

**GUIDE PRATIQUE POUR L'EMPLOI
DES L-MOMENTS EN HYDROLOGIE**

Rapport d'étape présenté dans le cadre du projet Ouranos

Rapport de recherche R-725

Mars 2004



GUIDE PRATIQUE POUR L'EMPLOI DES L-MOMENTS EN HYDROLOGIE

Rapport d'étape présenté dans le cadre du projet Ouranos

Par

Salaheddine El-Adlouni

André St-Hilaire

Bernard Bobée

Taha B.M.J. Ouarda

**Chaire industrielle Hydro-Québec / CRSNG en Hydrologie statistique
Institut national de la recherche scientifique (INRS)
Centre Eau, Terre et Environnement
2800 rue Einstein CP 7500 Sainte-Foy (Québec) G1V 4C7**

Rapport de recherche R-725

Mars 2004

Référence

El-Adlouni S., St-Hilaire A., Bobée B. et T.B.M.J. Ouarda (2004). Guide pratique pour l'emploi des L-moments en hydrologie. Rapport présenté dans le cadre du projet Ouranos. INRS-ETE, rapport de recherche R-725, 24 pages.

TABLE DES MATIÈRES

1.0	PROBLÉMATIQUE.....	1
2.0	OBJECTIFS GÉNÉRAUX DU PROJET	2
3.0	OBJECTIFS DU RAPPORT D'ÉTAPE	3
4.0	GUIDE PRATIQUE SUR LES L-MOMENTS	4
4.1	INTRODUCTION	4
4.2	THEORIE DES L-MOMENTS	5
4.3	ESTIMATEURS DES L-MOMENTS D'UN ECHANTILLON.....	6
4.4	ESTIMATION DES PARAMETRES PAR LES L-MOMENTS	8
4.5	UTILISATION DES L-MOMENTS	10
4.6	INTERVALLES DE CONFIANCE DES L-MOMENTS	11
4.6.1	<i>Méthode d'intervalle standard</i>	12
4.6.2	<i>Méthodes des premier et deuxième percentiles</i>	13
4.6.2.1	<i>Méthode du premier percentile</i>	13
4.6.2.2	<i>Méthode du deuxième percentile</i>	13
4.6.3	<i>Intervalles de confiance des rapports des L-moments</i>	14
4.7	EXEMPLE ETUDIE	17
5.0	CONCLUSION.....	21
6.0	REFERENCES.....	22

LISTE DES FIGURES

FIGURE 1 : COMPARAISON DE DEUX AJUSTEMENTS D'UNE SERIE DES DEBITS MAXIMUMS ANNUELS PAR DEUX DISTRIBUTION : LA LOI DES VALEURS EXTREMES GENERALISEE (GEV) ET LA LOI DES FUTTES.	4
FIGURE 2 : PRESENTATION DES COURBES THEORIQUES DE QUELQUES DISTRIBUTIONS (NORMALE, GUMBEL, LOG-NORMALE A 2 ET A 3 PARAMETRES, LA LOI GEV, PEARSON TYPE III, PARETO GENERALISEE) SUR LE DIAGRAMME DES L-MOMENTS.	9
FIGURE 3 : ILLUSTRATION DE LA METHODE DU PREMIER PERCENTILE POUR LA CONSTRUCTION DES INTERVALLES DE CONFIANCE PAR BOOTSTRAP.	15
FIGURE 4 : VARIANCE DE L-Cs DE LA LOI GEV, LN3 ET PIII POUR UN ECHANTILLON DE TAILLE $n=50$ ($N=1000$).....	16
FIGURE 5 : VARIANCE DE L-Ck DE LA LOI LN3 EN FONCTION DE LA TAILLE DE L'ECHANTILLON POUR DIFFERENTES VALEURS DE L-Cs (SIMULATION FAITE POUR $N=10000$).....	16
FIGURE 6 : SERIE DES DEBIT MAXIMUMS ANNUELS	17
FIGURE 7 : OBSERVATIONS ORDONNEES DANS UN ORDRE CROISSANT.....	18
FIGURE 8 : POINT REPRESENTANT DE LA SERIE DES DONNEES SUR LE DIAGRAMME DES L-MOMENTS AVEC 100 ECHANTILLONS OBTENUS PAR BOOTSTRAP.....	19

1.0 PROBLÉMATIQUE

Dans le cadre du projet «Développement d'outils et analyses statistiques des occurrences/réurrences hydroclimatiques dans un contexte non stationnaire», un axe de recherche vise à approfondir des concepts et méthodes statistiques afin d'être en mesure d'effectuer des analyses plus précises de l'occurrence et de la récurrence des variables.

L'analyse fréquentielle ayant été identifiée comme étant un des principaux outils de caractérisation statistique pour les travaux d'analyse des extrêmes, il est important de continuer les travaux de recherche dans ce domaine. C'est dans cette optique qu'un projet intitulé «développement d'outils pour le choix de lois statistiques représentatives des variables climatologiques et hydrologiques» a été proposé.

Le but de l'analyse fréquentielle est de trouver x_T , la variable ou quantile de période de retour T, de probabilité au non-dépassement p, où $T = \frac{1}{1-p}$. On utilise des observations d'événements extrêmes passés afin d'estimer les probabilités futures d'occurrence. On doit, au préalable, vérifier l'hypothèse d'indépendance et d'identique distribution (i.i.d.) des observations. On cherche ensuite à estimer les quantiles x_T de période de retour T ou x_p , de probabilité au non-dépassement tel que : $prob\{X \leq x_T\} = 1 - \frac{1}{T}$.

Pour ce faire, on sélectionne une loi statistique à partir de plusieurs modèles disponibles. La loi sélectionnée doit ensuite être ajustée aux séries de mesures, et l'estimation du quantile x_T par une estimation ponctuelle \hat{x}_T est alors donnée à partir de la loi ajustée.

2.0 OBJECTIFS GÉNÉRAUX DU PROJET

Ce projet vise à élaborer, à partir des travaux réalisés par l'équipe de la Chaire en Hydrologie statistique de l'INRS-ETE, des méthodes de sélection et d'ajustement de lois statistiques applicables dans un contexte d'analyse d'occurrence et de récurrence d'événements associés aux changements climatiques. Les travaux envisagés vont porter principalement sur les méthodes de sélection des lois statistiques ajustées aux séries d'observations de variables hydroclimatiques. Les forces et les faiblesses de certaines de ces méthodes ont déjà été soulignées (*e.g.* Klemesš, 2000; Bernier, 1993 communication personnelle), mais une analyse critique et quantitative plus complète est nécessaire dans un cadre pratique lié aux changements climatiques. Nous proposons donc d'entreprendre cette analyse.

