JOURNAL OF FLOOD ENGINEERING
Vol. 9 No. 1 (January-June, 2018)

DECISION SUPPORT SYSTEM FOR FLOOD RISK ASSESSMENT

Salaheddine El Adlouni[1] and Bernard Bobée[2]
[1]*Professor, Institut National de Statistique et d'Économie Appliquée, INSEA, Rabat-Morocco, Department of Mathematics and Statistics, University of Moncton.*
E-mail: salaheddine.el.adlouni@umoncton.ca
[2]*Emeritus Professor, Institut National de la Recherche Scientifique, centre Eau, Terre et Environnement, Québec-Canada*

ABSTRACT: The estimation of flood flows is of major importance for the design and management of hydraulic structures. Two main classes of distributions are used in hydrology frequency analysis: the class *D* of sub-exponential and the class *C* of regularly varying distributions with a heavier tail. No criteria were available for the choice between these two classes of fit the most appropriate especially at extreme values. A Decision Support System (DSS) based on the characterization of the right tail, corresponding to large return period *T*, of probability distributions used in frequency analysis, has been developed. The DSS allows discriminating between the class *C* and *D*. It is worth to emphasize that the class selection has a great importance for the extrapolation to estimate events with large return periods. An illustration of the DSS approach is presented using the maxima peak flow at the Potomac River, USA (n=80).

Keyword: Extreme values, Return period, Decision Support System, Asymptotic behavior.

1. INTRODUCTION

Flood Frequency Analysis (FFA) is of particular interest for the design and management of hydraulic structures. The principal objective of FFA is to obtain robust estimates of the occurrence of extreme events from a dataset of observations that are independent and identically distributed (IID). These hypothesis indicate that the observations are independent and they are generated by the same phenomenon which is assumed to be the same in the future. To check whether the observations are independent and the data series are stationary and homogeneous tests such as Wald-Wolfowitz, Kendall and Wilcoxon can be used (Bobée and Ashkar, 1991).

The FFA procedure is related to extreme value theory (EVT), which is often derived from asymptotic properties according to the Fisher-Tippet theorem (Fisher and Tippet, 1928). Conventional estimates of flood exceedance quantiles are highly dependent on the underlying flood frequency distribution. The form the right tail, difficult to estimate from observed data, is particularly important concerning event with large return periods. Several standard frequency distributions have been extensively studied in the statistical analysis of hydrologic data. Physical processes which generate extreme events are rarely considered for the choice of the model (Kidson et al., 2005; Singh and Strupczewski, 2002) i.e. the theoretical model of observed data can not be derived from physical consideration. The statistical selection of the most appropriate distribution of annual maximum or peaks over threshold series has received widespread attention. There are no rigid rules governing which type of distribution is most appropriate for a particular case, and a variety of probability distributions are commonly used as frequency-magnitude distributions in hydrology.

2. PROBABILITY DISTRIBUTIONS USED IN FFA

Brooks and Carruthers (1953) indicated that the Gumbel distribution (Gumbel, 1942), which is commonly used in flood frequency prediction, tends to underestimate the magnitude of the rarest rainfall events. Bernier (1959) suggested the Log-Gumbel distribution for hydrological series. The Log-Gumbel (i.e. the log of the peak flow follows the Gumbel or EV1 distribution) distribution is also called Fréchet distribution (Fréchet, 1927) and is known as EV2; a special case of the Generalized Extreme Value (GEV) distribution. In order to select the adequate distribution, empirical comparisons are commonly used for a given region. A comparison of the Lognormal (LN), Gamma (G), Gumbel (EV1), Fréchet (EV2), and Log-Pearson type 3 (LP3) fits for ten USA stream flow stations was presented by Benson (1968). Based on this study, the USA adopted a uniform approach to flood frequency estimation which consists of fitting LP3 distribution to the annual peak discharges (US Water Resources Council, 1967). Although some countries, such as Australia, have adopted the LP3 distribution (Eslamian, 2010), others have selected different distributions, such as the Generalized Extreme Value (GEV) distribution in the United Kingdom and the Lognormal (LN) in China, (Bobée 1999; Robson and Reed 1999). Discussions and reviews of the application of these and other statistical distributions to FFA are given in Stedinger et al. (1993), Bobée and Rasmussen (1995) and Rao and Hamed (2000). A comparison study for some distributions, given by Koutsoyiannis (2004), shows that the EV2 distribution is more adequate to represent extreme rainfall series. Halphen distributions (Halphen type *A*, type *B* and Inverse type *B*, HIB) have been introduced to fit a large variety of data sets (Halphen 1941; Morlat 1951, 1956). They constitute with their limiting forms, the Gamma and Inverse Gamma distributions, a complete system to model hydrological variables. Indeed, the (δ_1, δ_2) diagram (Morlat 1956, El Adlouni and Bobée, 2007) makes it possible to represent this family of distributions and their limiting cases in a plane the same way as the well-known Pearson system with the (β_1, β_2) diagram corresponding to the skewness and kurtosis coefficients. The (δ_1, δ_2) diagram is defined by the moment ratios $\delta_1 = \ln(A/G)$ and $\delta_2 = \ln(G/H)$, where *A*, *H*, and *G* are the arithmetic, harmonic and geometric means, respectively (Bobée et al., 1993). For any sample, the corresponding (δ_1, δ_2) point is associated to one and only one member of the Halphen family or their limiting cases.

Werner and Upper (2002) presented a classification of a distributions with respect to their tail behaviour and more details are given by (El Adlouni et al., 2008). These classes of distributions are nested $(A \subset B \subset C \subset D \subset E)$:

E : distributions with non-existence of exponential moments

D : sub-exponential distributions

C : regularly varying distributions

B : Pareto type tail distributions

A : α-stable (non-normal) distributions

In FFA, the annual maximum flow datasets which are generally independent and identically distributed are concerned. Indeed, let $Y_1;Y_2;\ldots;Y_p$ be a sequence of daily flow, then the sample maximum *X* is defined as $X = \max\left(Y_1;Y_2;\ldots;Y_p\right)$ and $p = 365$. Distributions that are usually used in FFA belong to the classes *C* and *D*. However, the class *E* will be also considered in the DSS for more generality. Indeed, even if the proposed approach is illustrated in the case of extremes in hydrology, its use is valid for any series of observations that are IID. Distributions that are widely used in hydrology to represent maximum annual flow series are:

- Class *C* (regularly varying distributions): Fréchet (EV2), Halphen Inverse type *B* (HIB), Log-Pearson type 3 (LP3), Inverse Gamma (IG).
- Class *D* (sub-exponential distributions): Halphen type *A* (HA), Halphen type *B* (B), Gumbel (EV1), Pearson type 3 (P3), Gamma (G).

