Larhyss Journal, ISSN 1112-3680, n°43, Sept 2020, pp. 57-72

EXPLORATION OF MAXIMUM LIKELIHOOD METHOD IN EXTREME RAINFALL FORECASTING USING FOUR PROBABILITY DISTRIBUTIONS - THE CASE OF NORTHERN ALGERIA

EXPLORATION DE LA METHODE DU MAXIMUM DE VRAISEMBLANCE DANS LA PREDETERMINATION DES PLUIES PAR QUATRE LOIS STATISTIQUES - CAS DU NORD DE L'ALGERIE

BENKACI T.[1], MEZENNER N.[2], DECHEMI N.[2]

[1] High National School of Hydraulic, Blida, Algeria
Mobilization and Development of Water Resources Research Laboratory
[2] National Polytechnic School, Algiers, Algeria.

t.benkaci@ensh.dz

Research Article – Available at http://larhyss.net/ojs/index.php/larhyss/index
Received March 5, 2020, Received in revised form July 12, 2020, Accepted July 18, 2020

ABSTRACT

In this research, we have compared four probability distributions: lognormal, Gumbel, gamma and GEV using method of moments (MM) and maximum likelihood (MLE) parameters estimation that we have applied on extreme rainfall in North of Algeria. The main objective of this study is to explore the advantages of MLE method in extreme rainfall frequency analysis. The comparison between the two methods showed that method of moments gives generally better performances than maximum likelihood, especially for GEV distribution comparing to others distributions, this model (GEV) appears least efficient when skewness of data exceeds 1.2. We have concluded that lognormal distribution is the most efficient and stable and gives better simulation of annual maximum daily rainfall using the two methods for the North of Algeria.

Keywords: Rainfall, Flood, Forecasting, Frequency Analysis, Maximum Likelihood, Algeria.

RESUME

Dans cet article, nous avons testé quatre lois de probabilité (log-normale, Gumbel, Gamma et GEV) sur les pluies annuelles extrêmes des régions situées dans le Nord de l'Algérie. L'objectif de la recherche, est l'étude de l'ajustement de ces lois par la méthode du maximum de vraisemblance. Apres analyse approfondie des résultats, nous avons conclu que la loi GEV ne donne pas forcément des résultats meilleurs que les autres lois. De plus, la méthode de vraisemblance pour la loi GEV apparait très instable quand on l'applique aux données hydrologiques présentant une forte variabilité et un coefficient d'asymétrie supérieur à 1,2.

De ce fait, les ajustements effectués par ces quatre lois ont montré que la loi log-normale parait plus robuste et s'adapte correctement aux données de pluies maximales des bassins du Nord Algérien.

Mots clés : Pluie, Crue, prédétermination, Analyse fréquentielle, Maximum de Vraisemblance, Algérie.

INTRODUCTION

The protection of populations and hydraulic structures against floods and their damages is an important step in urban design study and resources management. Flood forecasting is an important research axis in hydrology, since it allow to properly hydraulic works dimensioning (dam and reservoir) and to estimate the amount of runoff that will occur, and predict water levels in the damage-prone areas of drainage basins spatially and temporally.

In hydrology, flood forecasting can be explored by many techniques and depends heavily on the availability of historical data such as extreme rainfall. Accordingly, as noted by some hydrologists (Lang *et al.,* 2014) estimate of hydrological risk at a given site can be highly variable depending on the flood or rainfall forecasting method used.

In order to forecast floods, two techniques are frequently used by the hydrologists:

- Statistical methods based on flood frequency analysis (FFA) by utilizing systematic streamflow/Rainfall observations and are usually employed for estimation of flood quantiles corresponding to different return periods (100, 1000 years) (Laglaine *et al.,* 1994);
- Rainfall-runoff models that generally relate peak discharge to catchment size and other physiographical and climatic catchment characteristics, we mention as an examples the empirical methods (rational method), or including deterministic models like, SCS method or PMP/PMF (Probable Maximum Precipitation/Probable Maximum Flood) widely used in Anglophone countries.

Statistical approach is the most used, and consists to fit a probability distribution to a series of observations for defining the probabilities of future occurrences of some events an estimate of a flood magnitude corresponding to a chosen risk of failure (Meylan *et al.*, 2008). This technique needs selecting a priori an appropriate theoretical probability distribution and fitting it to the observed data.

However, in flood frequency analyses (FFA), the sampling of data is a crucial step, that consists to select the serial of the maximum values.