3.0 OBJECTIFS DU RAPPORT D'ÉTAPE

Dans le premier rapport (El Adlouni *et al.*, 2003), nous avons présenté en détail trois diagrammes pour représenter les distributions utilisées en hydrologie : le diagramme des moments, le diagramme (δ_1, δ_2) et le diagramme des L-moments. Ces trois diagrammes peuvent être utilisés comme des critères graphiques pour le choix de la distribution qui s'ajuste le mieux avec les données. Ceci est possible en comparant le point représentant l'échantillon aux courbes ou surfaces représentatives des distributions théoriques dans chacun de ces diagrammes. Au cours de la dernière décennie, l'utilisation des L-moments a connu un grand essor, tant au niveau du choix d'une distribution que de l'estimation des paramètres de la distribution sélectionnée. L'objectif de ce rapport d'étape est de fournir un guide pratique pour l'utilisation des L-moments. Après avoir rappelé la théorie des L-moments, nous présentons les différentes étapes d'utilisation des L-moments soit pour l'estimation des paramètres ou bien pour le choix du meilleur ajustement. La dernière partie de ce rapport sera consacrée à la construction des intervalles de confiance pour les L-moments par Bootstrap paramétrique.

4.0 GUIDE PRATIQUE SUR LES L-MOMENTS

4.1 INTRODUCTION

Dans le présent rapport, les travaux portent sur l'emploi des L-moments pour l'estimation des paramètres et aussi comme outil de choix du meilleur ajustement. La figure 1 présente un exemple d'ajustement d'une série des débits maximums annuels. Deux distributions ont été ajustées aux observations : la loi des valeurs extrêmes généralisée (GEV) et la loi des fuites. On remarque que plus la période de retour augmente, plus la différence entre les quantiles donnés par chacun des deux modèles devient significative. Ainsi, la période de retour 100 ans est associée à un débit maximum annuel de 500 m³/s dans le cas de l'ajustement avec la loi des fuites, alors que le quantile associé à cette même période de retour est de 600 m³/s avec la GEV (figure 1). Cette différence est encore plus importante pour la période de retour de 1000 ans (voir figure 1). Le choix du meilleur modèle dépend de (1) la distribution utilisée pour l'ajustement et de (2) la méthode d'estimation des paramètres. L'approche par les L-moments peut être utilisée pour effectuer ces deux étapes importantes.

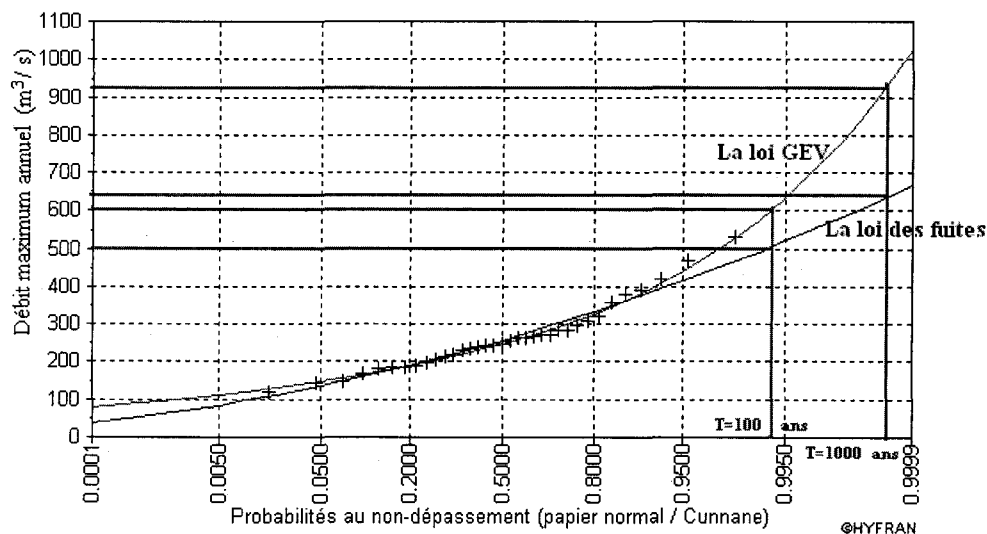


Figure 1 : Comparaison de deux ajustements d'une série des débits maximums annuels par deux distributions : la loi des valeurs extrêmes généralisée (GEV) et la loi des fuites.

Tout d'abord nous rappelons la théorie de base des L-moments. Ensuite, nous présentons un guide pratique pour l'estimation des paramètres par les L-moments des distributions utilisées pour l'ajuster un échantillon et quelques procédures basées sur le Bootstrap pour la construction d'intervalles de confiance pour les L-moments ce qui permet d'obtenir une incertitude sur l'estimation des débits de crue. Nous terminons par une conclusion générale et des travaux futurs.

4.2 THÉORIE DES L-MOMENTS

Soit une variable aléatoire X de fonction de distribution F , et soit $X_{1:n} \leq X_{2:n} \leq \dots \leq X_{n:n}$ les statistiques d'ordre $1, 2, \dots, n$ pour un échantillon de taille n . Les L-moments de X sont définis par :

$$\lambda_r = r^{-1} \sum_{k=0}^{r-1} (-1)^k \binom{r-1}{k} E(X_{r-k:r}) \quad r=1, 2, \dots \quad (1)$$

où $E(X_{j:r}) = \frac{r!}{(j-1)!(r-j)!} \int x \{F(x)\}^{j-1} \{1-F(x)\}^{r-j} dF(x)$.

$E(X)$ est la moyenne de la variable aléatoire X .

Les premiers L-moments sont donnés par :

$$\lambda_1 = E[X] = \int x dF \quad (2)$$

$$\lambda_2 = \frac{1}{2} E[X_{2:2} - X_{1:2}] = \int x(2F-1) dF \quad (3)$$

$$\lambda_3 = \frac{1}{3} E[X_{3:3} - 2X_{2:3} + X_{1:3}] = \int x(6F^2 - 6F + 1) dF \quad (4)$$

$$\lambda_4 = \frac{1}{4} E[X_{4:4} - 3X_{3:4} + 3X_{2:4} - X_{1:4}] = \int x(20F^3 - 30F^2 + 12F - 1) dF \quad (5)$$

λ_2 est une mesure de dispersion de la variable X . Parmi les avantages des L-moments, on note qu'ils existent si et seulement si $E[X] < \infty$ et la distribution est bien caractérisée par ses L-moments. Ce qui n'est pas le cas pour les moments ordinaires. Comme par exemple la loi de

Student (t) et la loi de Cauchy pour lesquelles les moments (classiques) d'ordre supérieur à 2 n'existent pas. Et donc, on ne peut pas les comparer à d'autres distributions.