Figure 1 presents exponential (*E*), sub-exponential (*D*) and regularly varying (*C*) distributions. Distributions are ordered from light tailed (from the left) to heavy tailed (to the right). The limiting cases (bottom squares) represented by distributions in the limits of classes. The tail of the class *C* distributions is heavier than that of the class *D* distributions, which is heavier than that of the class *E*. Thus, estimated quantiles can be ordered equivalently. Indeed, for a given sample, the *T*-event corresponds to the quantile of the probability of non-exceedance $p = 1 - 1/T$ estimated by distributions of the classes *C*, *D* and *E*, are *QT* (*C*), *QT* (*D*) and *QT* (*E*) respectively, which verify the following relation: $QT(E) < QT(D) < QT(C)$.

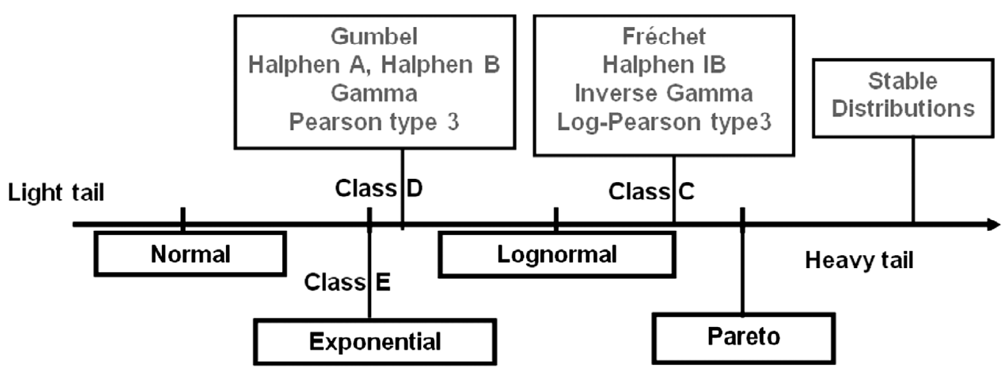

Figure 1: Distributions Ordered with Respect to Their Right Tails (El Adlouni et al., 2008)

The importance of the tail behaviour was illustrated by Hubert and Bendjoudi (1996) who showed that the event of return period $T = 1000$ estimated by the normal distribution corresponds to *T* = 100 years when the power type tail distribution is the most adequate to fit precipitation data series. Figure 2 shows that quantile estimated by the HIB distribution (class *C*) is larger than that of the Lognormal and the distribution P3 (class *D*) even if they have the same statistical characteristics (mean and variance). The difference is more important for a large non-exceedance probability (high return period). This comparison illustrates the importance of tail behaviour determination to estimate large return period quantiles. Indeed, the 1000-return period estimated from the P3 distribution correspond to the flow of return period 200 years when the HIB distribution is fitted.

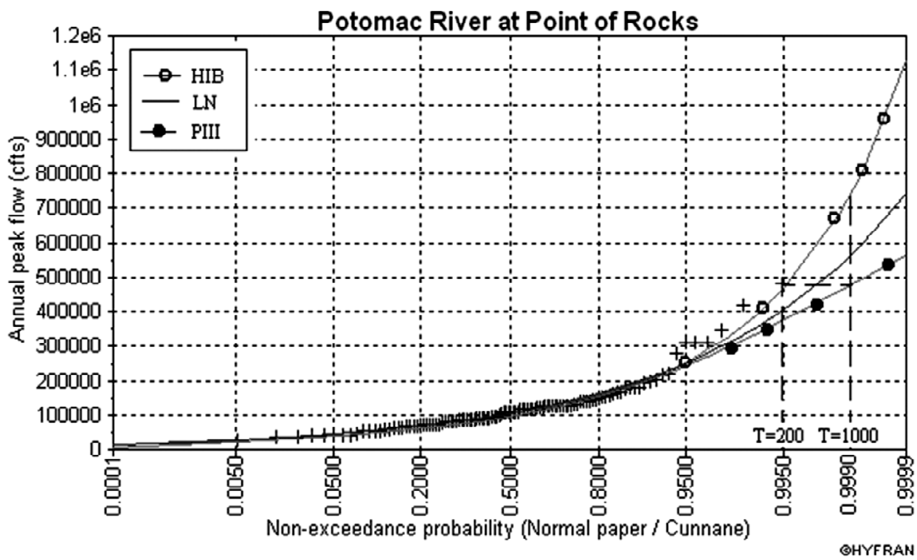

Figure 2: Comparison of the HIB, LN and PIII Fits for the Annual Peak Flood at Potomac River

Remark

The Lognormal distribution (*LN*) does not belong to any of these two classes *C* or *D*. It has an asymptotic behaviour which is in the frontier of the classes *C* and *D* (Figure 1). Indeed, the LN tail is lighter (respectively, heavier) than that of a distribution of the class *C* (respectively, class *D*). Thus, the quantiles (*QT*) estimated by a distribution belonging to the classes *C*, *D* and the *LN*, verify the following relation:

$QT(D) < QT(LN) < QT(C)$. Consequently:

- If the parent distribution is regularly varying (class *C*), and the *LN* distribution is considered for the fit, thus the estimated quantile, for a fixed return period, will be lower than the real value and there is a risk to underestimate this quantile;
- If the true distribution is sub-exponential (class *D*), and the *LN* distribution is considered for the fit, thus the estimated quantile, for a fixed return period, will be higher than the real value and there is a risk to overestimate this quantile.

In the DSS, and to have a safer choice, *LN* is considered by default as a distribution of the class *D*. However, the user could make a different decision and associate it to the class *C*. Research work has been done to determine powerful tools to test log-normality (Martel et al., 2011).

The majority of distributions used in the FFA are available in HYFRAN (CHS, 2002 and Hubert, 2005) software to fit data sets that are independent, homogenous and stationary (IID hypotheses). To select the most adequate fit, the Akaike (AIC, Akaike, 1974) and Bayesian (BIC, Fortin et al., 1998) Information Criteria are available in HYFRAN. These criteria and classical tests, give more weight to the central part of the sample. However, discrepancies between different models are significant for rare events (i.e. corresponding to large return periods). To solve this problem, a Decision Support System (DSS) is developed to help to the selection of the most appropriate class of distributions, with respect to extreme values.

