Analele Universităţii din Oradea, Seria Geografie Year **XXVI**, no. **1/2016** (June), pp. **107-117**
ISSN **1221-1273**, E-ISSN **2065-3409** Article no. **261110-707**

ESTIMATION AND MAPPING OF EXTREME RAINFALL IN THE CATCHMENT AREA OF BATNA (ALGERIA)

Guellouh SAMI*
Doctorant, Laboratoire LRNAT, University of Batna, 05 Avenue Chahid Boukhlouf 05000, Algeria
e-mail: guellouhsami@gmail.com

Dridi HADDA
Pr. Laboratoire LRNAT, University of Batna, 05 Avenue Chahid Boukhlouf 05000, Algeria
e-mail: hadda.dridi@gmail.com

Kalla MAHDI
Pr. Laboratoire LRNAT, University of Batna, 05 Avenue Chahid Boukhlouf 05000, Algeria
e-mail: m_kalla1@yahoo.fr

Abstract: Statistical estimation of rainfall associated with extreme events is of major interest for hydrologists in terms of risk prevention. Comprehending the spatial distribution of extreme rainfalls that cover the entire catchment area, the impluvium, of Batna, requires as a first step a frequency analysis of annual maximum daily rainfall time series with the application of empirical distributions, namely the GEV distribution, the Gumbel distribution and the log-normal distribution. This has allowed us to estimate the quantiles of extreme rainfall with return periods of 5, 10, 20, 50 and 100 years for ten rainfall stations. Subsequently, this has allowed us to map the quantiles matching the centennial return period using three types of interpolations.

Key words: extreme rainfall, frequency analysis, interpolation

* * * * * *

INTRODUCTION:

The extreme rainfall phrase characterises high intensity rainfall events occurring on a short time frame, and of low occurrence, both in time and in space (Berolo et al., 2003, p. 1). They are the main cause of devastating floods that generate significant, and often permanent, damage.

The estimation of rainfall associated with extreme events is a topic of growing interest in the field of water science (Hubert et al., 2004, p. 718). This is especially true in the field of structures design aiming at providing protection against flooding (water infrastructure). Studies of extreme rainfall events have great relevance for water resources management, as they provide insight into the understanding of the hydrological behavior of a given watershed under the flooding point (Samuel et al., 2015, p. 123-133). We have started off with a frequency analysis of the data of maximum daily rainfall recorded by ten rainfall stations, figure 1 which purpose is to

* Corresponding Author

study past events in order to define the future occurrence probabilities (Degla et al., 2014, p. 137-148). This analysis was carried out on all series of rainfall to ensure independence, homogeneity and stationarity of the latter. We relied on several distributions for estimating rainfall quantiles. The choice of the mathematical distribution best fitting our data was based on suitability tests. Several statistical tests are available for determining the best adjustment, and we have adopted two of them, the Chi-Square χ^2 and both the criteria known as the Bayesian Information Criterion BIC put forward by Schwarz in 1978 and the Akaike Information Criterion AIC developed by Akaike in 1974. This approach is completed by a spatial interpolation of rainfall that provides rainfall ratios in sites for which there is no recorded data. We used the inverse distance weighting IDW interpolation, the spline interpolation and the ordinary kriging OK methods.

Based on the result of the Root Mean Squared Error RMSE, interpolation by the spline method seems to be the most reliable.

