

Estimation des événements extrêmes de crue par l'analyse fréquentielle : une revue critique

Extreme flood events valuation using frequency analysis : a critical review

par B. Bobée

Université du Québec, Institut national de la recherche scientifique (INRS-Eau)
Chaire industrielle en Hydrologie statistique (CRSNG - Hydro-Québec)

In 1968, the U.S. Water Resources Council recommended, for flood frequency analysis, the systematic use of the Log-Pearson type III distribution, fitted by the indirect method of moments. This recommendation was the starting point of numerous studies involving

- proposing other parameter estimation methods for this distribution;*
 - taking into account other distributions, such as Generalized Extreme Value (GEV), for which Hosking (1990) [18] proposed an estimation method based on probability weighted moments (L-moments).*
- A critical review of these theoretical developments is presented herein. The effect of separation is being discussed, as it is considered a selection criterion for representative flood distribution functions. Also, we present recent results concerning the Halphen distributions.*

I ■ INTRODUCTION

L'analyse fréquentielle des événements extrêmes, et plus particulièrement des crues, a donné lieu à de nombreux travaux au cours des dernières décennies.

Dans l'estimation locale, on considère généralement un échantillon de n valeurs de la variable aléatoire débit maximum annuel (X). On fait l'hypothèse que ces observations sont :

- indépendantes (absence d'autocorrélation) et ;
- identiquement distribuées (homogénéité, stationarité, absence de valeurs singulières) selon une distribution statistique (D) de fonction densité de probabilité $f(x, \theta)$ et de fonction de distribution $F(x, \theta)$ où θ est un vecteur de paramètres que l'on doit estimer à partir de l'échantillon.

En pratique, on s'intéresse à la probabilité au dépassement $p = 1/T$ d'un quantile X_T de période de retour T ; on a :

$$Pr[X \geq x_T] = 1 - F(x_T; \theta) = 1 / T$$

L'événement X_T est donné par :

$$X_T = F^{-1}[1 - 1/T; \theta]$$

et son estimation \hat{X}_T est obtenue par :

$$\hat{X}_T = F^{-1}[1 - 1/T; \hat{\theta}] \text{ où } \hat{\theta} \text{ représente l'estimation du vecteur } \theta.$$

Il est reconnu [1] qu'il est nécessaire de considérer des lois à trois paramètres [$\theta = (\theta_1 ; \theta_2 ; \theta_3)$] si l'on veut caractériser la forme de la distribution même si la faible taille des échantillons généralement disponibles en hydrologie pourrait inciter à ne considérer que des lois à deux paramètres. Si l'échantillon n'est pas trop petit ($n > 50$), \hat{X}_T est non biaisé et distribué approximativement selon une loi normale de moyenne X_T et de variance $\text{var}(X_T)$; on peut en déduire l'intervalle de confiance au niveau $(1-\alpha)$ de la vraie valeur inconnue X_T .

Il existe en général plusieurs méthodes d'estimation (E) pour une distribution (D). Pour permettre la comparaison des (D/E), Cunnane [2] a introduit les notions de :

- capacité descriptive (*descriptive ability*) qui mesure l'aptitude à reproduire les valeurs observées de l'échantillon ; et
- capacité prédictive (*predictive ability*) qui mesure l'aptitude d'une distribution à l'extrapolation.

Pour déterminer la capacité prédictive d'une procédure (D/E), on procède par simulation de Monte-Carlo, ce qui permet de calculer le biais (B), la variance (V) et l'écart quadratique moyen EQM (on a : $EQM = B^2 + V$) de l'estimation d'un événement X_T .

De manière générale, les méthodes d'estimation (E) sont des variantes de la méthode des moments qui consiste à estimer des moments théoriques de la population (qui sont des fonctions des paramètres de la distribution) par les moments

correspondants de l'échantillon. La résolution du système d'équations obtenues permet d'estimer les paramètres.

La méthode du maximum de vraisemblance (MV) qui, pour un échantillon (x_1, \dots, x_n) de taille n consiste à maximiser la fonction vraisemblance (ou son logarithme $L = \prod_{i=1}^n f(x_i, \theta)$)

peut conduire à de mauvaises estimations pour la plupart des lois à trois paramètres considérées en hydrologie. En effet, ces lois dépendent, en général, d'un paramètre d'origine (m) qui détermine l'intervalle de définition de la variable X ($X > m$ par exemple); dans ce cas, le MV est seulement asymptotiquement optimal (i.e., sans biais et avec une variance minimum pour n élevé) [3]. Ainsi pour les échantillons de petite taille ($n < 50$), souvent obtenus en hydrologie, cette optimalité n'est pas atteinte en général.