The methods developed in the DSS allow the identification of the most adequate class of distribution (Figure 1) to fit a given sample, especially for extremes. These methods are (Figure 3):

- The Log-Log plot: used to discriminate between on the one hand the class *C* and on the other hand the classes *E* and *D*;
- The mean excess function (MEF) to discriminate between the classes *D* and *E*; and
- Two statistics: Hill's ratio and modified Jackson statistic, for confirmatory analysis of the conclusions suggested by the previous two methods.

Note that, the class *E* is also considered in the DSS for more generality and the approach can be used for any series of observations that are IID.

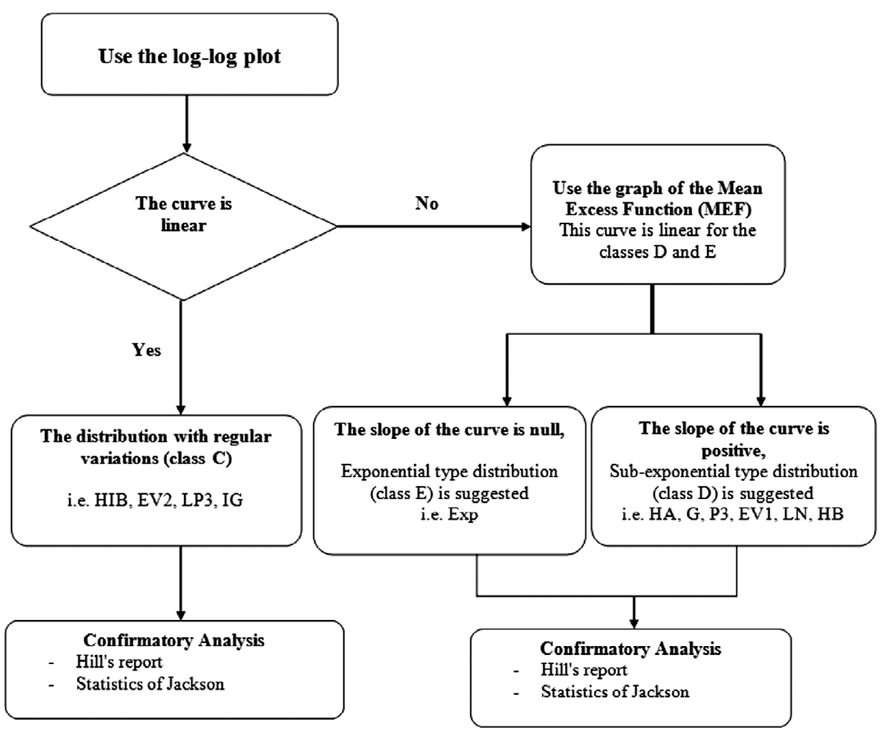

Figure 3: Diagram for Class Discrimination used in the DSS

More theoretical details of this classification and the criteria are available in El Adlouni et al. (2008).

In the four following sections, the details of the integration of these criteria to the DSS of HYFRAN-PLUS software are presented.

3. LOG-LOG PLOT

The log-log plot is based on the fact that the survival function $\bar{F}(u) = P(X > u)$, is given by $\bar{F}(u) = P(X > u) = \exp\{-u/\theta\}$ for exponential tail with mean θ; for regularly varying distribution with tail index α, $\bar{F}$ is equivalent to (for large quantile):

$$\bar{F}(u) = P(X > u) \approx C\int_u^\infty \frac{1}{x^\alpha}dx = C\left[\frac{x^{-\alpha+1}}{1-\alpha}\right]_u^\infty = C_1 u^{-\alpha+1}$$

(with $\alpha > 1$, which is equivalent to finite mean).

Therefore, taking the logarithm we have for regularly varying distributions $\log\left[P(X > u)\right] \approx \log C_1 - (\alpha - 1)\log(u)$. This suggests that, for the log-log plot, the tail probability is represented by a straight line for power-law (or regularly varying distributions, class *C*) but not for the other sub-exponential or exponential distributions (class *D* or *E*).

As illustrated in Figure 4, the curve represented in the Log-Log plot corresponds to a straight line for the distributions of the class *C* i.e. Fréchet (EV2), Halphen type IB (HIB), Log-Pearson type 3 (LP3) and Inverse Gamma (IG), but not for sub-exponential or exponential type tails (class *D* or *E*). When the diagram is not linear we suggest the use of the Mean Excess Function (MEF) to discriminate between the classes *D* and *E*.

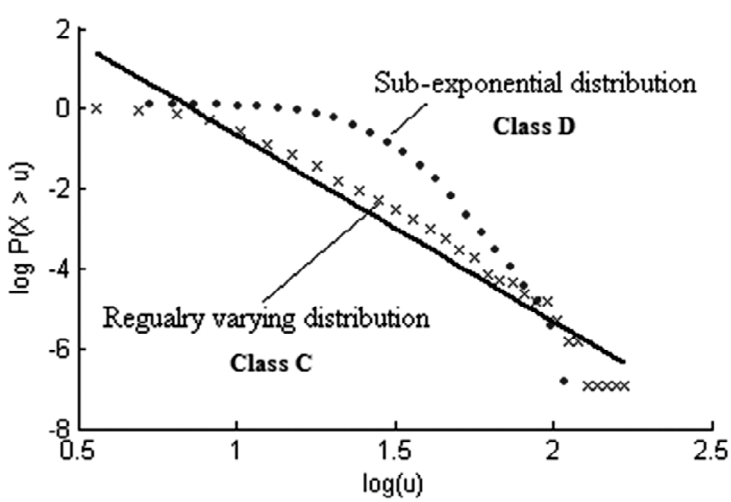

Figure 4: Illustration of the Log-Log Plot to Characterize the Regularly Varying Distributions

To check the linearity of the curve in the log-log diagram, a test on the associated coefficient of correlation is considered. Simulation studies allow the determination of critical values (*rc*) corresponding to significance levels of 5% and 1%, to test the HYPOTHESIS *H*0: THE DATA FOLLOW A DISTRIBUTION OF THE CLASS *C* (i.e. THE CURVE IS LINEAR). The hypothesis *H*0 is equivalent to H0′: THEORETICAL VALUE OF THE COEFFICIENT OF CORRELATION $\rho = 1$. These critical values are calculated according to the size *N* of the sample ($30 \leq N \leq 200$). Note that the decisions given by the DSS are based, by default, on the significance level 5%.

If the hypothesis *H*0 is rejected, at the significance level 5%, the use of the mean excess function plot (MEF) are suggested. However the critical value at the significance level 1% is also given for more flexibility and to allow the user to make another decision than that based, by default, on the significance level 5%.