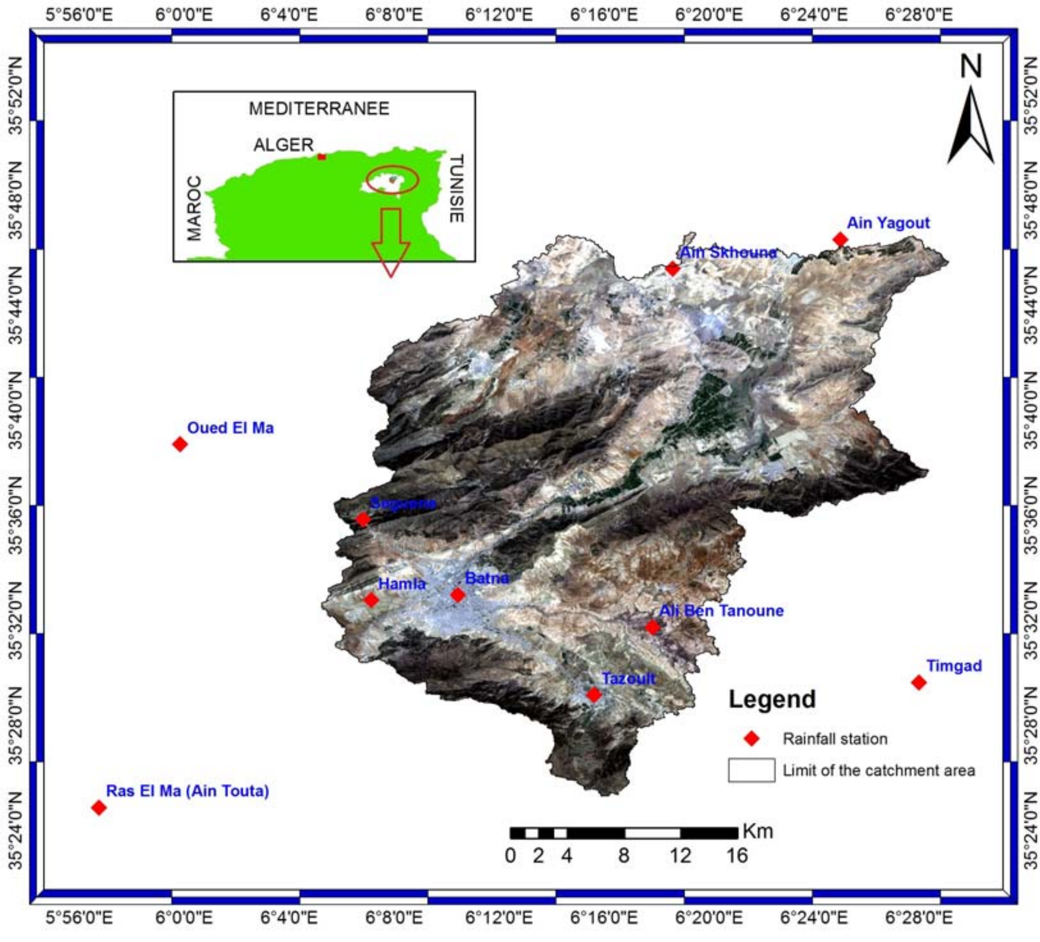

Figure 1. Study area
(Source: Guellouh Sami)

STUDY AREA (REPRESENTATION AND INFORMATION GATHERING)

The catchment area of Batna figure 1, located in eastern Algeria, covers an area of 802.68 km², between latitudes 35°25' and 35°47' North and between longitudes 6°5' and 6°29' East. The climate of the region is classified as semi-arid, characterized by irregular rainfall. The series of daily rainfall recorded at each rainfall station encompass a period of 44 years from 1969 to 2012, with the exception of the Ain Skhouna series that span a period of 42 years from 1971 to 2012, ANRH. The samples are created by extracting the maximum daily value of each year. The annual maximum values method appears to be generally preferred to the Peak Over Threshold POT method, both by researchers and by forecasters (Cunnane,1987, p. 49-95).

The statistical parameters of the data used are calculated and summarized in table 1.

Table 1. The statistical parameters of the data used

The series of daily rainfall	minimum value mm	maximum value mm	The average value mm	Coefficient of asymetrie CS	Coefficient of variation CV
Tazoult	20	94	46.3	0.903	0.423
Batna	12	64	36.7	0.293	0.366
Seguene	11	129	47.5	1.08	0.524
Hamla	10.1	94	34.6	1.4	0.522
Ben Tanoune	10	101	37.3	1.51	0.455
Oueds El Ma	12	94	38	1.2	0.45
Ras El Ma	10	94	36.9	1.23	0.502
Ain Yagout	11	120	39	1.8	0.48
Ain Skhouna	11	100	43	0.87	0.5
Timgad	13	93	35	1.5	0.41

MATERIALS AND METHODS

Frequency analysis

The stationary, independence and homogeneity tests

Frequency analysis requires first to appreciate the ability of the series to have a distribution function using Kendall stationarity tests, Aka et al., 1996, Wald-Wolfowitz runs test for randomness, Hache et al., 1999 and Wilcoxon Siegel signed-rank test, 1956 (Alain et al., 2015, p. 121-133).

The Kendall stationarity test quoted by (Manohar et al., 2005) is a rank correlation test, used to detect trends in the series (Yue et al., 2004, p. 21-38). The Wald-Wolfowitz test is useful for checking the existence of a sequential dependence within the observations that would lead, if proven, to define its type and level before continuing the study of the frequencies process. These statistical tests essentially allow the assessment of the representativeness of the distributions observed compared to known population values and are used to assess the significance of the observations (Albert et al., 2007, p. 49-67).

The Wilcoxon test is used to confirm the homogeneity of the data series to be analysed.

The distribution patterns of the maximum annual rains

A 1989 survey of the WMO involving 55 agencies from 28 countries reveals that 52% of them are accustomed to use a distribution, specifically, the generalized extreme value (GEV) distribution as the main reference. This distribution recognizes as specific cases, the Gumbel, Fréchet and Weibull distributions. Some 31% of agencies use either a Pearson III distribution P3 a Log-Pearson type III distribution LP3 or a Pearson type III LN distribution. Furthermore, (Rossi et al., 1984) have noticed that the two-component extreme value TCEV distribution is better fitted to series of daily maximum rainfall recorded than more traditional distributions are.