Pour certaines distributions (normale, exponentielle, Gamma), on peut trouver des statistiques conjointement exhaustives des paramètres. Dans ce cas, la méthode de MV conduit à une estimation optimale des paramètres, quelle que soit la taille de l'échantillon.

Dans ce qui suit, on présente succinctement les (D/E) qui ont été le plus couramment utilisées pour l'estimation locale.

II ■ ESTIMATION LOCALE

● 2.1 Loi Log-Pearson type 3 (LP)

2.1.1 Méthode indirecte des moments (MIM)

Les diverses agences locales, provinciales ou fédérales des Etats-Unis utilisaient pour l'estimation des débits des crues des procédures souvent différentes pouvant conduire à des résultats discordants. C'est pourquoi, à la suite d'une comparaison d'ajustement sur des données réelles, le Conseil des Ressources en eau a recommandé [4], pour l'estimation des quantiles de crue, l'utilisation systématique de la distribution Log-Pearson type 3 définie de la manière suivante, si $Y = \text{Log } X$ suit une distribution Pearson type 3 $P(\alpha, \lambda, m)$ alors X suit une distribution Log-Pearson type 3 LP (α, λ, m). Cette recommandation réconciliait les tenants des deux lois qui étaient alors les plus couramment utilisées pour représenter approximativement les crues :

- la distribution Pearson type 3 ;
- la distribution Log-normale (ou loi de Galton-Gibrat). La transformation logarithmique ayant pour effet de diminuer l'asymétrie généralement observée dans les séries de débit de crue.

La méthode d'estimation proposée (méthode indirecte des moments) [4] consiste à ajuster la distribution $P(\alpha, \lambda, m)$ à la série des logarithmes des valeurs observées par la méthode des moments en considérant la moyenne (μ), la variance (σ^2) et le coefficient d'asymétrie $C_s = \mu_3/\sigma^3$.

Cette méthode indirecte des moments [4] utilise le coefficient d'asymétrie de l'échantillon qui est une estimation biaisée et bornée du coefficient d'asymétrie de la population :

$$\left(|C_s| < \frac{n-2}{\sqrt{n-1}} \right)$$

Il en résulte une sous-estimation du coefficient d'asymétrie pour des échantillons de petite taille [5]. Afin d'améliorer l'estimation de C_s , le Conseil des Ressources en eau [4] propose d'utiliser une moyenne pondérée de l'estimation au site et d'une estimation régionale. Dans le but de prendre en compte la grande erreur d'échantillonnage de C_s associée à de petits échantillons, le coefficient de pondération dépend de la taille de l'échantillon.

2.1.2 Méthode directe des moments (MDM)

La méthode précédente conserve les moments de la distribution (P) mais non ceux de la distribution (LP) et donne un poids identique aux logarithmes des observations mais non aux observations elles-mêmes. C'est pourquoi Bobée [6] a proposé la méthode directe des moments qui consiste à éga-ler les trois premiers moments non centrés de la population à leur estimation obtenue à partir de l'échantillon. La résolution du système d'équations obtenues permet d'estimer les paramètres α, λ et m de la loi LP (α, λ, m).

Pour contourner les difficultés d'estimation associées à l'estimation d'un moment d'ordre 3 (qui est très sensible aux grandes valeurs de l'échantillonnage), plusieurs autres méthodes ont été proposées en considérant des moments d'ordre moins élevé.

2.1.3 Méthodes des moments mixtes

Ces méthodes (tabl. 1) considèrent pour estimer les trois paramètres de la distribution LP, trois moments choisis parmi :

- les deux premiers moments de la variable x qui suit une loi LP ; et
- les deux premiers moments de la variable $Y = \text{Log } X$ qui suit une loi P .

Tableau 1. — Méthodes des moments mixtes : loi Log-Pearson type 3.

	MIM	MDM	MM1	MM2	MM3	MM4
$\mu'_1 (LP)$		•	•	•	•	
$\mu'_2 (LP)$		•	•	•		•
$\mu'_3 (LP)$		•				
$\mu'_1 (P)$	•		•		•	•
$\mu'_2 (P)$	•			•	•	•
$\mu'_3 (P)$	•					

Les méthodes MM1 et MM2 proposées par Rao [7] conservent les deux premiers moments de la loi LP (espace réel) et l'un des deux premiers moments de la loi P (espace des Log). Les méthodes MM3 et MM4 proposées par Phien et Hira [8] conservent les deux premiers moments de la loi P (espace des Log) et l'un des deux premiers moments de la loi LP (espace réel). Il est intéressant de noter que :

$$\mu'_1(P) = E[\text{Log } X] = \text{Log } G$$

G peut être considéré [9] comme le moment μ'_0 d'ordre quasi-zéro

$$\left(G = \lim_{r \rightarrow 0} \left[\frac{1}{n} \sum x_i^r \right] \right) \text{ de la distribution LP.}$$

La méthode MM1 est donc en fait une méthode directe des moments qui préserve les moments (0, 1, 2) de la distribution LP.