Indeed, if the observed correlation coefficient (*ro*) is greater than critical value (*rc*) at the significance level 5%, then it is concluded that it is not significantly different from 1 at the significance level 5 % and the hypothesis *H*0 of linearity is accepted at this level. In this case, the most adequate choice corresponds to the class *C* of regularly varying distributions (power-law type): HIB, EV2, LP3, IG.

4. THE MEAN EXCESS FUNCTION DIAGRAM (MEF)

The mean excess function method is based on the function $e(u) = E[X - u \mid X > u]$. This function is constant for exponential tail distributions $(e(u) = \theta)$. However, in the case of regularly varying distribution with tail index $\alpha(\alpha > 2)$: $e(u) = \frac{u}{(\alpha - 2)}$. The Mean Excess Function (MEF) allows discriminating between the class *D* (sub-exponential distributions) and the class *E* (Exponential distribution). Indeed, the curve presented in the MEF diagram is linear for high observed values for distributions of both classes *D* and *E*. If in addition, the slope of this curve is (Figure 5):

- Equal to zero, the most adequate distribution belongs to the class *E* (Exponential law);
- Strictly positive, the most adequate distribution belongs to the class *D* of sub-exponential distributions: HA, EV1, HB, P3 and G.

Note that, in the DSS, this method should be used after the log-log plot method. Indeed, if the assumption *H*0 of the log-log plot method is rejected (distribution does not belongs to the class *C* of regularly varying distributions and then it belongs to class *D*) FME method allows testing whether the distribution is exponential or not.

The use of this diagram in the DSS is based on the slope of the MEF curve for the observations that exceed the median (50% of the highest observed value of the sample because it is an asymptotic test).

Simulation studies allow the determination of critical values of the slope corresponding to significance levels of 5 and 1 %, to test the HYPOTHESIS *H*0: THE DATA FOLLOW A

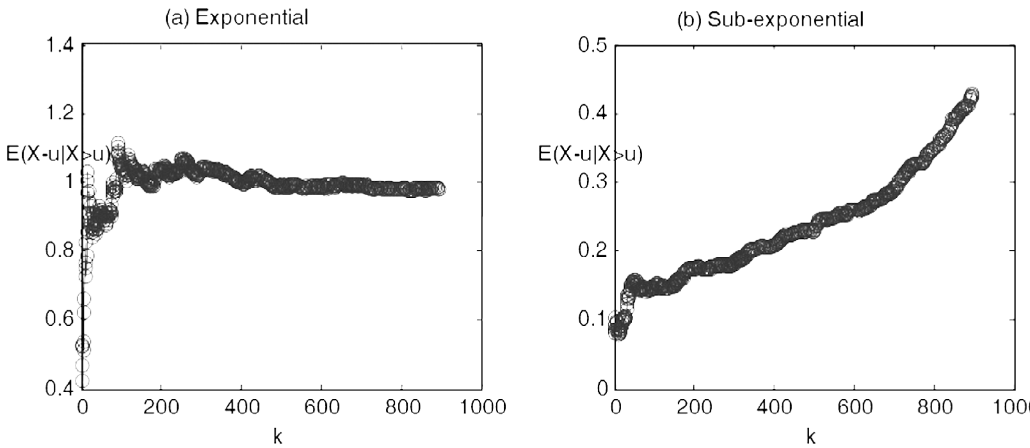

Figure 5: Mean Excess Function for (a) Exponential and (b) Sub-exponential Distributions

DISTRIBUTION OF THE CLASS *E* (i.e. THE SLOPE OF THE MEF IS EQUAL TO ZERO). These critical values are calculated according to the size *N* of the sample ($30 \leq N \leq 200$). Note that the decisions given by the DSS are based, by default, on the significance level 5 %.

When the hypothesis *H*0 is accepted, it is suggested the use of the Exponential distribution (class *E*). However, when it is rejected at the significance level 5 %, the use of a distribution of the class *D* (HA, EV1, HB, P3, G) is suggested.. Note that the critical values at the significance level 1 % are given for more flexibility and to allow the user to make possibly another decision than that recommended, by default, for the significance level of 5%.

5. HILL'S RATIO PLOT

The Hill ratio is defined by

$$a_n(x_n) = \frac{\sum_{i=1}^{n} \mathrm{I}(X_i > x_n)}{\sum_{i=1}^{n} \log(X_i / x_n)\mathrm{I}(X_i > x_n)}$$

where $\mathrm{I}(X_i > x_n) = \begin{cases} 1 & \text{if} \quad X_i > x_n \\ 0 & \text{if} \quad X_i < x_n \end{cases}$.

This method is based on the fact that a_n is a consistent estimator of α if the tail is regularly varying (Class *C*) with tail index α (Hill, 1975). In the expression of the Hill ratio, x_n is chosen to be large such that $P(X > x_n) \to 0$ and $nP(X > x_n) \to \infty$, and I is the indicator function. The standard Hill estimator, of the tail index, corresponds to the particular case where the observations are ordered $X_{(1)} \leq \ldots \leq X_{(n)}$ and $x_n = X_{(k_n+1)}$, where k_n is an integer which tends to infinity as *n* tends to infinity.

In practice, the values of the function $a_n(x_n)$ are plotted as function of x_n and the user is looking for some stable regions from which $a_n(x_n)$ can be considered as an estimator of α. Figure 6, presents the Hill ratio plot for a sample varying (6-a) and sub-exponential (6-b) distributions.

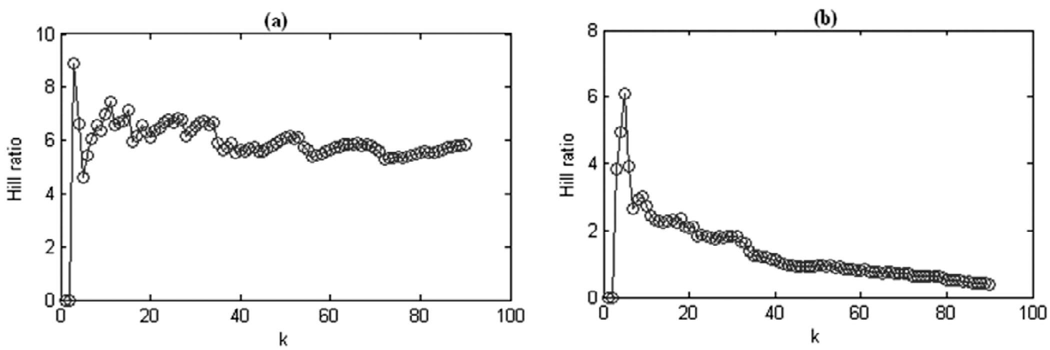

Figure 6: Generalized Hill Ratio Plot for (a) Regularly-varying and (b) Sub-exponential Distributions

This statistics is used in the DSS to confirm the suggested choice given by the first two diagrams (the distribution belongs to the class *C*, *D* or *E*).