We retained the law of GEV, the Gumbel distribution and the Log-Normal.

The Generalized Extreme Values (GEV) distribution

The GEV distribution is a statistical method widely used in meteorology to identify rainfall and wind extreme values (Coles, 2001, p. 23-31). This distribution allows you to adjust a regression curve according to the frequency of past events and predict the occurrence probability of a major event over a return period longer than the available range of measurements (Fallot, 2013, p. 224-229).

$$F(X) = \exp\left[-\left(1 - K\left(\frac{x-\mu}{\alpha}\right)\right)^{1/K}\right]$$

Where: α, μ and K are respectively the scale, location and shape parameters.

The Gumbel distribution

The Gumbel distribution is a special case of the GEV distribution also known as K=0. This statistic is a frequency distribution model often used to describe the statistical behaviour of extreme values (Wilcoxon, 1945, p. 21-38). This distribution is often used because it allows a linear fit from a double exponential distribution according to a method developed by Gumbel 1958 and illustrated in (Fallot et al., 2013).

$$F(x) = \exp\left[-\exp\left(\frac{x-\mu}{\alpha}\right)\right]$$

Log-Normal distribution

It is also known as the Galton distribution. It is believed that a statistical series better fits with a log-normal distribution when the logarithm of a random variable X is normally distributed. Several studies have rightly concluded that the best fit for arid and semi-arid hydrological regimes is obtained by the log-normal distribution (Dubreuil, 1974, p. 216).

The Formula of the log-normal distribution is as follows:

$$F(x) = \frac{1}{x.\sigma\sqrt{2\pi}} \exp\left[-\frac{(\mathrm{Ln}\, x-\mu)^2}{2\sigma^2}\right]$$

Where: μ is the mean of ln X and σ is the standard deviation of ln X.

Estimation of the distribution function parameters

There are various methods of parameters estimation such as the method of moments, maximum likelihood and L-moments. We have estimated the distribution function parameters using the maximum likelihood method type. This method consists of determining the parameters of the distribution chosen in order to obtain the most reliable possible sample (Roche, 1963, p. 196, 330).

The validity of the models

There are many fitting tests meant to identify the most appropriate adjustment distribution for each range. We chose the Chi-square χ^2 test, which is a first selection mode based on the comparison of the suitability of these distributions (Albert et al., 2007, p. 49-67).

It is difficult to graphically identify the most appropriate distribution, which led us to involve both the Bayesian Information Criterion BIC proposed by Schwarz 1978 and the Akaike Information criterion AIC formulated by Akaike, 1974.

The mathematical formulas (BIC) and (AIC) are:

BIC = -2 log (L) +2 k log (N).

Where: L is the likelihood, K is the number of parameters and N is the sample size.

AIC = -2 log (L) +2(k).

Where: L is the likelihood and K is the number of parameters.

Interpolation

The interpolation term is used to describe all the restitution methods of continuous spatial fields from a set of points (Daniel et al., 2009, p. 98). It is a statistical function that estimates the values of a phenomenon at non-sampled locations.

Three methods were used, both deterministic and probabilistic (stochastic).

IDW i.e. Inverse Distance Weighting

It was developed in the USA by the National Weather Service in 1972 and is classified as a deterministic mathematical interpolation technique (Zaninetti et al., 2005, p. 675-699).

IDW is a simple algorithm that gives more weight to nearby values than to far ones. Shepard 1968. This interpolation method ensures the allocation of estimated values on a grid, calculated by a weighted average of the distance between known values (Lu et al., 2008, p. 1044-1055 and Soenario et al., 2010, p. 4-31).

Spline i.e. optimized local polynomial method

This is an interpolation method in ensures that values are determined using a mathematical function meant to reduce the overall surfaces curvature and renders a smooth surface by going through the input points in a precise manner (ESRI, 2010).

Ordinary kriging, Cressie 1988

Also known as a stochastic geostatistics method; it is regularly used to characterize hydro-meteorological and climatic variables, within wide ranging space and time scales (Delhomme, 1976). It takes into account data point covariance by integrating the spatial dispersion of these points (Eric et al., 2009, p. 4-6).

RESULTS AND DISCUSSION:

The results of the Wald-Wolfowitz runs test indicate that the observations are independent with a 5% threshold for all stations except the Hamla and Ain Yagout stations where observations are accepted with a 1% threshold.

The other two tests are valid for the observations of all stations with a 5% threshold.

Results of both tests

We have calculated the values of both tests using the HYFRAN (Hydrological Frequency Analysis) software developed by INRS-EAU Canada.