2.1.4 Méthode SAM (Sundry Averages Method)

Cette méthode, proposée par Bobée et Ashkar [10], conserve la moyenne harmonique (ordre -1), géométrique (ordre 0) et arithmétique (ordre 1).

Une comparaison par simulation [11] des différentes méthodes d'ajustement de la loi LP montre que la méthode directe des moments donne les meilleurs résultats pour les grandes périodes de retour (extrapolation) alors que la

méthode SAM est adéquate pour l'estimation des bas débits.

Après les Etats-Unis, l'Australie [12] recommandait l'utilisation de la loi LP. Cependant, à la même période, le choix de cette loi était remis en question par :

- la proposition en Grande-Bretagne de la loi généralisée des valeurs extrêmes G.E.V. [13] ; et
- la mise en évidence de l'effet de séparation [14].

● 2.2 Loi G.E.V. (loi généralisée des valeurs extrêmes)

Si l'on note X la plus grande (ou la plus petite) valeur d'un échantillon (y_1, \dots, y_p) de taille p (par exemple, le débit maximum annuel journalier avec $p = 365$). Si p est grand et si les réalisations de y_j sont indépendantes et identiquement distribuées, la théorie des valeurs extrêmes [15] montre l'existence de trois formes asymétriques de la distribution de X . Jenkinson [16] a montré que les trois lois des valeurs extrêmes pouvaient être regroupées en une forme unique :

$$F(x) = \exp \left[- \left[1 - \frac{k}{\alpha}(x-u) \right]^{1/k} \right] \alpha > 0$$

- la distribution EV1, connue également sous les noms de distribution de Gumbel ou double exponentielle, qui comprend deux paramètres correspond à $k = 0$;
- la distribution EV2 (obtenue pour $k < 0$) est connue sous les noms de Frechet ou de Log-Gumbel (si $X \sim EV2$ alors $\text{Log } X \sim EV1$;
- la distribution EV3 obtenue pour $k > 0$ est appelée aussi Weibull négative (si $X \sim EV3$ alors $X \sim WEI$).

Ces deux dernières distributions à trois paramètres (EV2 et EV3) sont bornées (supérieurement pour EV3 et inférieurement pour EV2) par une fonction des paramètres $(u + \alpha/k)$. L'utilisation de la distribution G.E.V. pour représenter les crues a été recommandée en Grande-Bretagne par le Conseil de Recherches en environnement [13].

Les hypothèses qui sont à la base de la théorie des valeurs extrêmes ne sont pas rigoureusement respectées en hydrologie :

- le débit maximum annuel est le maximum de la série des débits annuels avec $p = 365$; cette valeur ne peut être considérée comme élevée dans le cadre d'une théorie asymptotique, et
- les débits journaliers ne sont généralement pas indépendants et identiquement distribués.

La distribution G.E.V. et ses cas particuliers ne sont pas *a priori* de plus mauvais candidats que les autres distributions utilisées en analyse fréquentielle. Le cas EV1 (quand $k \rightarrow 0$) de la distribution G.E.V. est donné par :

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

Pour les distributions G.E.V. et EV1, on peut obtenir une relation explicite $x(F)$:

G.E.V. : $x = u + \frac{\alpha}{k} [1 - (-\text{Ln } F)^k]$;

EV1 : $x = u - \alpha \text{Ln}(-\text{Ln } F)$

La méthode d'ajustement de la loi G.E.V. la plus utilisée est basée sur la méthode des moments pondérés (G.E.V./PWM). Greenwood *et al.* [17] ont défini les moments pondérés (*probability weighted moments*) par :

$$M_{p,r,s} = E[X^p F^r (1-F)^s] = \int_0^1 [x(F)]^p F^r (1-F)^s dF$$

En pratique, on utilise :

$$\alpha_r = M_{1,0,r} = \int_0^1 x(F) (1-F)^r dF$$

$$\beta_r = M_{1,r,0} = \int_0^1 x(F) F^r dF$$

Pour un échantillon de taille n classé par ordre croissant $x_1 \leq x_2 \leq \dots \leq x_n$, β_r est estimé par :

$$b_r = \frac{1}{n} \sum_{j=1}^n \frac{(j-1)!}{(j-r-1)!} \frac{(n-r-1)!}{(n-1)!} x_j$$

La méthode G.E.V./PWM consiste à résoudre :

$$\beta_r = b_r \quad r = 1, 2, 3$$

où β_r est une fonction des paramètres et b_r est obtenu à partir de l'échantillon.