- If the curve converges to a non-null constant value, the most adequate distribution belongs to the class *C* (regularly varying distribution). The recommendation is then to use of a distribution of the class *C*: Fréchet (EV2), Halphen Inverse type *B* (HIB), Log-Pearson type 3 (LP3), Inverse Gamma (IG).
- If the curve **decreases to zero**, the distribution belong to the Sub-exponential class (class *D*: Halphen type *A*, Gamma, Pearson type III, Halphen type *B*, Gumbel); and the Exponential class (class *E*: Exponential distribution).

6. JACKSON STATISTIC

This method is presented by Beirlant et al. (2006) and is based on the Jackson statistic (Jackson, 1967). It allows to test whether the sample is consistent with Pareto type distributions. Note that the distributions of the class *C* (regularly varying distribution) have asymptotically the same behaviour as that of the Generalized Pareto distribution. Originally the Jackson statistic was proposed as a goodness-of-fit statistic for testing exponential behaviour, and given the link between the Exponential and the Pareto distribution (if *X* has a Pareto distribution, the logarithmic transformation $Y = \log(X)$ is exponentially distributed) this statistic is used to assess Pareto-type behavior. The Jackson statistic is further modified by taking into account the second-order tail behavior of a Pareto-type model. Beirlant et al. (2006) give the limiting distribution of this statistic with corrected bias version for finite size samples. This adapted version of the Jackson statistic converges to 2 for power tail type distribution and has an

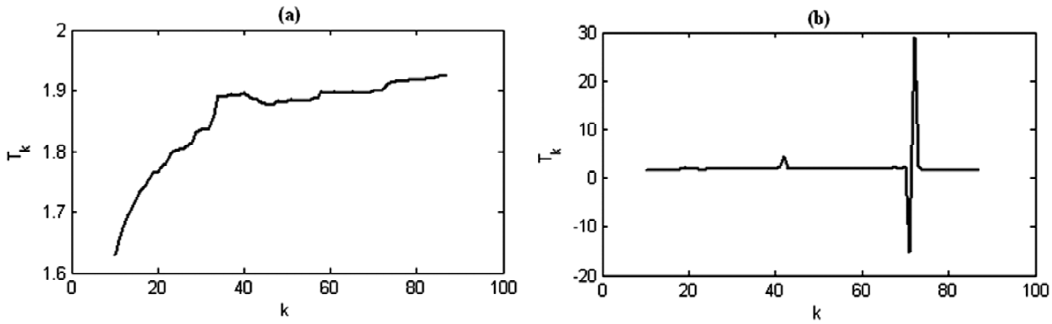

Figure 7: Adapted Jackson Statistic for (a) Regularly Varying and (b) Sub-exponential Distributions

irregular behavior for sub-exponential or exponential distributions (Figure 7). In the DSS, the Jackson statistic is used to characterize distributions of the class *C*. Indeed, regularly varying distributions (class *C*) have asymptotically a power tail.

In the DSS, this method is considered as a confirmatory method for recommended decision based on the Log-Log and the MEF. So:

- If the curve converges clearly and regularly to 2 (Figure 7-a), the studied distribution belongs to the class C (regularly varying distribution). The use of: Fréchet (EV2), Halphen type IB (HIB), Log-Pearson type 3 (LP3), Inverse Gamma (IG) then are then recommended:
- If the curve presents some irregularities and do not converge to 2 (Figure 7-b), then we recommend the sub-exponential class (class *D*: Halphen type *A*, Gamma, Pearson type III, Halphen type *B*, Gumbel); or exponential (class *E*: Exponential distribution).

7. CASE STUDY: POTOMAC RIVER PEAK FLOW

In this section, the annual instantaneous peak flows of the Potomac River at Point of Rocks for the time period 1895-2006 (water year October-September) are studied. Figure 8 shows the observed annual peak flow time series. Smith (1987) and Katz et al. (2002) analyzed the same time series for the time period 1895-1986 and 1895-2000, respectively. To check (IID hypothesis) whether the observations are independent and if the data series are stationary and homogeneous we applied, respectively, the Wald-Wolfowitz, Kendall and Wilcoxon tests (Bobée and Ashkar, 1991). These tests indicated that the observed peak flow data series can be considered as independent and identically distributed (i.e. stationary and homogeneous).

The log-log plot corresponding to annual peak flow at Potomac is given in Figure 9. The plot shows that the curve is not perfectly linear, but the correlation coefficient is not significantly different to 1 at the level 5%. The DSS gives a decision to the user. In the case of the Potomac peak flows, the decision is displayed in Figure 10 and indicates that the curve is linear especially for high values of log(*u*) and thus the distribution belongs to the class *C* of regularly varying distributions.