Il est toujours intéressant de standardiser les moments d'ordre supérieurs $\lambda_r, r \geq 3$, pour qu'ils soient indépendants de l'unité de mesure de X . On définit alors les rapports des L-moments (Hosking, 1990) :

$$\tau_r = \lambda_r / \lambda_2, \quad r = 3, 4, \dots \quad (6)$$

Une propriété intéressante des rapport des L-moments est qu'ils sont bornés ($|\tau_r| < 1$), τ_3 est une mesure d'asymétrie (L-asymétrie) et τ_4 est une mesure d'aplatissement (L-aplatissement).

4.3 ESTIMATEURS DES L-MOMENTS D'UN ÉCHANTILLON

Pour un échantillon ordonné dans un ordre croissant, les estimateurs des L-moments peuvent être déduits de ceux des moments de probabilité pondérés (Greenwood *et al.*, 1979).

Les moments de probabilité pondérés sont définis par :

$$\beta_r = E \left\{ X [F(x)]^r \right\} \quad (7)$$

un simple estimateur de β_r pour $r \geq 1$ est :

$$b_r^* = \frac{1}{n} \sum_{j=1}^n X_{j:n} \left[1 - \frac{(j-0.35)}{n} \right]^r \quad (8)$$

où $1-(j-0.35)/n$ sont des estimateurs de $F(X_{(j)})$ (Landwehr *et al.*, 1979). b_r^* est un estimateur biaisé mais l'erreur moyenne quadratique des estimateurs des quantiles est faible par rapport à l'autre estimateur sans biais donné ci-dessous :

$$b_0 = n^{-1} \sum_{j=1}^n X_j \quad (9)$$

$$b_r = n^{-1} \sum_{j=r+1}^n \frac{(j-1)(j-2)\dots(j-r)}{(n-1)(n-2)\dots(n-r)} X_j \quad (10)$$

On en déduit donc les estimateurs des quatre premiers L-moments (Hosking, 1990):

$$l_1 = b_0, \quad (11)$$

$$l_2 = 2b_1 - b_0, \quad (12)$$

$$l_3 = 6b_2 - 6b_1 + b_0 \quad (13)$$

$$l_4 = 20b_3 - 30b_2 + 12b_1 - b_0. \quad (14)$$

En divisant les L-moments d'ordre supérieur à 2 par la mesure de dispersion l_2 , on obtient les

estimateurs des rapports des L-moments $t_r = \frac{l_r}{l_2}$ $r = 3, 4, \dots$.

En pratique, le coefficient du L-asymétrie (t3) et du L-aplatissement (t4) peuvent s'écrire d'une manière directe de la façon suivante :

$$t_3 = \sum_{i=1}^n \frac{c_i x_{(i)} + \frac{\bar{x}}{n}}{l_2} \quad (15)$$

où $c_i = 6 \frac{(i-1)(i-2)}{n(n-1)(n-2)} - 6 \frac{(i-1)}{n(n-1)}$

et

$$t_4 = \sum_{i=1}^n \frac{d_i x_{(i)} - \frac{\bar{x}}{n}}{l_2} \quad (16)$$

où $d_i = 20 \frac{(i-1)(i-2)(i-3)}{n(n-1)(n-2)(n-3)} - 30 \frac{(i-1)(i-2)}{n(n-1)(n-2)} + 12 \frac{(i-1)}{n(n-1)}$

et \bar{x} étant la moyenne de l'échantillon.

Hosking et Wallis (1995) ont suggéré l'utilisation des estimateurs sans biais. L'utilisation des estimateurs basée sur la fonction de répartition empirique reste utile surtout dans le cas de l'estimation régionale des quantiles de période de retour élevée.

4.4 ESTIMATION DES PARAMÈTRES PAR LES L-MOMENTS

L'estimation des paramètres par la méthode des L-moments est basée sur le même principe que celui des moments ordinaires. Les estimateurs sont la solution d'un système d'équations correspondant à l'égalité entre les L-moments de l'échantillon et ceux de la distribution théorique à ajuster à l'échantillon. Pour les distributions à trois paramètres, le système correspondant à l'estimation des paramètres, comporte trois équations indépendantes, généralement on utilise l'égalité entre les L-moments théoriques d'ordre 1 et 2 avec ceux de l'échantillon. La troisième équation est basée sur l'égalité entre le coefficient du L-asymétrie de l'échantillon et celui de la distribution. Dans le diagramme ($L-C_s$; $L-C_k$), ceci équivaut à considérer une projection en parallèle avec l'axe des ordonnées du point représentant l'échantillon sur la courbe de la distribution à ajuster (figure 2).