Hosking [18] a introduit les L -moments λ_r qui sont des combinaisons linéaires des moments pondérés ; on a :

$$\lambda_{r+1} = (-1)^r \sum_{k=0}^r p_{r,k} \alpha_k \text{ avec } p_{r,k} = (-1)^{r-k} C_r^k C_{r+k}^k = \frac{(-1)^{r-k} (r+k)!}{(k!)^2 (r-k)!}$$

Les quatre premiers L -moments sont donnés par :

$\lambda_1 = \alpha_0$; $\lambda_2 = \alpha_0 - 2 \alpha_1$; $\lambda_3 = \alpha_0 - 6 \alpha_1 + 6 \alpha_2$; $\lambda_4 = \alpha_0 - 12 \alpha_1 + 30 \alpha_2 - 20 \alpha_3$ et ont un sens physique comparable aux moments classiques de même ordre (moyenne, dispersion, asymétrie, aplatissement). On peut également définir des rapports de L -moments similaires aux coefficients de variation, d'asymétrie et d'aplatissement :

$$LC_v = \frac{\lambda_2}{\lambda_1} ; \quad LC_s = \frac{\lambda_3}{\lambda_2} ; \quad LC_k = \frac{\lambda_4}{\lambda_2}$$

Les L -moments ont été utilisés au cours des dernières années pour l'estimation des paramètres, l'identification des distributions (diagramme des L -moments) et les études de régionalisation.

Les estimateurs des L -moments et des rapports de L -moments sont très peu biaisés et ont une faible variance. Pour cette raison, le diagramme des rapports $[LC_s ; LC_k]$ a été recommandé [19] pour identifier les distributions à la place du diagramme classique :

$$[\beta_1 = (C_s)^2 ; \quad \beta_2 = C_k - 3].$$

Par définition, les L -moments sont robustes aux valeurs extrêmes d'un échantillon. Mais dans le cas d'une valeur élevée observée et validée, cette robustesse peut être un inconvénient puisque l'on accorde un poids trop faible à une information importante concernant la queue de la distribution (Bernier 1993, communication personnelle ; [20]).

● 2.3 Effet de séparation

Matalas *et al.* [14] ont considéré les données de débit maximum annuel obtenues à 1 351 stations réparties dans 14 régions géographiques des Etats-Unis. Les séries obtenues ont ensuite été décomposées en échantillons disjoints de taille $n = 10, 20, \dots, 30$. On obtient ainsi, dans chaque région $k(n)$ échantillons de taille n , pour lesquels on calcule le coefficient d'asymétrie ($g = m_3/s^3$). On peut en déduire pour chaque région :

• la moyenne : $\bar{g} = \frac{1}{k(n)} \sum_{i=1}^{k(n)} g_i$

• l'écart-type : $\sigma(g) = \left[\frac{\sum_{i=1}^{k(n)} g_i^2 / k(n) - \bar{g}^2}{k(n)} \right]^{1/2}$

Pour différentes distributions (D) standardisées (moyenne 0, variance 1), Matalas *et al.* [14] ont simulé $p = 10^5$ échantillons de taille n correspondant à une asymétrie théorique γ et à partir des coefficients d'asymétrie obtenus pour l'ensemble des p échantillons, ils ont déduit la moyenne $\bar{\gamma}$ et l'écart-type $\sigma(\bar{\gamma})$.

Matalas *et al.* [14] ont porté sur le même graphique (fig. 1) la moyenne et l'écart-type des valeurs observées ($\bar{g}; \hat{\sigma}(g)$) et des valeurs simulées $\bar{\gamma}, \sigma(\gamma)$.

Matalas *et al.* [14] ont considéré sept distributions (normale, uniforme, Gumbel, Pareto, Log-normale, Weibull et Pearson type 3) et pour chaque taille d'échantillon ($n = 10, \dots, 30$), ils ont obtenu une représentation similaire à la figure 1 correspondant à $n = 20$. La courbe théorique relative à chaque distribution est située nettement sous les points correspondant aux données observées. La courbe supérieure étant obtenue pour la loi Log-normale, ils ont nommé ce « phénomène », effet de séparation, et en ont déduit qu'aucune des distributions considérées n'était adéquate pour représenter les données observées. Des études similaires effectuées en Italie [21] et en Grande-Bretagne [22] ont conduit à des résultats similaires.