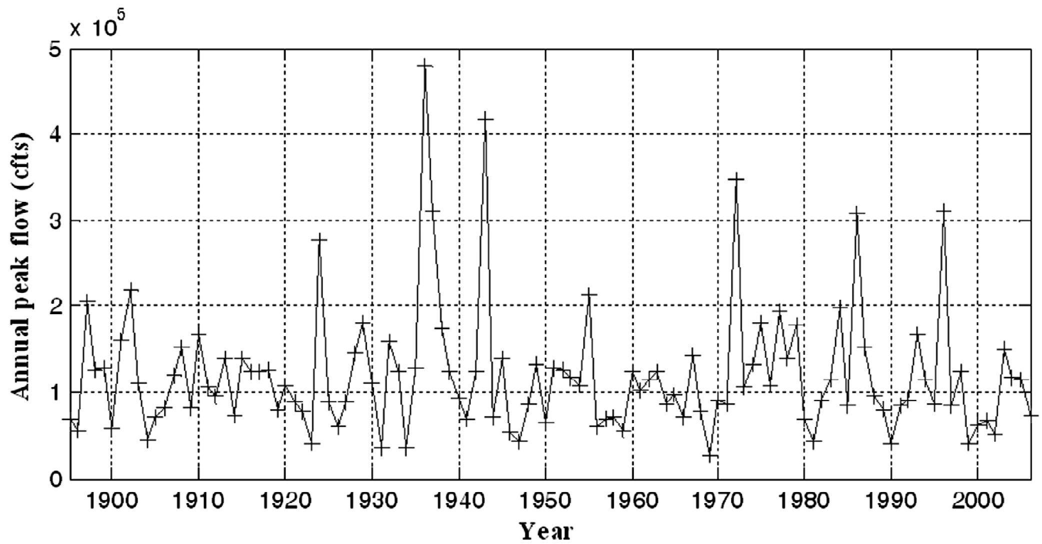

Figure 8: Annual Peak Flows of the Potomac River at Point of Rocks 1895-2006

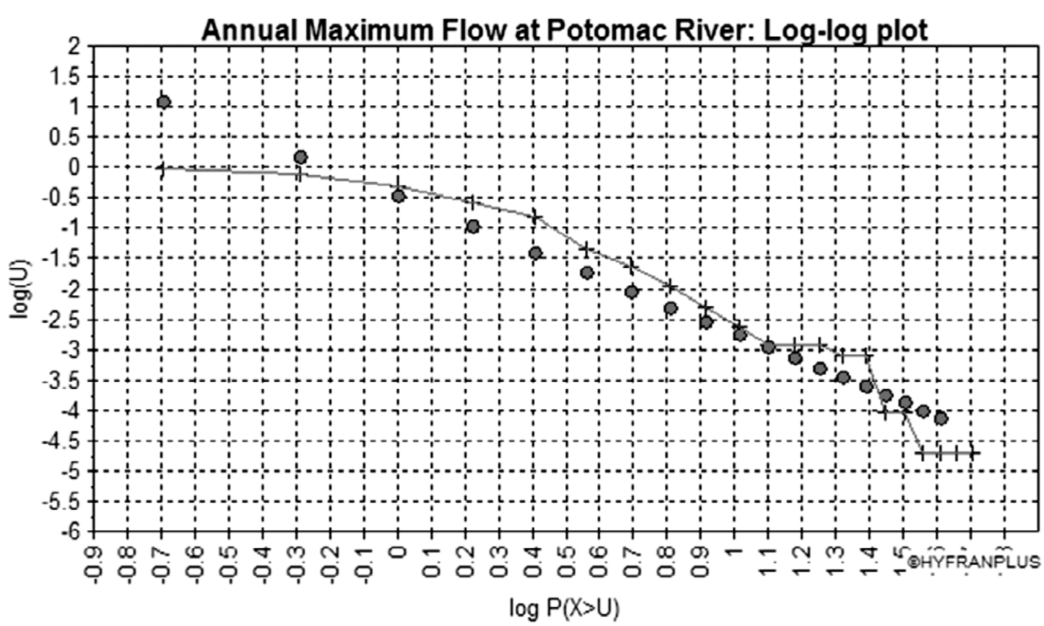

Figure 9: The Log-log Plot of the DSS for the Potomac Annual Maximum Flow

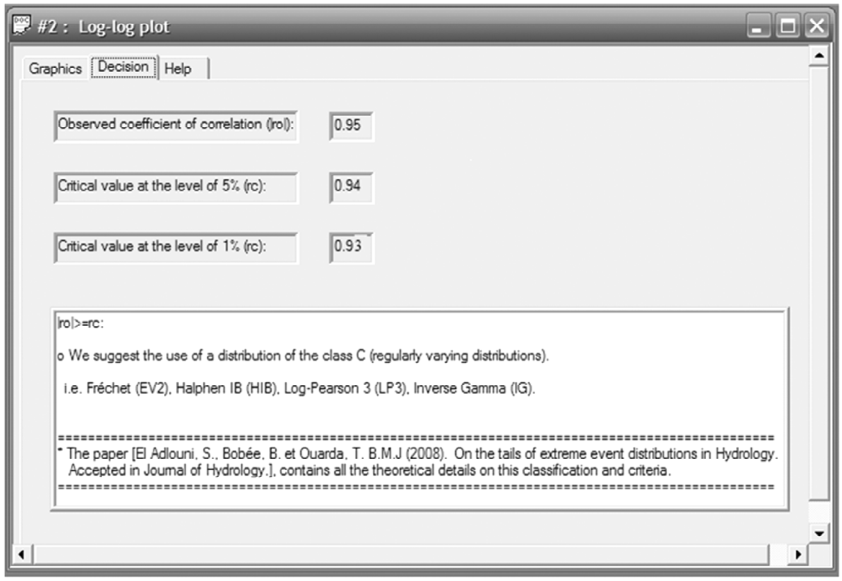

Figure 10: Output of the HYFRAN-PLUS When Executing The Log-Log Method of the DSS

At this stage, the DSS suggests the use of the Hill ratio plot and the Jackson statistic. These methods are considered for confirmatory analysis of the conclusions suggested by the previous method. The Hill ratio plot (Figure 11) shows a convergence towards a value different from zero.

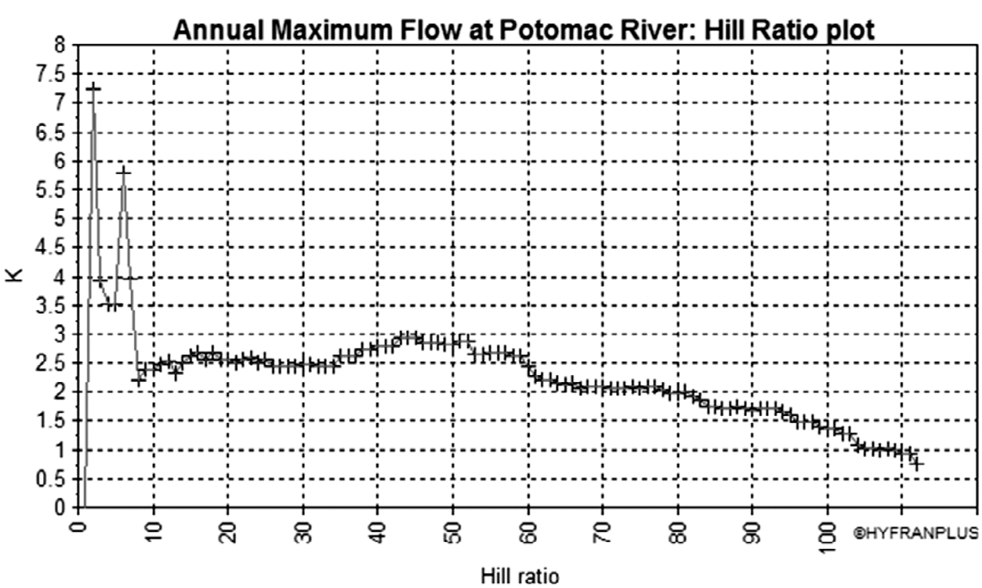

Figure 11: Hill's Ratio Plot of the DSS for Confirmatory Analysis

The adapted Jackson statistic (Figure 12) converges to its mean value 2 which implies that the studied data series have a power-law distribution (belong asymptotically to the class *B* which included in the class *C*). Both methods confirm that the most adequate distribution to fit the Potomac maximum peak data belongs to the class *C* of regularly varying distributions (EV2, Halphen type Inverse *B*, Inverse Gamma or Log-Pearson).