Table 2. Results of test Chi-square χ^2

The series of daily rainfall	statistical distribution		
	Log-Normale	**Gumbel**	**GEV**
Tazoult	6.60	5.77	8.28
Batna	6.32	4.27	4.68
Segeune	5.09	7.14	4.68
Hamla	9.18	7.95	4.68
Ben Tanoune	10.82	5.09	5.09
Oueds El Ma	7.95	5.91	7.14
Ras El Ma	5.50	3.86	4.27
Ain Yagout	3.86	2.64	3.86
Ain Skhouna	1.81	2.95	2.95
Timgad	6.73	2.64	2.64

Table 3. Results of test of both criteria AIC and BIC

The series of daily rainfall	criteria	Log-normal	Gumbel	GEV
Tazoult	BIC	391.617	388.624	388.777
	AIC	386.264	385.056	383.425
Batna	BIC	360.401	360.003	361.924

	AIC	356.833	356.435	356.572
Seguene	BIC	406.241	409.353	405.691
	AIC	402.673	404.001	402.123
Hamla	BIC	375.710	372.804	371.706
	AIC	370.358	369.236	368.138
Ben Tanoune	BIC	370.415	369.388	373.098
	AIC	366.847	366.847	367.745
Oueds El Ma	BIC	374.193	373.821	377.553
	AIC	370.625	370.252	372.201
Ras El Ma	BIC	377.648	377.517	381.139
	AIC	374.080	373.949	375.786
Ain Yagout	BIC	375.081	374.828	378.300
	AIC	371.512	371.260	372.948
Ain Skhouna	BIC	376.250	377.474	379.949
	AIC	372.775	373.999	374.736
Timgad	BIC	357.034	356.792	360.529
	AIC	353.466	353.223	355.176

The choice of distribution and calculated quantiles

Based on the visual analysis of the adjustment of the distributions and the results of both tests, we have been able to choose the right distribution data of each series.

The results at different return periods of extreme rainfall events for all rainfall stations are summarized in table 4.

Table 4. The results at different return periods of extreme rainfall events

The series of daily rainfall	Statistical distribution	Parameters	5 years	10 years	20 years	50 years	100 years
Tazoult	Gumbel	u=37.454 α=14.815	59.7	70.8	81.5	95.3	106
Batna	Gumbel	u= 30.198 α=11.8852	48	56.9	65.5	76.6	84.9
Seguene	GEV	α =18.589 k=-0.03525 u= 36.043	64.7	79.6	94.3	114	129
Hamla	GEV	α =12.215 k= -0.1035 u= 26.153	46	57.1	68.6	84.9	98.1
Ben Tanoune	Gumbel	u=29.8564 α =12.8092	49.06	58.7	67.9	79.8	88.7
Oueds El Ma	Gumbel	u= 30.5098 α =13.4468	50.7	60.8	70.4	83	92.4
Ras El Ma	Gumbel	u= 28.7063 α =13.8735	49.5	59.9	69.6	82.8	92.5
Ain Yagout	Gumbel	u= 31.1324 α =13.4394	51	61	71	84	93
Ain Skhouna	Log-Normale	mu : 3.62975 σ =0.5315	59	74.5	90.4	112	130
Timgad	Gumbel	u=28.6737 α =11.2371	45.5	54	62	72.5	80.4