L'effet de séparation est devenu un critère de sélection pour choisir une distribution représentative des débits de crue (la distribution idéale ne devant pas reproduire cet effet de séparation). D'autres auteurs ont examiné si certaines distributions non considérées par Matalas *et al.* [14] pouvaient rendre compte de la variabilité de l'asymétrie régionale mise en évidence (fig. 1). Landwehr *et al.* [23] en considérant la distribution LP 3 et Cunnane (1984 ; non publié) la distribution G.E.V. ont montré que ces lois n'étaient pas satisfaisantes à cet égard.

Cependant plusieurs auteurs ([24], [25]) ont émis de sérieuses réserves sur l'utilisation de l'effet de séparation comme critère de sélection. Ashkar *et al.* [26] ont démontré en appliquant la méthode du *bootstrap* à des données réelles de l'Utah (partie des régions 9 et 10 de l'étude de Matalas *et al.* [14]) que l'effet de séparation était en fait dû à la combinaison dans une même région d'échantillons ayant différentes valeurs du coefficient d'asymétrie (*spatial mixing*).

L'hypothèse de l'inadéquation (tout au moins sur la base de l'effet de séparation) des distributions couramment utilisées pour représenter les débits de crue doit donc être rejetée.

● 2.4 Lois de Halphen (Types A, B, B⁻¹)

Un historique des lois de Halphen est donné par Morlat [1]. Ces lois ont été introduites successivement depuis 1941 dans le but de respecter certaines contraintes :

- les fonctions de densité ne doivent pas avoir de paramètre d'origine (borne inférieure nulle) ;
- les distributions proposées doivent modéliser les comportements algébrique et exponentiel des queues de distribution ;
- l'estimation des paramètres doit être exhaustive (variance minimum).

Loi harmonique (1941) :

$$f(x) = \frac{1}{2xK_0(2\alpha)} \exp\left[-\alpha\left(\frac{x}{m} + \frac{m}{x}\right)\right] \quad x > 0$$

fonction définie pour $x > 0$ avec $m > 0$ (paramètre d'échelle) et $\alpha > 0$ (paramètre de forme)

Loi de Halphen de type A (1946) : HA

$$f_A(x) = \frac{1}{2m^\nu K_\nu(2\alpha)} x^{\nu-1} \exp\left[-\alpha\left(\frac{x}{m} + \frac{m}{x}\right)\right] \quad x > 0$$

Cette loi (avec $m > 0$ et $\alpha > 0$) généralise la loi harmonique en introduisant le paramètre de forme ν (qui peut être positif ou négatif). $K_\nu(2\alpha)$ est la fonction de Bessel modifiée de seconde espèce ($K_0(2\alpha)$ est le cas particulier correspondant à $\nu = 0$).

Loi de Halphen de type B (1948) : HB

Cette loi a été introduite par Halphen pour représenter des comportements asymptotiques différents de ceux obtenus pour la loi de type A :

$$f_B(x) = \frac{2}{m^{2\nu} e f_0(\alpha)} x^{2\nu-1} \exp\left[-\left(\frac{x}{m}\right)^2 + \alpha\left(\frac{x}{m}\right)\right] \quad x > 0$$

fonction définie pour $x > 0$ où $\nu > 0$ et α sont des paramètres de forme, et $m > 0$ un paramètre d'échelle.

La fonction $e f_0(\alpha) = \int_0^\infty x^{2\nu-1} e^{-x^2 + \alpha x} dx$

a été nommée, par Halphen, fonction exponentielle [27] et est en réalité une fonction de Hermite [1].