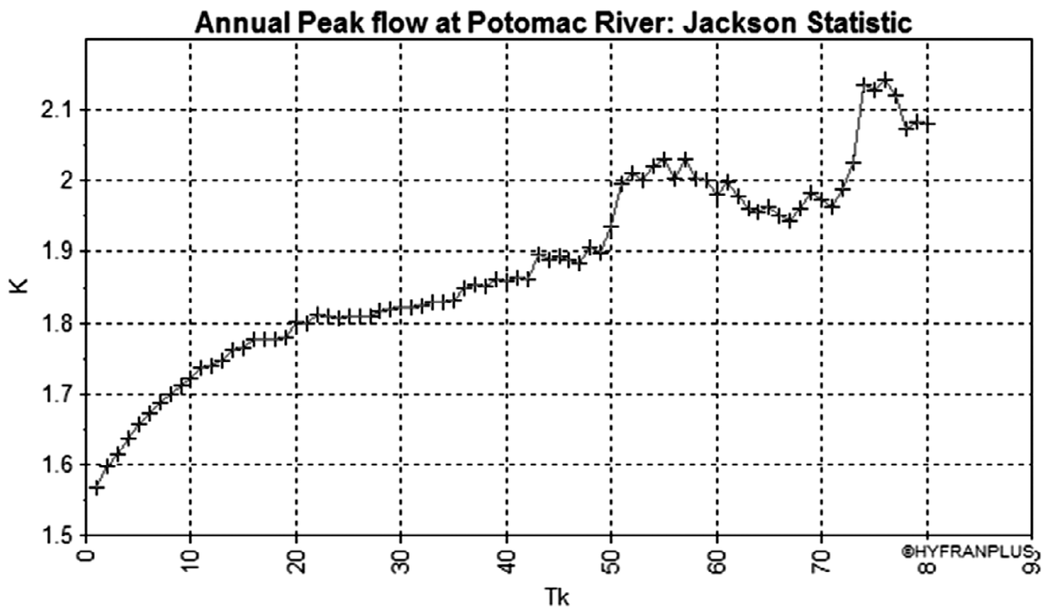

Figure 12: Modified Jackson Statistic of the DSS for Confirmatory Analysis

Once the class which represents adequately the studied dataset is selected, other criteria could be used to select the most adequate fit inside of each class. Classical tests and criteria, such as the Akaike Information Criterion (AIC, Akaike, 1974) or Bayesian Information Criterion (BIC, Fortin et al., 1998), can be used to select the distribution (EV2, HIB, IG and LP3) within the selected class. The AIC of the Inverse Gamma distribution is the smallest when compared to the distributions of the class *C* (EV2 and Log-Pearson type 3). These criteria are available in the software HYFRAN (CHS, 2002).

Table 1 presents the output of the Inverse Gamma distribution fit produced by the HYFRAN-PLUS software, which is the new version of HYFRAN including the Decision Support System (DSS). Return period events are estimated with their confidence interval at the level 95%. The software allows also presenting the fit results graphically.

Table 1
Results of the Fit of the Inverse Gamma Distribution to the Annual Peak Flow at Potomac River

Results of the fitting					
Inverse Gamma (Maximum Likelihood)					
Number of observations	*112*				
T	*q*	***XT***	*Standard deviation*	*Confidence intervals (95%)*	
1000.0	0.9990	**896**	256	640	1152
200.0	0.9950	**557**	184	441	761
100.0	0.9900	**453**	117	223	682
50.0	0.9800	**366**	72.7	223	508
20.0	0.9500	**271**	36.7	199	3437
10.0	0.9000	**212**	20.6	171	252
5.0	0.8000	**160**	10.9	139	182
3.0	0.6667	**126**	6.76	113	139
2.0	0.5000	**99.7**	4.80	90.3	109

8. CONCLUSIONS

Conventionally, for hydrological frequency analysis, a number of theoretical probability functions are fitted to the sample data and then statistical criteria such as Akaike Information Criterion (AIC) or Bayesian Information Criterion (BIC) are used to investigate the best fit. However, AIC and BIC are based on the likelihood function and thus give large weights to the central part of the distribution which contains a high proportion of the observations. Given that the main objective of the frequency analysis is to estimate the quantiles with high return periods (small probability of exceedance), the best fit should be checked for extremes rather than the center of the distribution.

In this article, it has been summarized how certain graphical methods used to discriminate distributions of the classes *C*, *D* and *E* can be used to select the most adequate fit especially for extremes. Although it is not useful in FFA, the class E is considered for more generality. Indeed the DSS can be used for any set of observations that are IID. The approach has been developed in C++ and integrated into the software HYFRAN-PLUS (http://www.wrpllc.com/books/hyfran.html) to help user to select the most adequate class. Within the selected class, the criteria such as the Akaike criterion and the Bayesian information may be used to choose the most adequate distribution.

The Log-normal distribution does not belong to any of the classes *C* and *D*. The *LN* has an asymptotic behavior which lies in the border of classes *C* and *D* (Figure 1). Indeed, the right tail of the *LN* distribution is lighter (respectively, heavier) than distributions of the class *C* (respectively of the class *D*). This problem has been recently studied to develop tools to test the Log-normality before the use of the DSS and a new version of the diagram (Figure 3) has been proposed (Martel et al. 2011).

Although HYFRAN and HYFRAN-PLUS have been developed for the fitting of hydrological data, they can be considered of observations that are independent and identically distributed (IID).

ACKNOWLEDGMENTS

The authors would like to thank Mr Ouejdene Samoud for his invaluable technical support to develop the C++ code version of the DSS.

References

[1] Akaike, H. (1974), "A New Look at the Statistical Model Identification", *IEEE Transactions on Automatic Control*, **19** (6): 716–723.

[2] Beirlant, J., de Wet, T. and Goegebeur, Y., (2006), "A Goodness-of-fit Statistic for Pareto-type Behaviour", *Journal of Computational and Applied Mathematics*, **186**, 99-116.