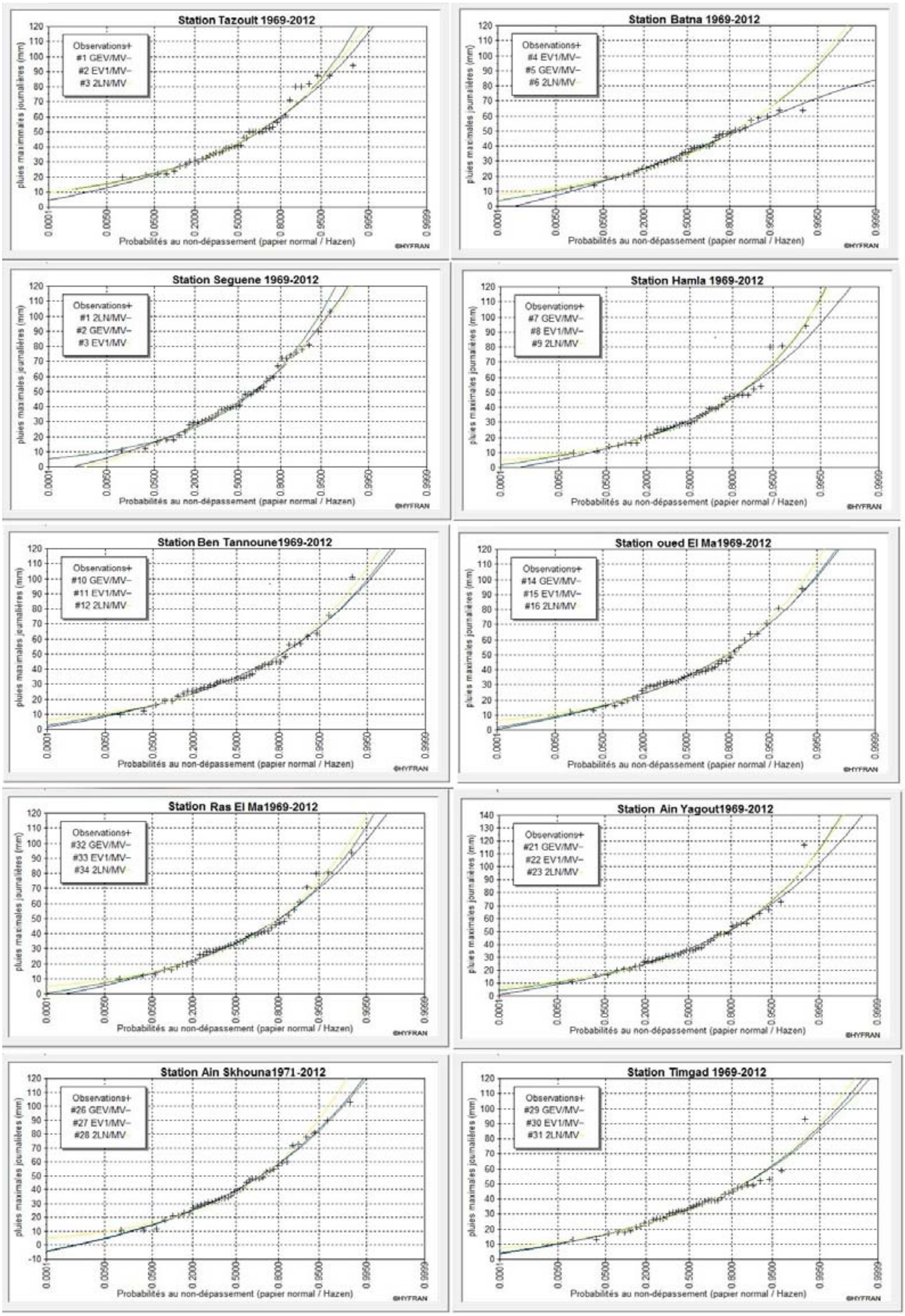

Figure 2. Graphic results
(Source: Hyfran software)

The results display the prevalence of the Gumbel distribution that has been adopted as an adjustment model for extreme values time series in many studies worldwide (Koutsoyiannis, 2004, p. 148-157). In Algeria, this distribution is favoured by most official meteorological services, to describe the frequency of extreme precipitation. It has the advantage of being well known by engineers who use it as part of their work on the reliability of water infrastructure (Benkhaled, 2007, p. 83-91).

The interpolation of estimated rainfall extremes

Direct rain measurements at ground level from a network of rain gauges are punctual. They primarily indicate the volume of rain collected at a specific point. This discrete spatial information is insufficient, or not easily visualised, in particular when it comes to interpreting extreme rainfall events (Comby, 1998, p. 668).

Therefore, these values need to be spatially interpolated in order to evaluate the surface rain in a precise manner (Laborde, 2000, p. 27).

Using the Arc Gis software extension called Geostatistical Analyst, developed by the ESRI (Environmental Systems Research Institute), we have been able to perform three types of interpolation (IDW, Spline and ordinary kriging) we have taken the values of rainfall centennial return as sample.