Diagramme des L-Moments

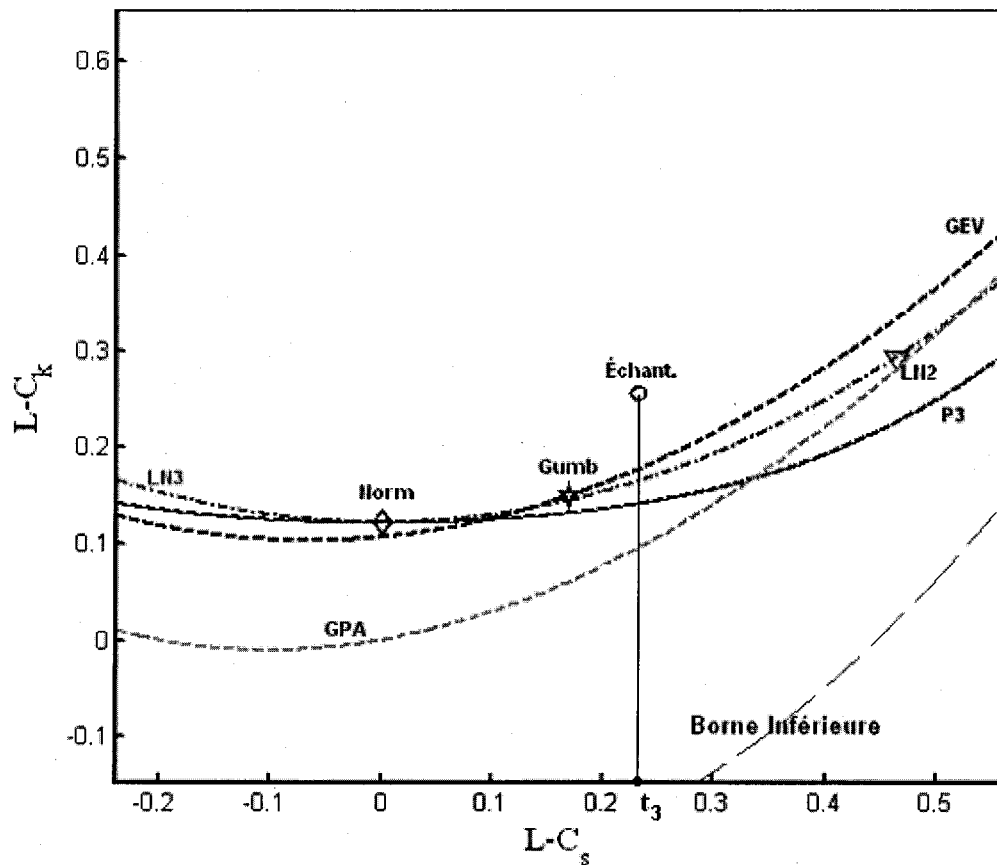


Figure 2 : Présentation des courbes théoriques de quelques distributions (Normale, Gumbel, Log-Normale à 2 et à 3 paramètres, la loi GEV, Pearson type III, Pareto généralisée) sur le Diagramme des L-moments.

Les étapes de l'estimation des paramètres par la méthode des L-moments peuvent se résumer de la façon suivante :

Algorithme d'estimation

1. Ordonner l'échantillon d'une manière croissante $x_{1:n} \leq \dots \leq x_{n:n}$
2. Calcul des estimateurs des moments de probabilité pondérés b_0, b_1, b_2 et b_3 (équations 9 et 10)
3. Calcul des trois premiers L-moments l_1, l_2 et l_3

4. En déduire le coefficient de L-asymétrie t_3 (Pour les distributions à trois paramètres, on n'utilise que le coefficient de L-asymétrie, [équation 15])
5. Pour chaque distribution, on utilise l'expression des paramètres en fonction des L-moments.

La résolution de ce système n'est pas explicite pour la majorité des distributions considérées dans ce diagramme.

Par exemple, pour une loi GEV (*Generalized Extreme Value*) de fonction densité de probabilité :

$$F(x) = \exp \left[- \left(1 - k \frac{x - \xi}{\alpha} \right)^{1/k} \right] \quad (15)$$

où $k \neq 0$ est un paramètre de forme α un paramètre d'échelle et ξ un paramètre de position.

Les estimateurs des paramètres par L-moments de la loi GEV sont données par :

$$\hat{k} \approx 7.8590 c + 2.9554 c^2 \quad \text{où} \quad c = \frac{2}{3+t_3} - \frac{\log 2}{\log 3}$$

$$\hat{\alpha} = \frac{l_2 \hat{k}}{(1-2^{-\hat{k}}) \Gamma(1+\hat{k})} \quad \text{et} \quad \hat{\xi} = l_1 - \hat{\alpha} \left\{ 1 - \Gamma(1+\hat{k}) \right\} / \hat{k}$$

Pour les distributions pour lesquelles on ne peut pas exprimer explicitement les paramètres en fonction des L-moments, des approximations polynomiales sont données dans Hosking et Wallis (1997).

4.5 UTILISATION DES L-MOMENTS

Les L-moments d'une variable aléatoire réelle X, existent si et seulement si X a une espérance finie. La distribution peut être caractérisée par les L-moments même si certains des

moments ordinaires n'existent pas. Les avantages, des L-moments par rapport aux moments classiques sont : 1) leur capacité de caractériser plusieurs distributions; 2) ils sont plus robustes à la présence des horsains (valeurs singulières dans l'échantillon) et 3) leurs estimateurs sont moins biaisés et sont proches de leurs approximations asymptotiques par une loi normale. Un exemple souple des L-moments est de permettre la caractérisation et l'ajustement de certaines distributions qui ne sont pas caractérisées par les moments classiques telles que la loi de Student et la loi de Cauchy. Concernant le deuxième avantage cité ci-dessus, Bobée et Rasmussen (1995) expliquent que, bien que cette faible sensibilité peut sembler souhaitable, il se peut que ces estimateurs soient trop robustes. Si on accepte la prémisse que les valeurs des échantillons donnant de l'information sur la queue de la distribution sont légitimes, c'est-à-dire qu'elles ne sont pas des horsains, il est alors préférable de choisir une méthode d'estimation qui soit suffisamment sensible pour refléter cette information.

Actuellement les L-moments ont remplacé les moments classiques pour caractériser les distributions dans plusieurs domaines : l'analyse des ressources hydriques, les études climatiques, l'astronomie et l'hydrologie. Même s'il n'y a pas de test d'adéquation développé en relation avec les L-moments, la distance entre l'échantillon et les courbes des différentes distributions peut être considérée comme une mesure de la qualité de l'ajustement (Onoz et Bayazit, 1995).

4.6 INTERVALLES DE CONFIANCE DES L-MOMENTS

Les L-moments sont distribués asymptotiquement suivant une loi normale multivariée. Cependant, la matrice de variance/covariance dépend de la distribution de la variable aléatoire. Nous avons donc besoin de connaître la loi de la variable étudiée pour construire les intervalles de confiance. Puisque les L-moments sont distribués suivant une loi normale quand n tend vers l'infini, nous pouvons utiliser l'expression des intervalles de confiance correspondant à une loi

normale. La construction des intervalles de confiance par Bootstrap donne de bons résultats surtout pour les petits échantillons. Dans cette partie, nous présentons trois méthodes basées sur le Bootstrap pour déterminer des intervalles de confiance pour les L-moments. La méthode du Bootstrap, développée par Efron en 1979, est une méthode de ré-échantillonnage qui simule N sous-échantillons de taille n à partir de la distribution ajusté à l'échantillon observé, afin d'en tirer toutes les propriétés de la statistique étudiée.

Les étapes de la méthode du Bootstrap peuvent être résumées de la façon suivante :

1. Ajustement d'une distribution F (de paramètre θ) à l'échantillon étudié (de taille n), on obtient donc un estimateur $\hat{\theta}$ de θ ;
2. Simulation de N échantillons de taille n à partir de $F(\hat{\theta})$;
3. Ajustement de la distribution F aux N échantillons. On obtient donc $\hat{\theta}_1, \dots, \hat{\theta}_N$ estimateur de θ ;
4. À partir de cette série $\hat{\theta}_1, \dots, \hat{\theta}_N$ d'estimateurs, on peut tirer toutes les propriétés statistiques sur le paramètre θ .

