, Paris.

- Royston P. (1982). An Extension of Shapiro and Wilk's W Test for Normality to Large Samples.

  Applied Statistics, 31, pp. 115–124.

- Vicente-Serrano SM, Saz-S´anchez MA, Cuadrat JM. 2003. Comparative analysis of interpolation

  methods in the middle Ebro Valley (Spain): application to annual precipitation and temperature.

  Climate

Mapping rainfall

The two annual average rainfall maps (Fig 5 and 6) for the period 1972 to 2012 show that the most temperate regions are located in the massifs of the north and the massifs of South, at altitudes is more than 500 meters. At the same time, this cartographic result confirmed that the trend of the rainfall regime is decreasing from East Green to West. The differences between these two maps are well noticed when the rainfall is greater than 500 mm, that is to say in the mountainous regions, in contrast in the plain is almost the same average.