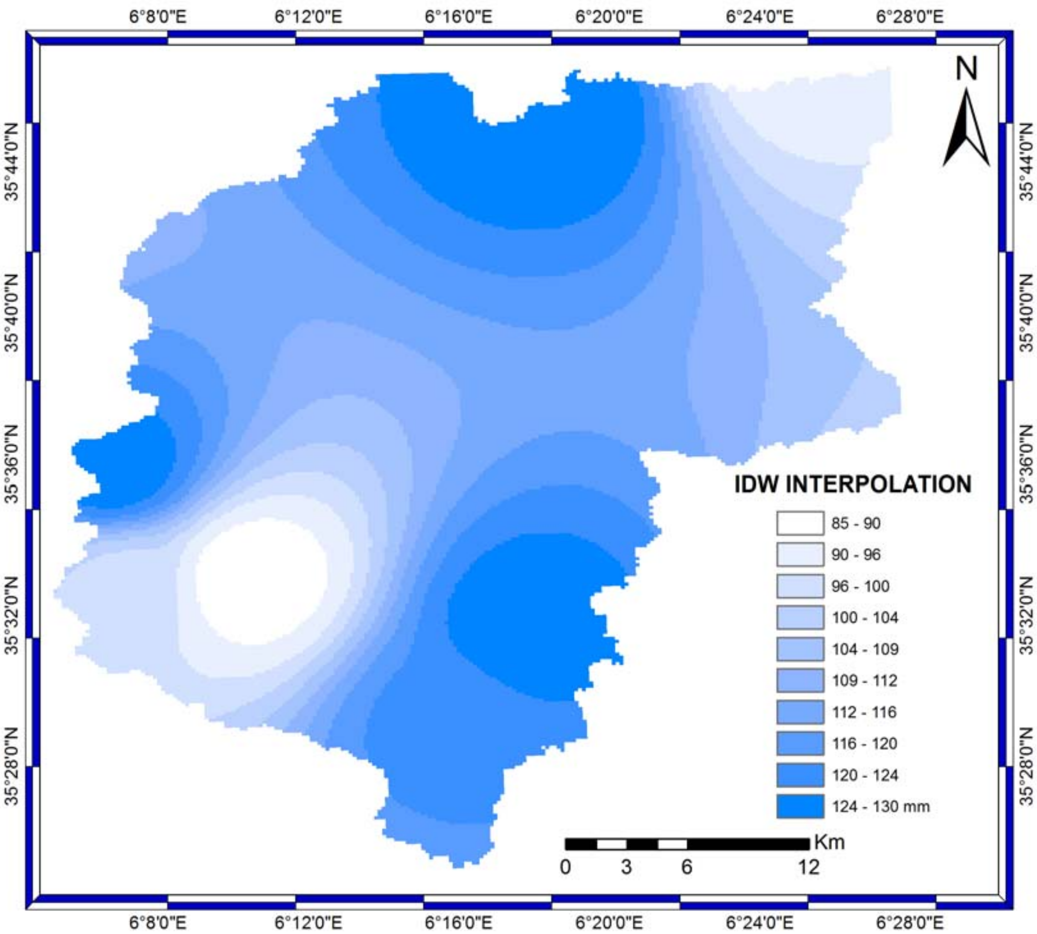

Figure 3 IDW interpolation
(Source: Guellouh Sami)

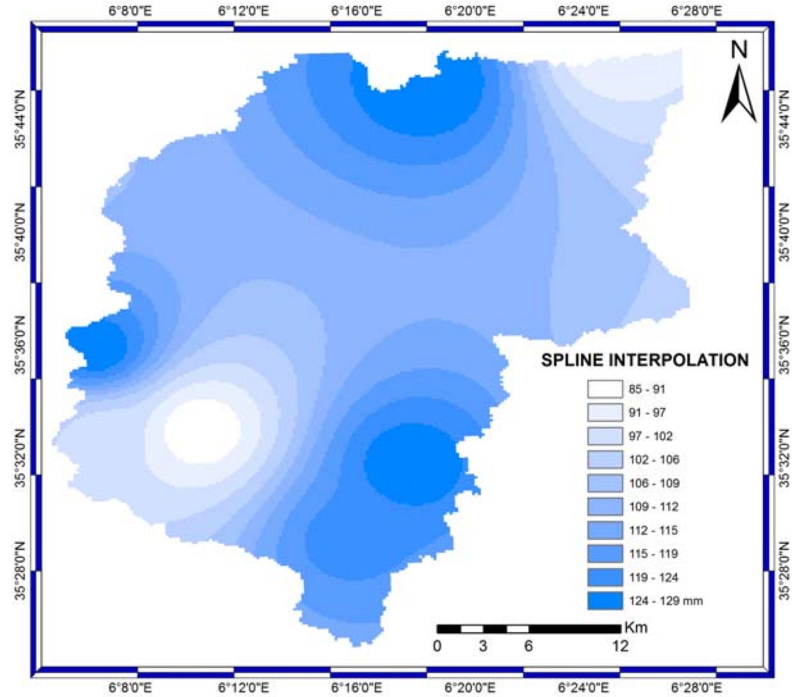

Figure 4 SPLINE interpolation
(Source: Guellouh Sami)

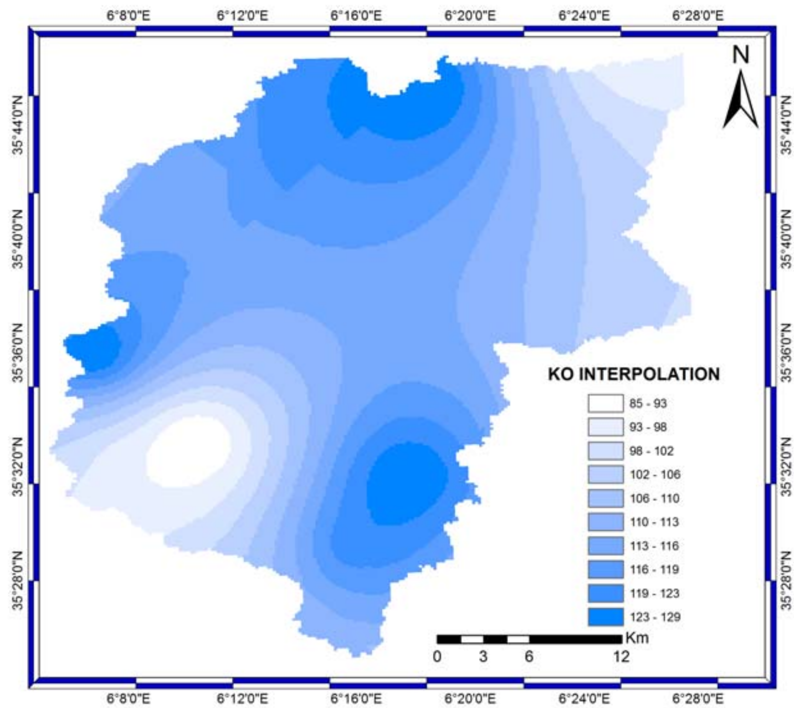

Figure 5. KO interpolation
(Source: Guellouh Sami)