This sampling can be made by using two main categories:

- Annual maximum method (AM): the sample is composed of maximum peak data (extreme rainfall or discharge) for each year, this sampling method by its simplicity is still the most used in many countries;
- *Peak Over Threshold* method (POT), (Cunnane, 1973, Miquel, 1984, Madsen *et al.*, 1997): in that case, the sample of data is composed of peak values that lie above level or threshold, which is generally set arbitrarily. One of advantages of this method is to take account more information (values) about historical flood for each year (Lang *et al.*, 1999).

However, in semi-arid countries, the applying of this technique (POT) may run up against a threshold or level choice. That last must be lower compared to that humid countries, this was because of drought cycles that act on extreme rainfall.

The validity and the quality of the results of flood frequency analysis depend on the choice of the probability model and more specifically on its type: the objective is to select a statistical distribution function that can calculate the best fit of observed data.

Estimation of parameters of selected probability model is another crucial problem in flood prediction, since its enable to accept or reject the adopted model.

Method of moments (MM) by its simplicity, was the technique the most used for practically all probability distributions, but criticized by some hydrologist for its unreliability and lack of robustness for certain data (Greenwood *et al.*, 1979).

Maximum likelihood estimation (MLE) is another estimation parameter technique, relatively complex, based on an optimisation process allows reducing the difference between observed data and estimated quantiles by the probabilistic model, it seems as theoretically (or mathematically) most robust than method of moments (MM).

Updating flood protection forecasting studies remains a current issue for meteorological and hydrological services to provide information at the different levels and to anticipate a possible defence against flood. In fact, recent extreme weather events and climate change occurred in recent years, have highlighted in large parts of the world vulnerability of the population to extreme natural events such as flooding and storms.

Algeria, which is part of the Mediterranean basin, has been identified as one of the most vulnerable region to the impacts of climate change and to extreme weather events. Indeed, in the past two decades, several cities of the country have endured destructive flood disasters that have resulted in a loss of lives: Algiers (2001, 2007), Ghardaïa (2008), or even Bejaia city (2011, 2015, and 2018).

In this paper, we explore annual maximum daily rainfall forecasting by four probability distributions, based on maximum likelihood estimation method (MLE), trying to evaluate the advantage brought by this method on the results of probability fitting comparing to the method of moments.

MATERIEL AND METHOD

Probability distributions are basic concepts in statistics, and the results of statistical experiments and their probabilities of occurrence are linked by probability distributions.

Let X denote a random variable and x be a particular value of X. The cumulative distribution function (CDF) of X, F(x) is the probability that X is less or equal to x:

$$F(x) = P(X \leq x) = 1 - F(x) \quad (1)$$

In flood frequency analysis, the main objective is to find exceedance probability of an extreme event, and using probabilistic models seeks to calculate the function:

$$F'(x) = P(X > x) = 1 - F'(x) \quad (2)$$

Where $F'(x)$ is the cumulative distribution function (CDF) of x that represents values of studied phenomena (observed extreme rainfall or discharge).

PROBABILITY DISTRIBUTION MODELS USED

For selecting the best-fit probability distribution for a certain basins, the choice of probability distribution models is important, and various probability distributions are currently used to predict expected rainfall in different return periods. However, we mention that there is not a common methodology to select a priori one type of distribution for a specific climate.

In our case, for extreme rainfall prediction, we have selected four probability distributions the most used in hydrology:

PARAMETERS ESTIMATION

Probability distribution models contain (unknown) parameters which must be estimated based on statistical characteristics of observed data (rainfall or discharge). In hydrology, the commonly used to estimate parameters of some distribution are: Method of moments (MM) and Maximum Likelihood estimation (MLE).

Method of moments (MM)

This method is the widely employed, based on the assumption that if the distribution parameters have been correctly estimated, the moments of probability density function (PDF) will equal to the corresponding moments of the sample data, we obtain therefore a system of N equations with N unknowns (Laborde, 2009).

Maximum Likelihood Estimation (MLE)

Let $f_\theta(x_1, x_2, \dots, x_n) = f(x_1, x_2, \dots, x_n|\theta)$ be the PDF of random variables X, Given observed values $X_1 = x_1$, $X_2 = x_2$, . . . , $X_n = x_n$, the likelihood of θ is the function $L(\theta) = f(x_1, x_2, \dots, x_n|\theta)$.