Exemple la variance par Bootstrap de $\hat{\theta}$ est :

$$\sigma_{\hat{\theta},B}^2 = \frac{1}{N-1} \sum_{i=1}^N (\hat{\theta}_i - \bar{\hat{\theta}})^2 \quad \text{où } \bar{\hat{\theta}} = \frac{1}{N} \sum_{i=1}^N \hat{\theta}_i \text{ est la moyenne de } \hat{\theta}_1, \dots, \hat{\theta}_N.$$

On distingue trois méthodes pour la construction des intervalles de confiance par Bootstrap : la méthode d'intervalle standard, la méthode du premier percentile et la méthode du deuxième percentile.

4.6.1 Méthode d'intervalle standard

Pour un paramètre d'intérêt θ , la méthode d'intervalle standard, qui suppose que la loi de l'estimateur $\hat{\theta}$ est asymptotiquement normale permet de construire des intervalles de confiance en utilisant l'estimateur de l'écart type par Bootstrap de $\hat{\theta}$. L'écart type par Bootstrap est l'écart

type des N valeurs de la statistique $\hat{\theta}$ obtenues pour tous les sous-échantillons. Sous l'hypothèse de normalité, un intervalle de confiance de θ peut s'écrire de la façon suivante $\hat{\theta} - z_{\alpha/2} \hat{\sigma}_{\hat{\theta},B}, \hat{\theta} + z_{\alpha/2} \hat{\sigma}_{\hat{\theta},B}$, où $\hat{\sigma}_{\hat{\theta},B}$ est l'estimateur de l'écart type par Bootstrap. Les conditions nécessaires pour que cette méthode donne de bons résultats sont les suivantes :

(1) $\hat{\theta}$ doit être distribué approximativement suivant une loi normale, (2) $\hat{\theta}$ doit être un estimateur non biaisé de θ afin d'avoir des résultats fiables pour la valeur moyenne des sous-échantillons simulés et (3) le ré-échantillonnage par Bootstrap doit donner une bonne approximation de $\sigma_{\hat{\theta},B}$. La dernière condition dépend surtout du nombre de sous-échantillons simulés. La méthode d'intervalle standard donne de bons résultats juste pour $N=100$. D'autres méthodes demandent plus de simulations.

4.6.2 Méthodes des premier et deuxième percentiles

Les méthodes du premier et deuxième percentiles sont toutes les deux basées sur la construction des intervalles de confiance par Bootstrap paramétrique, qui utilisent les percentiles de la distribution « Bootstrapée » afin d'avoir une approximation des percentiles de la distribution d'un estimateur. Contrairement à la méthode des intervalles standard, ces deux méthodes ne font aucune hypothèse sur la distribution de l'estimateur. L'idée de base des deux méthodes est presque la même, sauf que la première utilise les percentiles obtenus pour les estimateurs, alors que la seconde considère les erreurs d'estimation.

4.6.2.1 Méthode du premier percentile

Soit $\hat{\theta}_1, \dots, \hat{\theta}_N$ les estimateurs de θ correspondants aux sous-échantillons simulés par Bootstrap. Après avoir ordonné ces estimateurs, la méthode du premier percentile considère les deux valeurs contenant la moyenne avec une probabilité de $100(1-\alpha)\%$.

4.6.2.2 Méthode du deuxième percentile

La méthode du deuxième percentile considère le même ensemble d'estimateurs ordonné, et calcul l'erreur entre les estimateurs par Bootstrap et l'estimateur donné par l'échantillon de base $\hat{\theta}$. On obtient donc N valeurs pour l'erreur de l'estimation, $\{\varepsilon_i = \hat{\theta}_i - \hat{\theta}; i = 1, \dots, N\}$ qui seront

utilisées comme approximation de l'erreur de la distribution de $\hat{\theta}$. À partir des N réalisations de l'erreur, on détermine les deux quantiles ε_l et ε_s correspondants aux probabilités au non-dépassement $\alpha/2$ et $(1-\alpha/2)$, respectivement. L'intervalle de confiance de θ par la méthode du deuxième percentile est $\hat{\theta} - \varepsilon_l; \hat{\theta} - \varepsilon_s$ ce qui est équivalent à $Prob(2\hat{\theta} - \hat{\theta}_l < \theta < 2\hat{\theta} - \hat{\theta}_s) = 1 - \alpha$.

4.6.3 Intervalles de confiance des rapports des L-moments

Dans le présent rapport nous présentons les résultats de la construction des intervalles de confiance pour les coefficients du L-asymétrie et L-aplatissement, par Bootstrap en utilisant la méthode de l'intervalle standard (cf. 4.6.1).

Pour illustrer l'idée de base de la construction des intervalles de confiance pas Bootstrap, la figure 3 présente les estimateurs de L-Cs et de L-Ck pour un échantillon donné (Échant.), les autres points correspondent à 50 sous-échantillons issus de la même population que l'échantillon étudié. À partir des valeurs du L-Cs et L-Ck des sous-échantillons, on présente les intervalles de confiances des deux rapports des L-moments par la méthode du premier percentile. Ceci correspond sur le graphique aux deux intervalles qui encadrent les valeurs de L-Cs et L-Ck de l'échantillon étudié.