[3] Benson, M. A. (1968), "Uniform Flood-frequency Estimating Methods for Federal Agencies", *Water Resources Research*, **4**, 891-908.

[4] Bernier, J. (1959), "Comparaison des lois de Gumbel et de Fréchet sur l'estimation des débits maxima de crue", *La Houille Blanche (SHF)*, (1), Janv.-Fév. [681].

[5] Bobée, B. and Ashkar, F. (1991), "The Gamma Family and Derived Distributions Applied in Hydrology", *Water Resources Publications*, Littleton, Colo., 217p.

[6] Bobée, B. and Rasmussen, P.F. (1995), "Invited Paper : Recent Advances in Flood Frequency Analysis", U.S. National Report Contributions in Hydrology to International Union of Geodesy and Geophysics 1991-1994. *Review of Geophysics*: 1111-1116.

[7] Bobée, B. (1999), "Extreme Flood Events Valuation Using Frequency Analysis: A Critical Review", *Houille Blanche*, **54** (7-8), 100-105.

[8] Brooks, C. E. P., and Carruthers, N. (1953), *Handbook of Statistical Methods in Meteorology*, London: Her Majesty's Stationery Office, 85-111.

[9] CHS-Chaire en Hydrologie Statistique (2002), "HYFRAN : Logiciel Pour l'analyse Fréquentielle en Hydrologie", INRS-Eau, Rapport Technique. *Water Resources Publications* (http://wrpllc.com/books/hyfran.html).

[10] El Adlouni, S. et B. Bobée (2007), "Sampling Techniques for Halphen Distributions", *Journal of Hydrologic Engineering*, **12** (6): 592-604.

[11] El Adlouni, S., Bobée, B. and Ouarda, T.B.M.J. (2008), "On the Tails of Extreme Event Distributions", *Journal of Hydrology*, **355**, 16-33.

[12] Eslamian, S. S., 2010, "Flood Regionalization Using a Modified Region of Influence Approach", *Journal of Flood Engineering*, **1**(1): 51-66.

[13] Fisher, R. and L. Tippet (1928), "Limiting Forms of the Frequency Distribution of the Largest or Smallest Member of a Sample", *Proceedings of the Cambridge Philosophical Society*, **24**, 180-190.

[14] Fortin V., Bernier J. and Bobée B. (1998), Determination des crues de conception, Chaire en Hydrologie Statistique, Rapport N0 R-532.

[15] Fréchet, M. (1927), Sur la loi de Probabilité de l'Écart Maximum. Annales de la Société Polonaise de Mathématique, Cracovie **6**, 93-116. (in Johnson et al. 1995).

[16] Gumbel, E. J. (1942), "On the Frequency Distribution of Extreme Values in Meteorological Data", *Bull. Am. Meteor. Soc.*, **23**, 95.

[17] Halphen, E. (1941), Sur un nouveau type de courbe de fréquence. Comptes Rendus de l'Académie des Sciences, Tome 213, 633-635. Due to War Constraints, Published Under the Name "Dugué".

[18] Hill, B. M. (1975), "A Simple General Approach to Inference about the Tail of a Distribution", *Annals of Statistics*, **3**, 1163-1174.

[19] Hubert, P. (2005). La prédétermination des crues C. R. Geoscience 337.

[20] Hubert, P. and Bendjoudi H. (1996), "Introduction à l'étude des longues séries pluviométriques. In: Journées hyrologiques de l'ORSTOM", , *Montpellier*, pp. 20.

[21] Jackson, O.A.Y., (1967), "An Analysis of Departures from the Exponential Distribution", *Journal of the Royal Statistical Society B*, **29**, 540-549.

[22] Katz, R. W., Parlange M. B. and Naveau P., (2002), "Statistics of Extremes in Hydrology", *Advances in Water Resources*, **25:**1287-1304.

[23] Kidson, R., K.S. Richards and P.A. Carling (2005), "Reconstructing the ca. 100-year Flood in Northern Thailand", *Geomorphology*, **70**, 279-295.

[24] Koutsoyiannis, D. (2004), "On the Appropriateness of the Gumbel Distribution for Modelling Extreme Rainfall", In *Hydrological Risk: Recent Advances in Peak River Flow Modelling, Prediction and Real-time Forecasting*, Assessment of the Impacts of Land-use and Climate Changes (Ed. by A. Brath, A. Montanari & E. Toth), 303–319. Editoriale Bios, Castrolibero, Bologna, Italy.

[25] Martel, B., S. El Adlouni and B. Bobée (2011), "Comparison of the Power of Log-Normality Tests with Different Right Tail Alternative Distributions", *Journal of Hydrologic Engineering*, Under Review, August 2011.

[26] Morlat, G. (1951), Note sur l'estimation des débits de crues. La Houille Blanche, N° Spécial B, p. 663-681.

[27] Morlat, G. (1956). Les lois de probabilité de Halphen. Revue de Statistique Appliquée, 3, 21-43.

[28] Rao, A. R. and Hamed, K. H., (2000), "Flood Frequency Analysis", *CRC Press*, Boca Raton, Florida, USA.

[29] Robson, A., and Reed, D.W., 1999, "Flood Estimation Handbook", *Statistical Procedures for Flood Frequency Estimation*, **3**, Centre for Ecology and Hydrology, Wallingford, UK, p. 108.

[30] Singh, V.P., and Strupczewski W.G., (2002), "On the Status of Flood Frequency Analysis, Hydrological Processes, **16** (18), 3737-3740.

[31] Smith J.A. (1987), "Estimating the Upper Tail of Flood Frequency Distributions", *Water Resour Res*; **23:**1657-66.

[33] Stedinger, J.R., Vogel, R.M., and Foufoula-Georgiou, E. (1993), "Frequency Analysis of Extreme Events", *Handbook of Hydrology*, Chapitre 18. Éditeur : D.R. Maidment. McGraw Hill, New York.

[34] US Water Resources Council, (1967), "A Uniform Technique for Determining Flood Flow Frequencies", *Bulletin #15 of the Hydrology Committee*, December 1967.

[34] Werner, T. and C. Upper (2002), "Time Variation in the Tail Behaviour of Bund Futures Returns", Working Paper N°199, European Central Bank.

[35] Bobée, B., Perreault, L. and Ashkar, F. (1993), "Two Kinds of Moment Ratio Diagrams and Their Applications in Hydrology", *Stochastic Hydrology and Hydraulics*, **7**, Issue 1, pp.41-65.