The maximum likelihood estimate (MLE) of θ is that value of θ that maximizes $L(\theta)$: it is the value that makes the observed data the "most probable" (Laborde, 2009):

$$L(\theta) = \Pi_{i=1}^{n} f(x_i|\theta) \tag{3}$$

In some cases, rather than maximizing this product which can be quite tedious, we often use the fact that the logarithm is an increasing function so it will be equivalent to maximize the log likelihood:

$$L(\theta) = \Sigma_{i=1}^{n} \log\left[f(x_i|\theta)\right] \tag{4}$$

Thus the probability distributions and their parameters estimation are described as follow:

Log-normal distribution

Derived from a normal distribution, the log-normal distribution is a distribution of random variables with a normally distributed logarithm. The probability density function (PDF) of the log-normal distribution is calculated as follow:

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-0.5\left[\frac{\ln(x)-m}{\ln(\sigma)}\right]^2} \tag{5}$$

Where the range of variable $x > 0$, the two parameters of log-normal distribution are m mean (parameter of location), and σ standard deviation (scale parameter) of the logarithmic transformation of observed data.

-For MM estimation, the parameters are the first two *r*th moments are respectively m and σ mean and standard deviation of observed data.

-The maximum likelihood (MLE) function for lognormal distribution is given by (Meylan *et al.*, 2008):

$$\ln L(\mu_y \sigma_y^2) = -\frac{n}{2}\ln(2\pi\sigma_y^2) - \sum_{i=1}^{n}\ln(x) - \frac{1}{2\sigma_y^2}\sum_{i=1}^{n}\left[\ln(x_i) - \mu_y\right]^2 \qquad (6)$$

Gumbel distribution

The Gumbel distribution (Gumbel, 1958) also called extreme value *type 1* is often to represent a maximum process, and widely used in hydrology to predict maximum rainfall or discharge. The Gumbel probability density function (PDF) is calculated as follow:

$$F(x) = \frac{1}{\beta}\exp\left[-\frac{(x-\alpha)}{\beta}\exp\left\{-\exp\left(-\frac{x-\alpha}{\beta}\right)\right\}\right] \qquad (7)$$

Where β and α are the parameters of Gumbel distribution. For Moment methods the two parameters β and α are estimated as follow (Chow *et al.*, 1988):

$$\beta = \frac{\sqrt{6}}{\pi} \text{ and } \alpha = m - 0.5772\beta \qquad (8)$$

The maximum likelihood estimation is obtained by the function:

$$L(\alpha,\beta) = f(x_1,...,x_n \mid \alpha,\beta) = \prod_{i=1}^{n} f(x_i \mid \alpha,\beta)$$
$$= \prod_{i=1}^{n}\frac{1}{\beta}\exp\left\{-\frac{(x_i-\alpha)}{\beta}\right\}\exp\left\{-\exp\left\{-\frac{(x_i-\alpha)}{\beta}\right\}\right\} \qquad (9)$$

$$= \left(\frac{1}{\beta}\right)^n \exp\left\{-\sum_{i=1}^{n}\left[\frac{(x_i-\alpha)}{\beta} + \exp\left\{-\frac{(x_i-\alpha)}{\beta}\right\}\right]\right\} \qquad (10)$$

To simplify the calculations, we often use the logarithmic of this function by substituting the product with a sum:

$$l = \ln(L) = -n\ln(\beta) - \sum_{i=1}^{n}\left[\frac{(x_i-\alpha)}{\beta} + \exp\left\{-\frac{(x_i-\alpha)}{\beta}\right\}\right] \qquad (11)$$

$$\begin{bmatrix} \dfrac{\partial l}{\partial u} = 0 \\ \dfrac{\partial l}{\partial \alpha} = 0 \end{bmatrix} \qquad (12)$$

Gamma Distribution

In our case we have used a two-parameter Gamma distribution (α and λ), with probability density function:

$$F(x) = \frac{\alpha^{\lambda}}{\Gamma(\lambda)} \exp(-\alpha x) x^{\lambda-1} \tag{13}$$

Where Γ is the gamma function.

For method of moments the two parameters α and λ are calculated as (Bobee and Des Groseilliers, 1985):

$$\alpha = \left(\frac{2}{C_s}\right)^2 \text{ and } \lambda = \frac{\sigma}{C_s} \tag{14}$$

C_s is the skewness of the data.