Diagramme des L-Moments

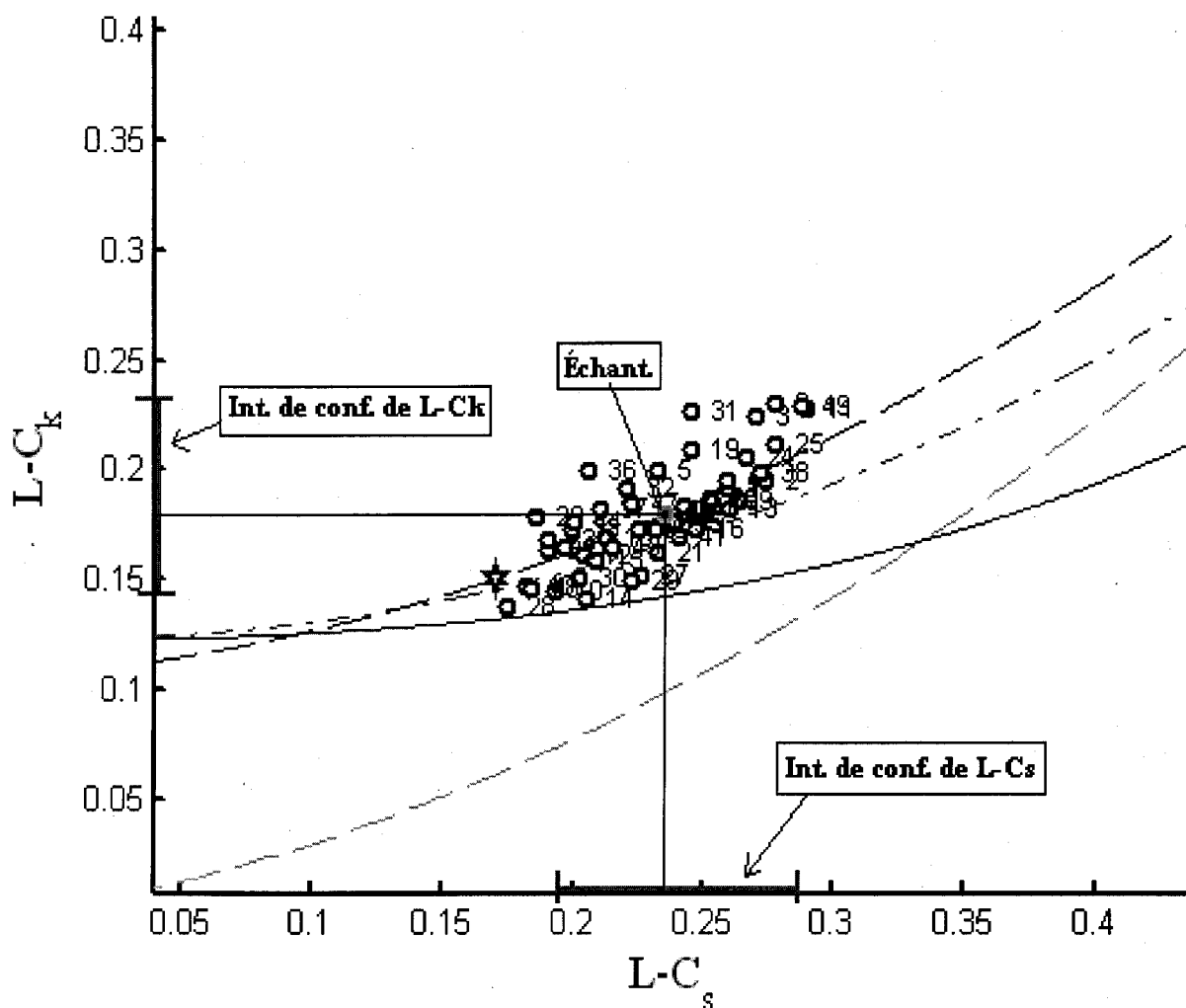


Figure 3 : Illustration de la méthode du premier percentile pour la construction des intervalles de confiance par Bootstrap.

La figure 4 représente la variance de $L-C_s$ (t3) pour les distributions GEV, Log-normale à 3 paramètres (LN3) et Pearson type 3 (PIII) correspondante à un échantillon de taille $n=50$ et différentes valeurs du L-asymétrie τ_3 . Le nombre de simulation considéré est $N=1000$. On remarque que la loi du coefficient du L-asymétrie dépend de la distribution de la variable étudiée et de la valeur de l'asymétrie.

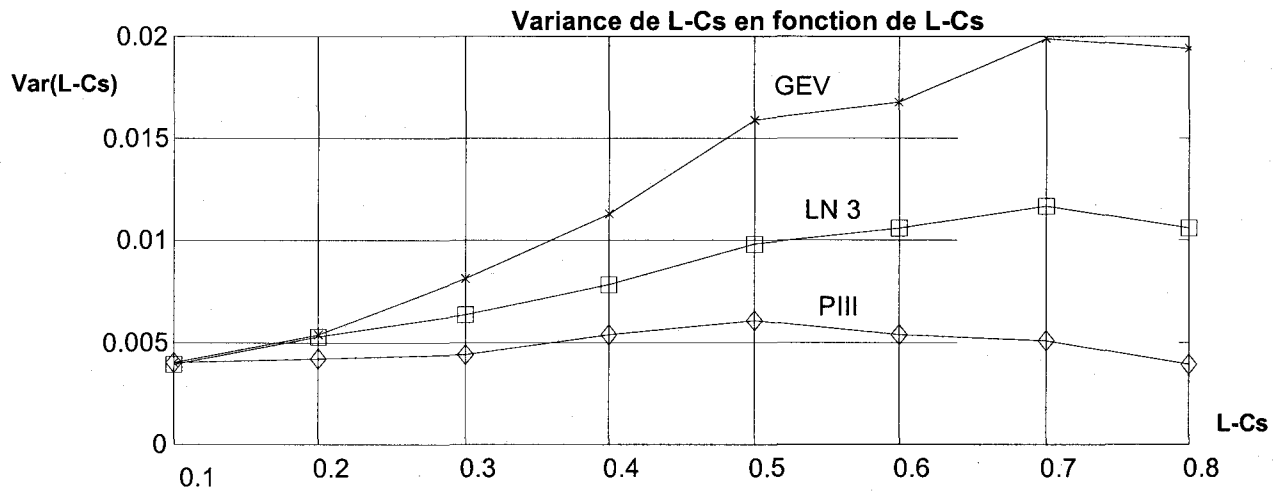


Figure 4 : Variance de L-Cs de la loi GEV, LN3 et PIII pour un échantillon de taille $n=50$ ($N=1000$)

La figure 5, présente la variance de $L-C_k$ (t_4) pour des échantillons issus d'une loi LN3, en fonction du coefficient L-Cs (t_3) et la taille de l'échantillon n . On constate que la distribution de l'estimateur de τ_4 dépend de la taille de l'échantillon et de la valeur de l'asymétrie. Ces résultats peuvent être utilisés pour déterminer des intervalles de confiance pour $L-C_s$, $L-C_k$ ou bien des estimateurs des quantiles obtenus par la méthode des L-moments.