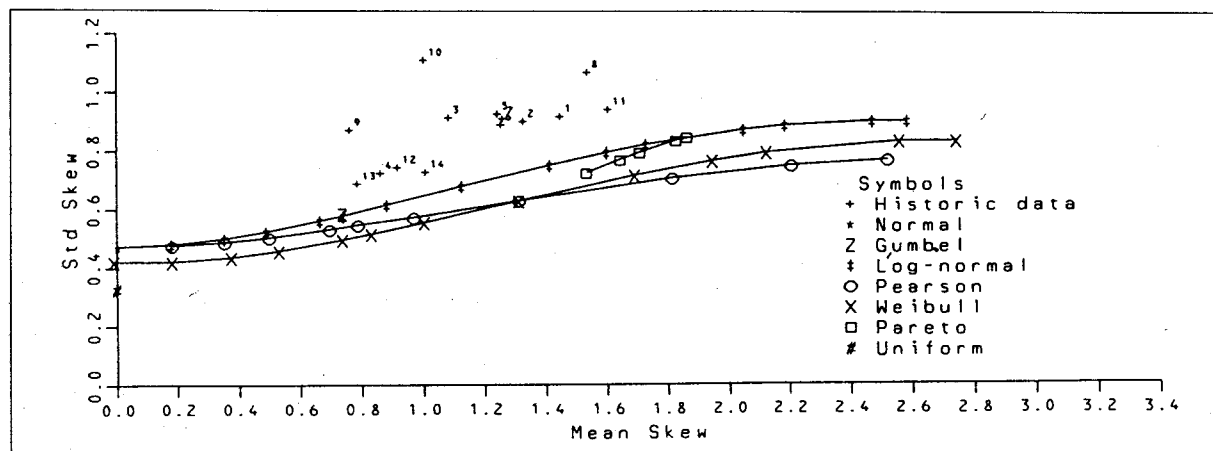
Loi de Halphen de type B⁻¹ : HB⁻¹

Cette loi qui complète le système de Halphen a été introduite par Larcher [1] :

$$f_{B^{-1}}(x) = \frac{2}{m^{-2\nu} e f_0(\alpha)} x^{-2\nu-1} \exp\left[-\left(\frac{m}{x}\right)^2 + \alpha\left(\frac{m}{x}\right)\right] \quad \text{avec } x > 0 ;$$

Cette distribution est déduite de la loi de type B par une transformation inverse ($X \sim \text{HB} \rightarrow 1/X \sim \text{HB}^{-1}$). Les lois de type A ($\nu > 0$) et B comprennent comme cas limite la loi Gamma alors que les lois de type A ($\nu < 0$) et B⁻¹ convergent vers la loi Gamma inverse.

Le grand intérêt des lois de Halphen et de leurs cas limites (Gamma et Gamma inverse) est de posséder des statistiques conjointement exhaustives des paramètres (tabl. 2). Il en résulte que, quelle que soit la taille de l'échantillon considéré, l'estimation des paramètres est optimale et correspond aux estimateurs du maximum de vraisemblance lorsqu'une solution existe.



1. Effet de séparation $n = 20$ [tiré de Matalas *et al.* (1975) [14]].

Tableau 2. — Statistiques exhaustives des lois du système Halphen

$$Q = \frac{1}{n} \sum x_i^2; (Q)^{-1} = \frac{1}{n} \sum \frac{1}{x_i^2}$$

Moments	LOIS				
	HA	HB	HB ⁻¹	G	GI
- 2 (QI)			•		
- 1 (H)	•		•		•
0 (G)	•	•	•	•	•
1 (A)	•	•		•	
2 (Q)		•			

On peut montrer [28] que :

- dans le cas de la loi HA, la solution du MV existe si

$$|v| < U = \frac{A}{H} / \left(\frac{A}{H} - 1 \right)$$

si $|v| > U$, la dispersion des observations est trop grande et les estimateurs du maximum de vraisemblance convergent vers ceux de la loi Gamma ($v > 0$) ou de la loi Gamma inverse ($v < 0$) ;

- dans le cas de la loi HB, la solution du MV est obtenue pour $v < V = 1/2[QA^2 - 1]$

si $v > V$, la dispersion des observations est trop grande et les estimateurs du MV convergent vers ceux de la loi Gamma ;

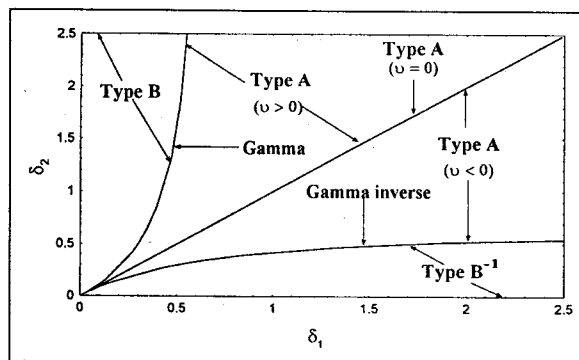
- dans le cas de la loi HB⁻¹, il y a solution du MV si $v < W = 1/2[(H^2/QI) - 1]$ autrement (si $v > W$) il y a convergence vers les estimateurs de la loi Gamma inverse.