The MLE function for Gamma probability fit is detailed as:

$$V = \Pi_{i=1}^{n} f(x_i) = \frac{\alpha^{\lambda n}}{\left[\Gamma(\lambda)\right]^n} \exp(-\alpha \Sigma_{i=1}^{n})(x_1, x_2, \ldots x_n)^{\lambda-1} \tag{15}$$

It is more practical to maximize the logarithm of MLE (Meylan *et al.*, 2008):

$$\ln(V) = n\lambda \ln(\alpha) - n \ln\left[\Gamma(\lambda)\right] - \alpha \Sigma_{i=1}^{n} x_i + (\lambda - 1)\Sigma_{i=1}^{n} \ln(x_i) \tag{16}$$

Generalized Extrême Value (GEV) distribution

The generalized extreme-value (GEV) distribution was introduced by Jenkinson (1955) and widely accepted distribution for describing flood frequency data from the United Kingdom, and its probability density function is described as follow:

$$f(x) = \frac{1}{s}\left[1 - \frac{k(x-x_0)}{s}\right]^{\frac{1}{k}-1} \exp - \left[1 - \frac{k(x-x_0)}{s}\right]^{1/k} \quad k \neq 0 \tag{17}$$

Where $s > 0$ and x_0 are respectively the scale and location parameters, and k is a shape parameter. The range of X depends on the value of k. The shape parameter k determines which extreme value distribution is represented. If k=0 GEV becomes a Gumbel distribution (Extreme value type 1).

In our study we have taken into consideration parameters estimation using Weighted or L-moments (Hosking *et al.*, 1985):

$$k = 7.859c + 2.9554c^2 \quad (18)$$

$$c = \frac{2}{3+\tau_s} - \frac{\ln(2)}{\ln(3)} \quad (19)$$

$$s = \frac{\lambda_2}{(1-2^{-k})\Gamma(1+k)} \quad (20)$$

$$x_0 = \lambda_1 - \frac{\varphi\left[1-\Gamma(1+k)\right]}{k} \quad (21)$$

Where λ_1 is the mean of data, λ_2 is the 2nd *L*-moment, τ_s is the *L* skewness. For Maximum Likelihood GEV estimation, Hosking (1985) recommends to resolve equation system (Martins and Stedinger, 2000):

$$\frac{1}{\alpha}\Sigma_{i=1}^{s}\left[\frac{1-k-(y_i)^{1/k}}{y_i}\right] = 0 \quad (22)$$

$$-\frac{s}{\alpha} + \frac{1}{\alpha}\Sigma_{i=1}^{s}\left[\frac{1-k-(y_i)^{1/k}}{y_i}\left(\frac{x_i-\xi}{\alpha}\right)\right] = 0 \quad (23)$$

$$-\frac{1}{k^2}\Sigma_{i=1}^{s}\left\{\ln(y_i)\left[1-k-(y_i)^{1/k} + \frac{1-k-(y_i)^{1/k}}{y_i}k\frac{(x_i-\xi)}{\alpha}\right]\right\} \quad (24)$$

Quantiles calculation

The main goal of flood frequency analysis is to estimate the quantile *X%* for a chosen return period. Generally quantiles estimation is based on followed equation:

$$X_{\%} = \overline{X} + k\sigma \quad (25)$$

Where $X_{\%}$ is the calculated quantile, $\overline{X}$ and σ are respectively the mean and standard deviation of observed data, *k* is a parameter that depends on the probability (or return period). For Gama distribution, *k* depends on the probability and skewness of the used data.

Basic conditions: independence and homogeneity hypothesis

Frequency analysis and Probability fit distributions need stationarity independence, and homogeneity hypothesis verification (Perreault *et al.*, 1994). In our case, independence and stationarity of data hypothesis was verified using Wald-Wolfowitz test, on other hand we have introduced Wilcoxon test to check for homogeneity of data. Practically, all samples of extreme rainfall data satisfy theses basic conditions.

STUDY AREA

Algeria, situated in the centre of the Maghreb, is the largest country in Africa, covering an area of 2,380,000 km², and has 1200km of coastline from East to west. The North of Algeria enjoys a semi-arid to sub-humid Mediterranean climate, and present high variability of annual rainfall, varying from averages 800 mm in East, and 650 mm in the centre, and decrease to 450 mm for the regions located in the extreme west of the country (ANRH, 1993).

Hydrologically, precipitations regime of Algeria that impacts on the flow is composed of two main different seasons: wet season causing high water level observed from October to march months; more than 70% of rainfall occurs during this period, and dry-season characterised by a lower flow observed during the rest of the year. As mentioned previously, many regions of the country are susceptible to flash flood.