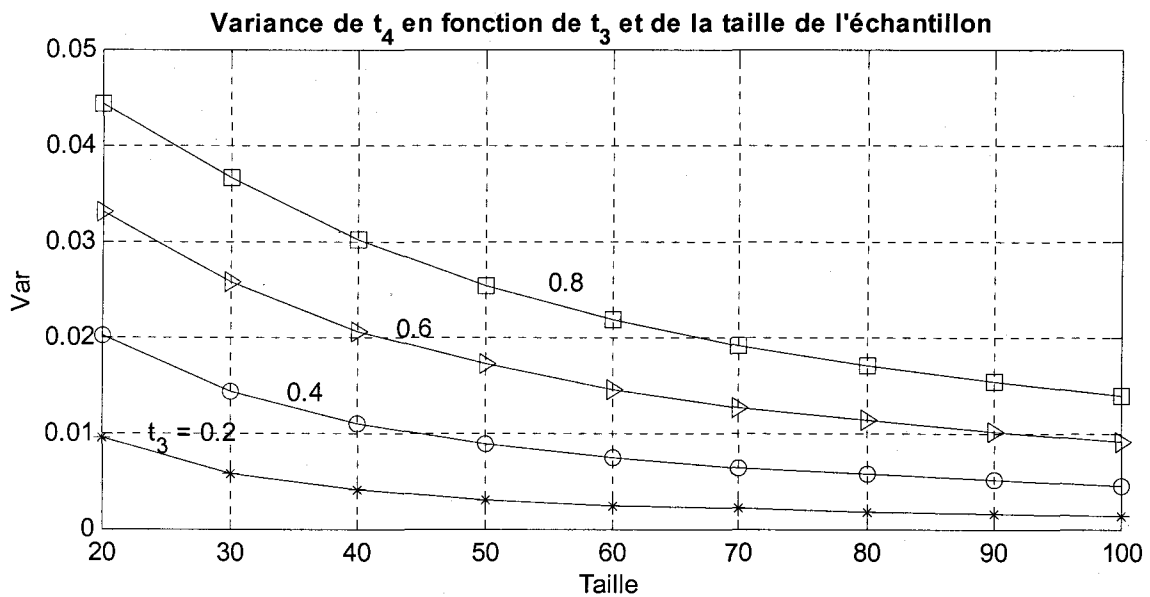


Figure 5 : Variance de $L-C_k$ de la loi LN3 en fonction de la taille de l'échantillon pour différentes valeurs de L-Cs (simulation faite pour $N=10000$)

4.7 EXEMPLE ÉTUDIÉ

Dans cette exemple nous présentons une étude d'une série de débit de taille $n = 30$ (figure6) afin d'illustrer les différentes méthodes décrites dans ce guide.

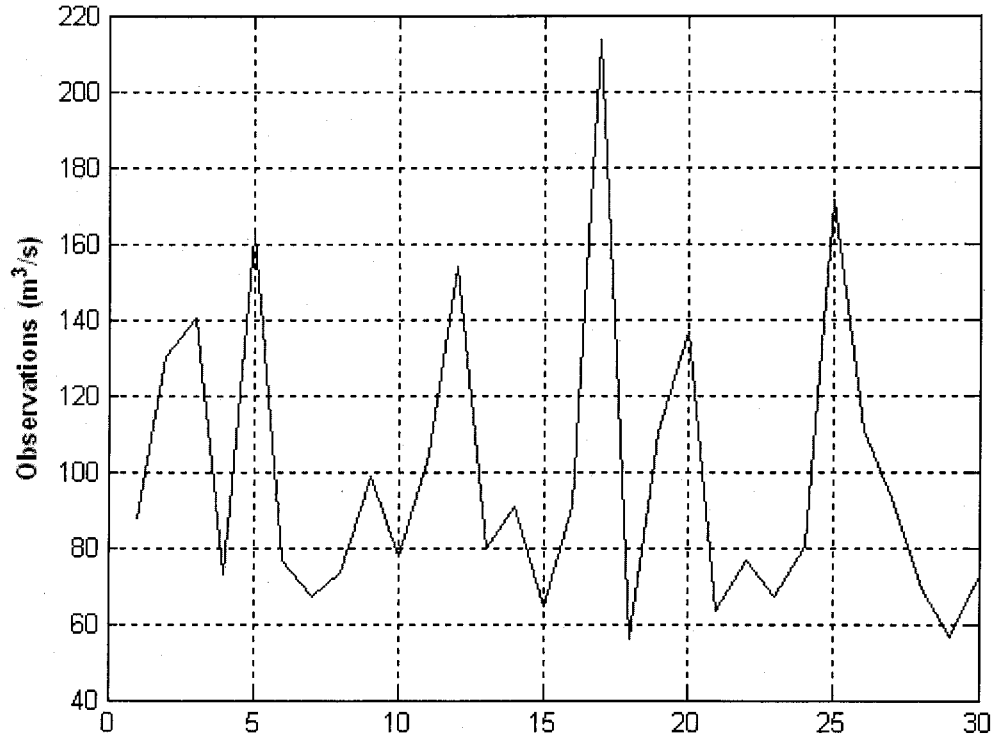


Figure 6 : Série des débit maximums annuels

Pour estimer les L-moments de cette série, on applique l'algorithme d'estimation (cf. 4.4). Après avoir ordonné la série de données dans un ordre croissant, $X_{1:n} \leq X_{2:n} \leq \dots \leq X_{n:n}$ (figure 7),

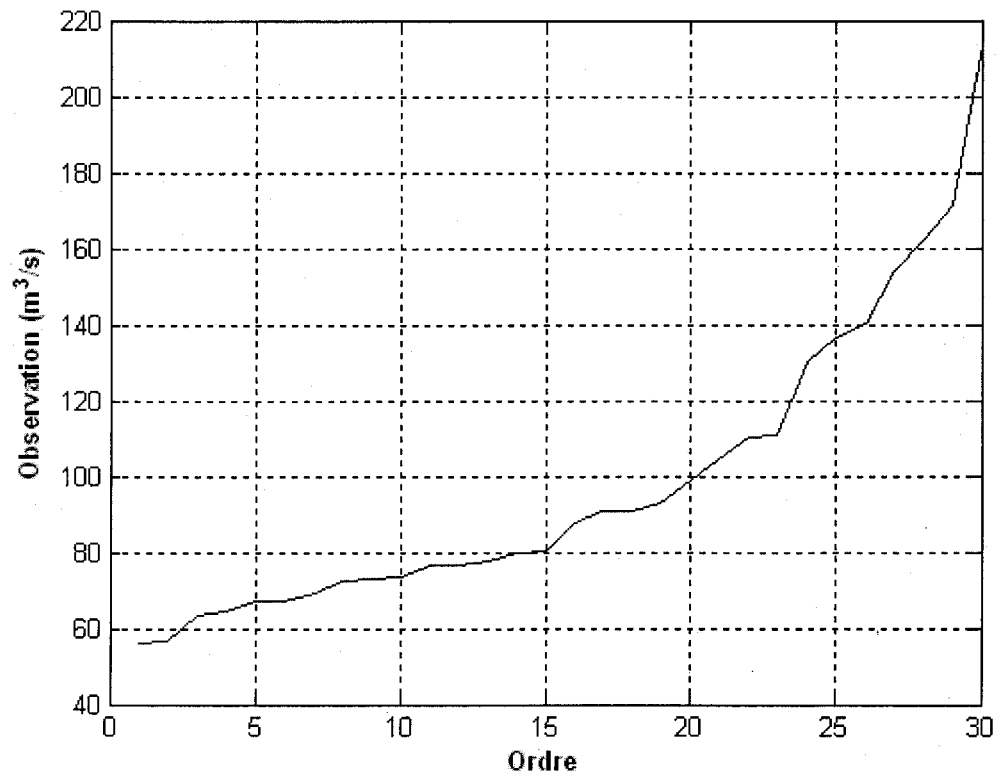


Figure 7 : Observations ordonnées dans un ordre croissant

on calcule les estimateurs des moments de probabilités pondérés (b_0, b_1, b_2 et b_3) donnés par les équations (9) et (10) :