On peut représenter [1] les distributions de cette famille dans le diagramme des rapports de moments : $\{\delta_1 = \text{Ln } A/G ; \delta_2 = \text{Ln } G/H\}$ (fig. 2a). La représentation dans ce diagramme (δ_1, δ_2) de diverses distributions de la famille Gamma utilisées en hydrologie (fig. 2b) a été effectuée par Bobée *et al.* [29] et a été comparée avec le diagramme (β_1, β_2) de Pearson.

Perreault *et al.* [30] ont étudié les propriétés des estimateurs des lois de Halphen et effectué la détermination des variances asymptotiques des quantiles X_T .

III ■ CONCLUSION ET PERSPECTIVES

Pour une distribution donnée (D), il est possible de comparer les diverses méthodes d'estimation (E) disponibles par simulation d'un nombre important d'échantillons. On peut en déduire le biais et la variance de l'estimation X_T associé à la période de retour T . Il est alors possible de caractériser chaque méthode (E) en fonction de ses capacités descriptive (interpolation) et prédictive (extrapolation).



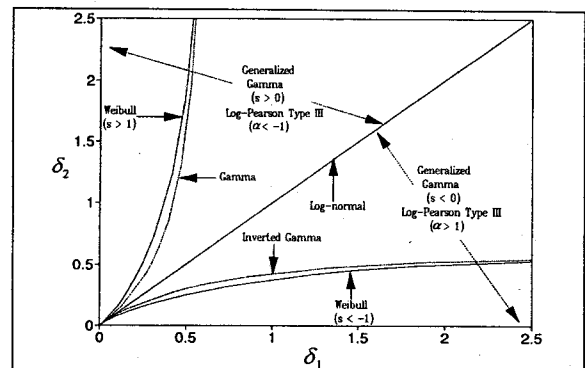
2. Diagramme (δ_1, δ_2) a : Lois de Halphen ; b : Famille Gamma.

Les tentatives (théorie des valeurs extrêmes, par exemple) visant à déterminer la loi théorique $g(x)$ que devrait suivre la variable débit maximum annuel n'ont pas donné de résultats incontestés. On admet de plus en plus que la distribution théorique de la variable est inconnue et que l'on cherche à en obtenir la meilleure approximation possible en puisant dans l'arsenal des distributions connues. On est ainsi amené à considérer diverses distributions plausibles. Les tests d'adéquation (Chi-carré, Kolmogorov-Smirnov) sont très peu puissants lorsque les paramètres de la distribution ne sont pas connus *a priori*, ce qui est le cas en pratique. Pour la loi normale et par extension la loi Log-normale, on peut considérer des tests plus puissants (Wilks et Shapiro, par exemple). Le critère d'information d'Akaike $AIC = -2 \text{Log } L(\hat{\theta}, \dots, x) + p$ qui combine la fonction vraisemblance L et le nombre (p) de paramètres de la distribution permet une comparaison globale des distributions.

Une étude plus approfondie du comportement des queues de distribution est en cours pour choisir les lois qui donnent une meilleure représentation des valeurs extrêmes d'un échantillon. Une approche bayésienne a également été considérée pour obtenir la probabilité *a posteriori* $p[H/x]$ de l'hypothèse $H_r (r = 1, \dots, R)$ qu'une distribution f_r représente la distribution $g(x)$ inconnue en tenant compte de l'information contenue dans l'échantillon x .

BIBLIOGRAPHIE

- [1] MORLAT G. (1956). Les lois de probabilité de Halphen. *Revue de Statistique appliquée*, 3 : 21-43.
- [2] CUNNANE C. (1987). Dans : *Hydrologic frequency modeling*. V.P. Singh (Ed.), D. Reidel Publishing Co., Dordrecht, pp. 49-95.
- [3] BOBÉE B. (1979). Comment on : The Log-Pearson type 3 distribution : The T-year event and its asymptotic standard error by maximum likelihood theory. *Water Resources Research*, 15(1) : 189-190.
- [4] WRC (1967). A uniform technique for determining flood flow frequencies, Bulletin 15 ; (1977) Guidelines for determining flood flow frequency, Bulletin 17A. *US Water Resources Council, Hydrology Committee*, Washington, DC.
- [5] GREIS N.P. (1983). Flood frequency analysis : A review of 1979-1982. *Reviews of Geophysics and Space Physics*, 21(3) : 699-706.
- [6] BOBÉE B. (1975). The Log-Pearson type 3 distribution and its application in hydrology. *Water Resources Research*, 11(5) : 681-689.
- [7] RAO D.V. (1980). Log-Pearson type 3 distribution : Method of mixed moments. *J. Hydraul. Div. ASCE*, 106(6) : 999-1019.