Data used

In order to better broach rainfall forecasting, we have selected a serial of data composed of observed extreme rainfall of ten (10) rainfall stations, uniformly distributed in North of Algeria. For the centre, we have chosen Bejaia station, Bouzareah, Mahelma stations (Algiers), and Baghlia station (Boumerdes city), and stations of Chlef, Mostaganem and Ghazaouet (Tlemcen) cities for the West region. Concerning the East of the country, that is characterized by intense rainfall and flash flood during winter months, we taken into consideration rainfall stations of Ain Assel (El Taref) Pont-Bouchet (Annaba), and Zardezas (Skikda) cities. The series of observed data were provided as the total rainfall (mm) and were collected from Algerian National hydraulic Resources Agency and National Weather Office (ONM) of Algeria. Figure 1 shows the localization of studied rainfall stations, from East to West:

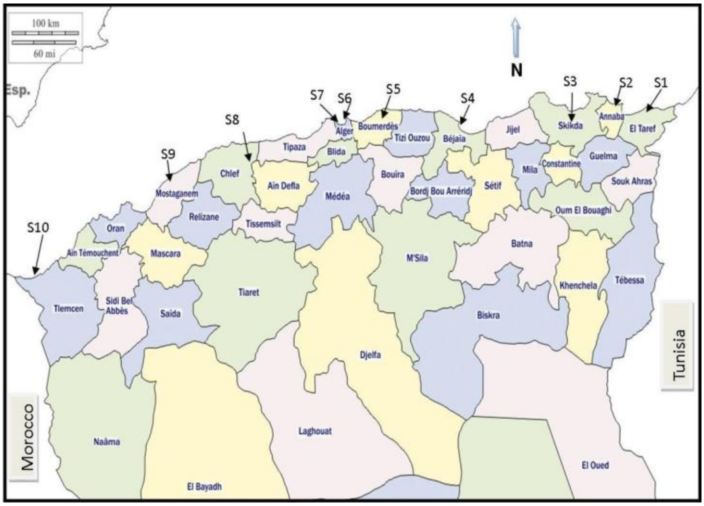


Figure 1: Localization of rainfall stations

Concerning data observation period, the database is composed of a continuous samples or extremes rainfall occurred during the period 1980-2016, therefore the observation period of series of rainfall exceeds 32 years for the major of rainfall stations, except two stations: Bouzareah (28 years) and Baraki (27 years) which have less than 30 years of observations.

The choice of this observation period (1980-2016) is justified by the fact that the major of intense rainfall events occurred during this period, and the serial data have no lack, which enables probability distribution fitting without any constraints.

For the sampling, the database provides the highest total daily rainfall (mm) for each month; the maximum rainfall observed during one year is selected as a data of the sample of annual extreme rainfall.

The basic statistical parameters of rainfall data are summarized in Table 1.

Table 1: Statistical characteristics of rainfall stations

R. Station	Code	Coordinates(*)			Mean	Max	Min	Std	Skew
		X(km)	**Y(km)**	**Z(m)**					
S1 (El Tarf)	031602	1005.4	399.8	35.0	62.8	142.7	28.0	26.2	1.36
S2 (Annaba)	140631	949.97	402.82	3.0	54.6	114.6	23.6	25.2	1.13
S3 (Skikda)	030903	875.3	374.6	195	57.1	142.0	21.0	26.6	1.38
S4 (Bejaia)	151009	712.7	386.5	6.0	47.7	121.0	18.6	19.7	1.78
S5 (Baghlia)	022002	603.6	390.0	30.0	61.8	97.4	32.7	19.3	0.17
S6(Bouzareah)	020506	529.7	388.3	354	86.1	260.0	28.4	57.7	1.60
S7 (Mahelma)	020511	517.0	376.8	150	56.2	131.0	25.5	23.9	1.42
S8 (Chlef)	011715	443.9	318.0	320	32.4	60.0	15.9	11.1	0.71
S9(Mostaganem	040612	266.4	293.4	151	50.7	173.7	15.4	27.8	2.68
S10 (Tlemcen)	040101	81.13	201.3	4.0	59.0	115.3	20.4	25.2	0.84

(*): Lambert coordinate system

RESULTS AND DISCUSSION

Generally frequency analysis of hydrological data is showed graphically using correlation coefficient. In our case, to better analyse the results and to compare between the different probabilistic models, we took into consideration an assessment criterion that consists to estimate difference between the observed data and the calculated quantiles by the four probabilistic models: log-normal distribution, Gumbel distribution, gamma distribution and GEV distribution. Various numerical criteria have been used by the hydrologists. In this study, we applied root mean squared error (RMSE), that is the standard deviation of the residuals and measures how much error there is between two data sets: observed and predicted. RMSE criterion is expressed as:

$$RMSE = \sqrt{\frac{\Sigma_{i=1}^{n}(P_i - x_i)^2}{n}} \tag{26}$$

Where P_i and is the observed rainfall and X_i is the calculated quantile. Probability fit distribution is correct when RMSE value is near to zero.