The best data interpolation method is the one that matches the lower value the Root Mean Squared Error RMSE. Geostatistical Analyst provides a cross-validation tool that calculates the error of the estimated values to assess and compare the appropriateness of the methods used.

Table 5. Value of Root Mean Squared Error RMSE
(Source: Guellouh Sami)

Interpolation Methods	the Root Mean Squared Error RMSE
IDW	23.63
KO	25.54
SPLINE	23.11

The spatialisation of rainfall data using the Spline interpolation figure 4 seems to be the most reliable method when compared to other methods and shows a better RMSE result. The statistical methods demonstrated the spatial dimension of extreme rainfall, showing variability and uneven distribution of rainfall in the catchment area of Batna.

CONCLUSION

In this study, we have first conducted a frequency analysis of extreme precipitation data from daily data of ten rainfall stations.

This analysis allowed the estimation of extreme precipitation quantiles for different return periods in order to determine the proper adjustment law for each data series.

Ultimately, the interpolation has provided results for sites where there are no values from sampled data (measuring stations). The comparison of different interpolation methods using a cross-validation tool has shown the relevance of the Spline method compared to other methods.

This mapping of extreme precipitation can be a valuable source of information and a decision-support tool for designers of water infrastructure, thus ensuring protection against floods.

REFERENCES

Alain I., Ague Afouda Abel (2015), *Analyse fréquentielle et nouvelle cartographie des maxima annuels de pluies journalières au Bénin,* International journal of biological and chemical sciences [online] 2015 vol 9 n° 1, 121-133.

Albert G. Bi Tie, Brou Konan., Yao Telesphore Brou., Savane Ssiaka., Vamoryba Fadika, Bernard Srohourou (2010), *Estimation des pluies exceptionnelles journalières en zone tropicale,* Cas de la Côte d'Ivoire par comparaison des lois Log-Normale et de Gumbel Hydrological Sciences Journal. [Online] 2010 vol 52 n° 1, 49-67.

Benkhaled A. (2007), *Distributions statistiques des pluies maximales annuelles dans la région du Cheliff, comparaison des techniques et des résultats*, Courrier du Savoir. [Online] 2007 n°08, 83-91.

Berolo W., Laborde J. P. (2003), *Statistiques des précipitations journalières extrêmes sur les Alpes-Maritimes*, Equipe de Gestion et Valorisation de l'Environnement de l'UMR 6012 espace, Université de Nice, France 1 p.

Coles S. (2001), *An Introduction to Statistical Modeling of Extreme Values*, Springer Verlag, [Online] 2001 vol 211, 23-31.

Comby J. (1998), *Les paroxysmes pluviométriques dans le couloir Rhodanien,* Thèse de géographie et aménagement, Université Jean Moulin Lyon III, 668 p.

Cunnane C. (1987), *Review of Statistical Models for Flood Frequency Estimation*, Inst, Symp on Flood Frequency and Risk Analysis. Baton Rouge. La., Publ. In Singh, V.P. (Ed.), Hydrology Frequency Modelling, Reidel Publ. CO. Dordrecht: 49-95 p.

Daniel Joly, Brossard Thierry, Herve Cardot, Cavailhes Jean, Cavailhes Jean, Hilal Mohamed, Wavresky Pierre (2009), *Interpolation par régressions locales, application aux précipitations en France*, L'Espace géographique vol [Online] 2009 38 n°2, 98.

Degla K. H., Tchibozo A. E., Vissin E., Houssou C. (2014), *Analyse fréquentielle des évènements hydro-pluviométriques extrêmes dans le bassin de la Sota au Bénin*, Afrique SCIENCE [Online] 2014 vol 10 n° 2, 137-148.

Delhomme J. P. (1976), *Application de la théorie des variables régionalisées dans les sciences de l'eau (Variabilités partiale des grandeurs hydro-climatiques et hydrogéologiques & précision de leurs connaissances),* Thèse de docteur Ingénieur, université Pierre et Marie Curie, Paris-6.

Dubreuil P. (1974), *Introduction à l'analyse hydrologique,* Ed. Masson-ORSTOM, Paris p. 216.

Eric Bernard., Tolle Florian., Griselin Madeleine., Laffly Dominique, Christelle Marlin (2009), *Quantification des hauteurs de neige et des températures de l'air à la surface d'un glacier: du terrain à l'interpolation, confrontation de méthodes*, Neuvièmes Rencontres de Théo Quant Besançon, 4-6 p.