- [8] PHUEN H.N. and HIRA M.A. (1983). Log-Pearson type 3 distribution parameter estimation. *J. Hydrol.*, 54 : 25-37.
- [9] BOBÉE B. et ASHKAR F. (1988a). The generalized method of moments applied to the Log-Pearson type 3 distribution. *J. Hydraul. Engineering ASCE*, 114(8) : 899-909.
- [10] BOBÉE B. and ASHKAR F. (1988b). Sundry averages method for the estimating parameters of the Log-Pearson type distribution. *INRS-Eau*, rapport scientifique R-251.
- [11] ASHKAR F. and BOBÉE B. (1987). The generalized method of moments as applied to problems of flood frequency analysis : Some practical results for the Log-Pearson type 3 distribution. *J. Hydrol.*, 90(3-4) : 199-217.
- [12] I.E.A. (1977). Australian rainfall and runoff flood analysis and design. *Institute of Engineers, Australia*, A.C.T., 149 p.
- [13] NERC (1975). Flood studies report — vol. 1 : Hydrological studies. *Natural Environment Research Council*, London.
- [14] MATALAS N.C., SLACK J.R. and WALLIS J.R. (1975). Regional skew in search for a parent. *Water Resources Research*, 11(6) : 815-826.
- [15] FISHER R.A. and TIPPETT L.H.C. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Proceedings Cambridge Philosophical Society*, 24 : 180-190.
- [16] JENKINSON A.F. (1955). The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Quarterly Journal of the Royal Meteorological Society*, 81(347) : 158-171.
- [17] GREENWOOD J.A., LANDWEHR J.M., MATALAS N.C. and WALLIS J.R. (1979). Probability weighted moments : Definition and relation to parameters of several distributions expressible in inverse form. *Water Resources Research*, 15(5) : 1049-1054.
- [18] HOSKING J.R.M. (1990). *L*-moments : Analysis and estimation of distributions using linear combinations of order statistics. *J. R. Stat. Soc.*, ser. B, 52 : 105-124.
- [19] VOGEL R.M. and FENNESSEY N.M.(1993). *L*-moment diagrams should replace moment diagrams. *Water Resources Research*, 29(6) : 1745-1752.
- [20] BOBÉE B. and RASMUSSEN P.F.(1995). *Invited paper* : Recent advances in flood frequency analysis. US National Report contributions in hydrology to International Union of Geodesy and Geophysics 1991-1994. *Review of Geophysics* : 1111-1116.
- [21] ROSSI F. FIORENTINO M. and VERSACE P. (1984). Two-component extreme value distribution for flood frequency analysis. *Water Resources Research*, 16(3) : 481-490.
- [22] ARNELL N.W. and GABRIELE S. (1988). The performance of the two-component extreme value distribution in regional flood frequency analysis. *Water Resources Research*, 24(6) : 879-887.
- [23] LANDWEHR J.M., MATALAS N.C. and WALLIS R. (1975). Some comparisons of flood statistics in real and Log space. *Water Resources Research*, 14(5) : 902-920.
- [24] KLEMES V. (1975). Comment on : Regional skew in search of a parent. *Water Resources Research*, 12(6) : 1325-1326.
- [25] ASHKAR F. and BOBÉE B.(1989). A discussion of statistical distributions and fitting techniques used in flood frequency analysis. *Ninth Canadian Hydrotechnical Conference*, St. John, Newfoundland.
- [26] ASHKAR F., BOBÉE B. and BERNIER J. (1992). Separation of skewness : Reality or regional artifact. *J. Hydraul. Eng. ASCE*, 118(3) : 460-475.
- [27] HALPHEN E. (1955). Les fonctions factorielles. *Institut de Statistique de l'Université de Paris*, IV(1) : 21-39.
- [28] PERREAULT L., BOBÉE B. and RASMUSSEN P.F.(1999a). The Halphen distribution system. Part 1 : Mathematical and statistical properties. *J. Hydrol. Engineering ASCE*, (accepté pour publication)
- [29] BOBÉE B., PERREAULT L. and ASHKAR F.(1993). Two kinds of moment ratio diagrams and their applications in hydrology. *Stochastic Hydrology and Hydraulics*, 7 : 41-65.
- [30] PERREAULT L., BOBÉE B. and RASMUSSEN P.F.(1999b). The Halphen distribution system. Part 2 :Parameter and quantile estimation. *J. Hydrol. Engineering ASCE* (accepté pour publication)