Since the study deal to comparative analysis between method of moments (MM, (L-moment for GEV distribution) and Maximum likelihood estimation (MLE), we present in table 2, probability fit results obtained by the four distributions for each rainfall station, this for the two estimation methods.

Table 2: Results of probability Distributions (RMSE values)

Stations	M. Moments (MM)				Maximum Likelihood (MLE)			
	Lognorm	Gumb	Gamma	GEV	Lognorm	Gumb	Gamma	GEV
S1 (El Tarf)	4.29	4.47	5.38	**3.60**	**4.52**	4.98	5.85	7.36
S2 (Annaba)	**4.79**	4.95	5.30	5.01	**4.90**	5.65	5.44	12.11
S3 (Skikda)	3.89	4.06	4.61	**3.0**	**3.73**	4.94	5.27	9.42
S4 (Bejaia)	4.70	4.75	5.41	**4.41**	**4.80**	4.95	5.65	5.84
S5 (Baghlia)	5.17	5.18	3.95	**3.58**	5.02	5.58	**3.96**	22.10
S6(Bouzareah)	10.61	13.85	10.97	**10.28**	**11.60**	18.50	14.67	23.41
S7 (Mahelma)	3.66	3.66	4.54	**2.30**	**3.93**	4.76	5.22	10.05
S8 (Chlef)	1.86	1.94	**1.84**	2.13	**1.87**	1.98	1.90	2.82
S9(Mostaganem)	9.63	10.52	10.82	**8.73**	**9.36**	10.57	10.90	12.80
S10 (Tlemcen)	5.22	4.83	**4.61**	5.30	5.10	4.87	**4.69**	5.06
Mean	5.38	5.82	5.74	**4.83**	**5.48**	6.68	6.36	11.10

The results obtained from extreme rainfall frequency analysis, presented in Table 2 show that the four probabilistic models give different results; by comparing the two methods parameters estimation, we can distinguish that these models give best approximated quantiles using moments method (MM) especially for GEV model, since the RMSE values are lower than those obtained using MLE method.

As for example, method of methods give better quantiles for the rainfall stations of East and the centre of country (Annaba, Skikda, Bejaia, Bouzareah, and Mahelma stations); the values of RMSE criterion is therefore lower comparing to those obtained using maximum likelihood method especially for lognormal distribution, where the mean of RMSE is around 11.4 for the MM method, and pass to 17.5 for MLE method for Bouzareah station. The same is for Gumbel distribution where the RMSE value of MM method (13,85) vary considerably from that of MLE method (18,5), and that for others probability fit Gamma and GEV distributions.

As an example, figure 2 shows the QQ-plot of Skikda station created for the four probability distributions using MLE method, lognormal model matches correctly the observed extreme rainfall, contrarily to GEV model that appears graphically less efficient.

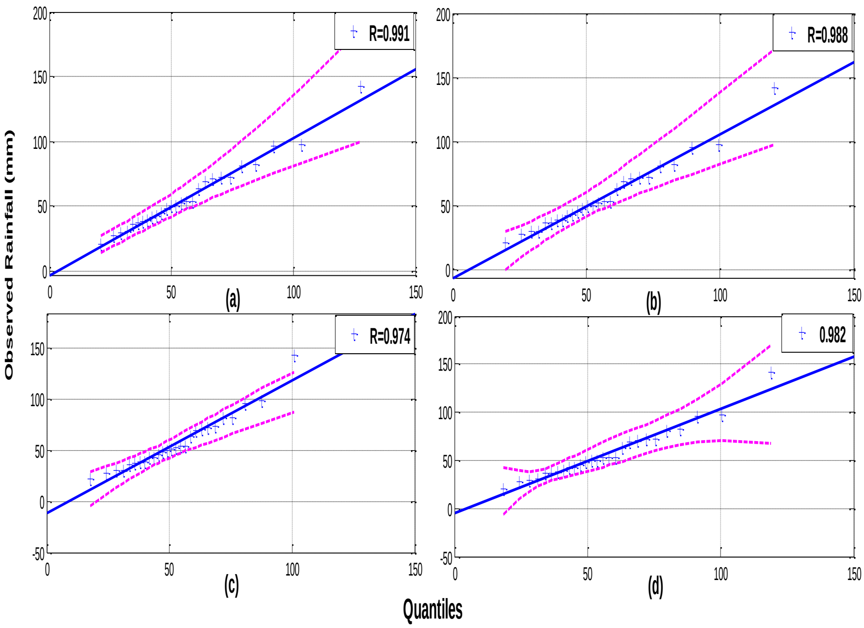


Figure 2: QQ-plot for Skikda rainfall station using MLE method: (a) Lognormal, (b) Gumbel, (c) GEV, (d) Gamma distributions