ESRI (2010), *Arc GIS 10 Help. Fonctionnement de l'outil Spline*. http://pro.arcgis.com/fr/pro-app/tool-reference/3d-analyst/how-spline-works.htm.

Fallot J. M. (2013), *Evaluations des différentes formules de la loi de Gumbel pour l'estimation des précipitations journalières extrêmes en Suisse*, XXVIème colloque de l'Association Internationale de Climatologie. Institut de Géographie, Université de Lausanne, CH-1015 Lausanne. 224.229 p.

Fallot J. M., Hertig J. A. (2013), *Détermination des précipitations extrêmes en Suisse à l'aide d'analyses statistiques et augmentation des valeurs extrêmes durant le 20ème siècle*, Mém. Soc. Vaud. Sc. Nat. [Online] 2013 vol 25, 13-24.

Hache M., Perreault L., Remillard L., Bobee B. (1999), *Une approche pour la sélection des distributions statistiques*, application au bassin hydrographique du Saguenay Lac St Jean. Revue Canadienne de Génie Civil, [Online] 1999 vol 26 n°2, 216-225.

Hubert Onibon., Taha B. M., Ouarda J., Marc Barbet, Andre S. T., Hilaire, Bernard Bobee, Pierre Bruneau (2004), *Analyse fréquentielle régionale des précipitations journalières maximales annuelles au Québec, Canada*, Journal des Sciences Hydrologiques. [Online] 2004 vol 49 n° 4, 718.

Javier Polo, Ferrer (1992), *Analyse statistique de pluies maximales journalières, Comparaison de différentes méthodes et application au bassin Guadalhorce (Espagne) Hydrologie continental*, [Online] 1992 vol 7, n° 1, 23-31 p.

Koutsoyiannis D. (2004), *Exploration of long records of extremes rainfall and design rainfall inferences*, Proceedings of the International conference of British Hydrological Society, Imperial College London, 148-157 p.

Laborde J. P. (2000), *Méthodes d'interpolation et géostatistiques pour la cartographie automatique*, Université de Nice, Sophia Antipolis, CNRS – UMR 5651 « Espace », 27 p.

Lu Gy, Wong DW (2008), *An adaptive inverse-distance weighting spatial interpolation technique*, Comput Geoscience. [Online] 2008 vol 34 n° 9, 1044-1055.

Manohar A., Goel N., Ksingh P. (2005), *Evaluation of temperature trends over India*, Hydrological Sciences, [Online] 2005 vol 50 n°1, 81-94.

OMM (1989), *Analyse statistique de pluies maximales journalières Comparaison de différentes méthodes et application au bassin Guadalhorce (Espagne) Hydrol*, Connent ml. 7, no 1, 1992.

Renard Florent, Comby Jacques (2007), *Caractérisation de l'alea pluviométrique en milieu urbain a partir d'interpolations spatiales. Le cas du grand LYON,* Revue Climatologie. [Online] 2007 vol 4, 133-141 p.

Roche M. (1963), *Hydrologie de surface*, Office de la recherche scientifique et technique outre-mer (PARIS). Gauthier - Villars éditeur PARIS.196, 330 p.

Rossi F., Fiorentino M., Versace P. (1984), *Component Extreme Value Distribution for Flood Frequency Analysis*, Water Resources. [Online] 1984 vol 20 n°7, 847-856.

Samuel Beskow., Tamara L., Caldeira Carlos Rogério de Mello., Lessandro C. Faria, Hugo Alexandre S. Guedes (2015), *Multiparameter probability distributions for heavy rainfall modeling in extreme southern Brazil*, Journal of Hydrology: Regional Studies. [Online] 2015 vol 4 (2015) 123–133.

Shepard D., (1986), *A two-dimensional interpolation functions for irregularly-spaced data*, Proceedings of the 1968 ACM National Conference, 517–524 p.

Siegel S. (1956), *Non-Parametric Statistics for the Behavioral Sciences*, McGraw-Hill: New York.

Soenario I., Sluiter R. (2010), *Optimization of Rainfall Interpolation*, Interne rapport, De Bilt, 4-31 p.

Wilcoxon F. (1945), *Individual comparisons by ranking methods Biometrics 1*. 80-83 p.

Yue S., Pilon P. A. (2004), *Comparison of the power of the t test, Mann–Kendall and bootstrap tests for trend detection*, Hydrological. Sciences Journal. [Online] 2004 vol 49 n°1, 21–38 .

Zaninetti J. M. Colten C. E. (2012), *Shrinking New Orleans Post-Katrina population adjustments*, Urban Geography. [Online] 2012, vol 33 n°5, 675-699.

Submitted:	Revised:	Accepted and published online
December 22, 2015	May 19, 2016	June 01, 2016