For the rainfall stations of the West of the country (Chlef, Mostaganem stations), we note that the results are generally too close, with a relative decline for GEV distributions using MLE. Same observations are obtained for Bejaia rainfall station, where the results do not differ significantly between the two methods.

If we compare between the four distributions, particularly in terms of mean, we remark easily that GEV distribution using method of moments gives generally lower RMSE values in the majority of cases, therefore best results, followed by lognormal distributions.

Method of moments or maximum likelihood method?

Maximum likelihood estimation is a method based on an optimisation process, relatively complex; consequently for a certain distributions (GEV case) algorithm must iterate many times to obtain an optimisation zone and to find the best parameters. GEV is a three-parameter distribution, thus the use of MLE method would face the problem of equi-finality of parameters, especially the third parameter "parameter of scale" which take a values close to zero (a positive or negative values), this makes difficult to obtain a stable results, and this is particularly important when skewness of data is greater than 1,2: El Tarf, Annaba, Skikda, Mahelma (figure 3) and Bouzareah stations, that last present a worst probabilistic fitting with MLE. However, GEV distribution using method of moments (Hosking *et al.*, 1985) allows to better calculate the quantiles and became more efficient.

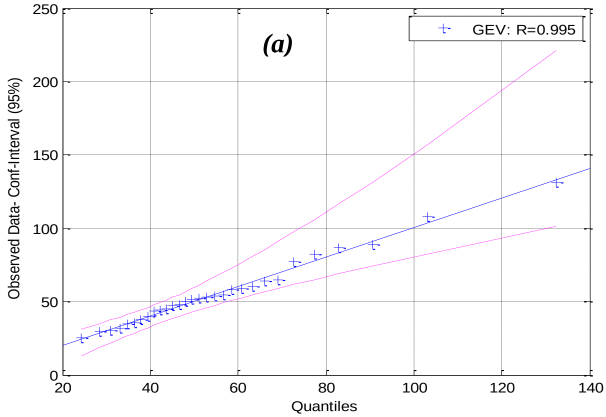
The best fit probabilistic model: Lognormal distribution

The selected data in this research showed the rainfall range varied widely in all the stations for the North of Algeria.

To choice of distribution model is very hard, since several factors affect the results fitting, and Rainfall patterns vary from country to country as well as from weather station to station (Ashraful *et al.,* 2018).

In our case, according the results of the four probability distributions studied for the maximum daily of stations, and taking into account the two methods: MM and MLE, two-parameter lognormal model appears the most efficient and most stable, and this whatever the characteristics of data: standard deviation, skewness, etc.

After many tests of lognormal distribution fitting, the RMSE criterion values obtained are generally lower and acceptable for the major of rainfall stations.



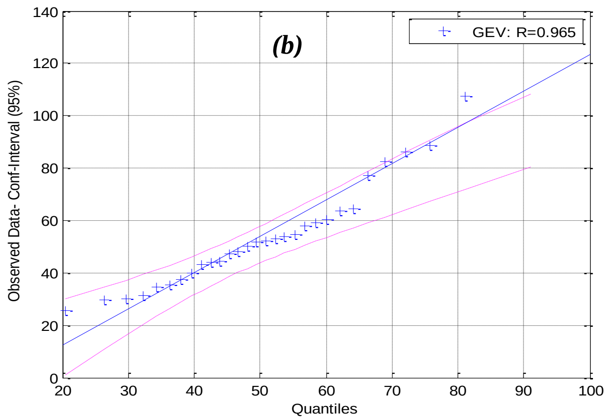


Figure 3: Simulation of quantiles by GEV distribution (Mahelma station): (a) Method of moments, (b) Maximum likelihood estimation

CONCLUSIONS

Floods attract attention for authorities because they have large social consequences for the populations, include loss of human life, damage to property, and deterioration of health conditions. Flood forecasting research is of great importance in hydrology.

In this paper, we have tested four probability distributions lognormal, Gumbel, Gamma and GEV, applied on extreme rainfall of regions situated in North of Algeria, by using two parameters estimation methods: method of moments (MM) and Maximum likelihood (MLE).

The results of different probability fits show that MLE method does not give better results than method of moments, especially for GEV distribution, where the estimation of parameters (with MLE) remains very tedious and relatively uncertain, and when data used (extreme rainfall) presents high variability with a skewness coefficient greater than 1.2, GEV distribution using MLE appears in that case very unstable.

On the contrary, when we used method of moments, GEV probability distribution for extreme rainfall become most robust and gives best estimated quantiles.

After comparison between the four probabilistic models, and given the results obtained by the two methods MM and MLE for all rainfall stations studied, two-parameter lognormal (2P) fit distribution appears most robust, most stable and fit correctly daily extreme rainfall of North of Algeria.

REFERENCES

ASHRAFUL A.M, EMURA K., FARNHAM C., YUAN J. (2018). Best-Fit Probability Distributions and Return Periods for Maximum Monthly Rainfall in Bangladesh, Climate Journal, Vol.6, Issue 9, pp. 1-16.

ANRH : Agence Nationale des Ressources Hydrauliques (1993). Carte pluviométrique de l'Algérie du Nord. Une carte au 1/500000^{e} et sa notice explicative. Ministère de l'Équipement, Alger, Algérie

BOBEE B., DES GROSEILLIERS L. (1985). Ajustement des distributions Pearson type 3, gamma, gamma généralisée, log-Pearson type 3 et log-gamma, Rapport scientifique no 105, INRS-Eau, Université du Québec, 154 p.

CHOW V.T., D. MAIDMENT, MAYS L.W (1988). Applied hydrology, McGraw-Hill Book Company, 572 p.

CUNNANE C. (1973). A particular comparison of annual maxima and partial duration series methods of flood frequency prediction, Journal of Hydrology, Vol. 18, pp. 257–271.

GUMBEL J. (1958). Statistics of Extremes, Columbia University Press, New York, 375p.

GREENWOOD A., LANDWEHR J., MATALAS N., WALLIS J. (1979). Probability weighted moments: Definition and relation to parameters of several distributions, Water Resources Research, Vol.15, pp. 1049–1054.

HOSKING J. (1985). Algorithm AS 215: Maximum-likelihood estimation of the parameter of the generalized extreme-value distribution, Applied Statistics, Vol.34, pp. 301–310.

HOSKING J., WALLIS J., WOOD E. (1985). Estimation of the generalized extreme-value distribution by the method of probability weighted moments, Technometrics, Vol. 27, Issue 3, pp. 251–261.

JENKINSON A.F. (1955). The frequency distribution of the annual maximum (or minimum) values of meteorological elements, Quarterly Journal of the Royal Meteorological Society, Vol. 81, No. 348, pp. 158-171.

LABORDE J. (2009). Éléments d'hydrologie de surface. École polytechnique de l'université de Nice- Sophia Antipolis, France, 199 p.

LAGLAINE V., BEROD D., DEVRED D., MUSY A. (1994). Modélisation déterministe de la crue extrême d'un bassin versant de montagne avec application de la description géomorphologique du réseau hydrographique, Revue Sciences de l'eau, Vol. 7, pp. 285-308.

LANG M., OUARDA T., BOBÉE B. (1999). Towards operational guidelines for over-threshold modeling, Journal of hydrology, Vol. 3, pp. 103-196.

LANG M., ARNAUD P., CARREAU J. (2014). Résultats du projet ExtraFlo (ANR 2009-2013) sur l'estimation des pluies et crues extrêmes, La Houille Blanche, Vol. 2, pp. 5-13.

MADSEN H., RASMUSSEN P., ROSBJERG D. (1997). Comparison of annual maximum series and partial duration series methods for modeling extreme hydrologic events, Water Resources Research, Vol.33, Issue 4, pp. 747-757.

MARTINS E., STEDINGER J. (2000). Generalized maximum-likelihood generalized extreme-value quantile estimators for hydrologic data. Water Resources Research, Vol.36, No 3, pp. 737-744.

MEYLAN P., FAVRE A., MUSY A. (2008). Hydrologie fréquentielle: une science prédictive, Presses polytechniques et universitaire romandes, EPFL, Lausanne, suisse, 173 p.

MIQUEL J. (1984). Guide pratique d'estimation des probabilités de crues, Eyrolles, France, 160 p.

PERREAULT L., BOBEE B., LEGENDRE P. (1994). Rapport général du logiciel ajuste-II: théorie et application, Institut National de la Recherche Scientifique, INRS-Eau, Québec, 102 p.