$$b_0 = 98.2929 ; b_1 = 59.4412 ; b_2 = 44.1969 ; b_3 = 35.7001$$

Ensuite, on utilise les expressions (équations 11-14) des L-moments en fonction des b_i ($i = 0, 1, 2, 3$). On obtient les résultats suivants :

$$l_1 = 99.0418 ; l_2 = 23.0971 ; l_3 = 6.6035 ; l_4 = 4.1136$$

Les estimateurs des rapports des L-moments (t_3 et t_4) peuvent, donc, être calculés en utilisant les

formules suivantes : $t_3 = \frac{l_4}{l_2}$, $t_4 = \frac{l_3}{l_2}$ on obtient donc les valeurs suivantes :

$$t_3 = 0.2859 \text{ et } t_4 = 0.1781$$

Avec ces valeurs de (t_3 et t_4), la série peut être représentée dans le diagramme des L-moments (figure 8).

Pour estimer les paramètres d'une distribution à trois paramètres, on utilise un système de trois équations indépendantes souvent celles de l_1 , l_2 et l_3 . Par exemple, pour la loi GEV on peut calculer les estimateurs des paramètres en utilisant les formules données au paragraphe 4.4. On obtient alors les estimateurs suivants :

$$\hat{k} = -0.2378 \quad \hat{\alpha} = 22.5894 \quad \hat{\xi} = 78.3891$$

Diagramme des L-Moments

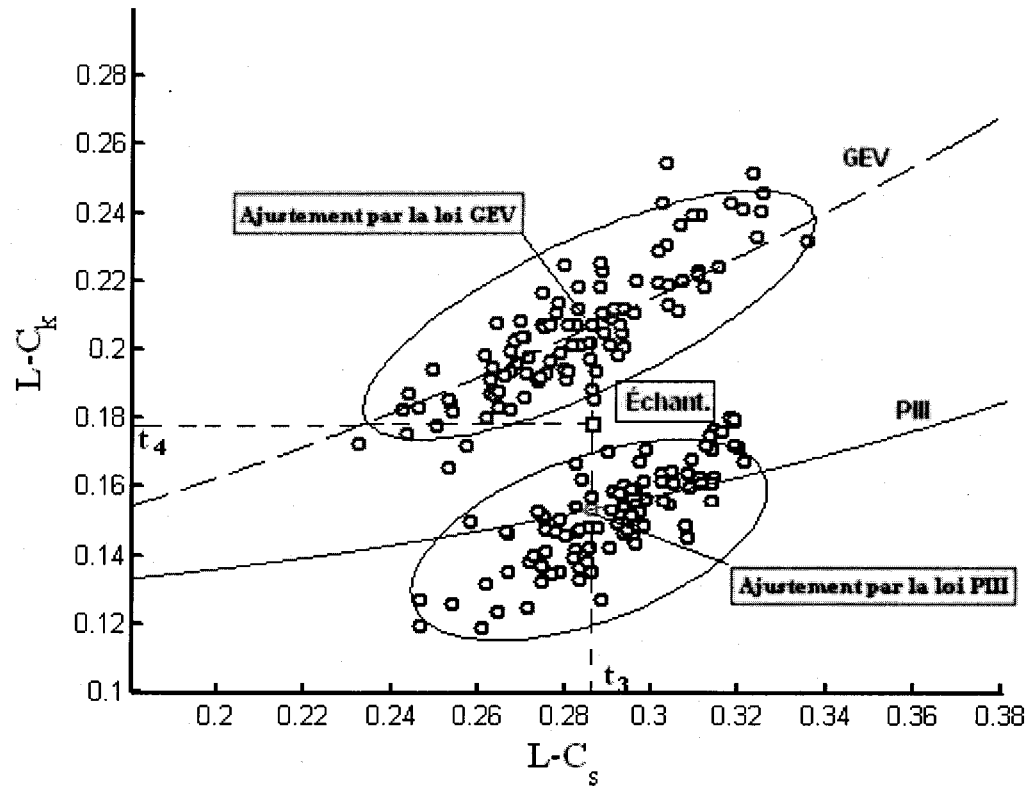


Figure 8 : Point représentant de la série des données sur le diagramme des L-moments avec 100 échantillons obtenus par Bootstrap

Comme nous l'avons mentionné au début de ce guide, l'approche par L-moments peut être utilisée comme test d'ajustement ou comme critère de sélection du meilleur ajustement. La figure 8, où l'on présente le point $(t_3 ; t_4)$ de la série étudiée dans le diagramme des L-moments ainsi que 100 échantillons par Bootstrap obtenus par la loi GEV et la loi Pearson type 3 (PIII), montre comment ce diagramme peut servir pour la comparaison de deux distributions. La méthode la plus simple est de considérer les distances entre le point représentant de la série et les courbes de chacune des distributions étudiées, ou bien à l'aide des intervalles de confiance obtenus par Bootstrap.

5.0 CONCLUSION

Dans ce rapport, nous avons présenté un guide pratique pour l'utilisation des L-moments. La théorie des L-moments peut être utilisée aussi bien comme méthode d'estimation des paramètres que pour la discrimination entre les distributions. Nous avons mis l'accent surtout sur l'utilisation des L-moments pour l'estimation des paramètres. La construction des intervalles de confiance joue un rôle très important dans toute analyse statistique. Les intervalles de confiance permettent de quantifier l'erreur admissible à un certain niveau de confiance $(1-\alpha)$. Les méthodes basées sur le Bootstrap permettent d'obtenir des intervalles de confiance plus fiables, surtout dans le cas des échantillons de petite taille. Il serait intéressant de comparer les différentes méthodes pour construire des intervalles de confiance, présentées dans ce rapport, à partir de données simulées.

6.0 RÉFÉRENCES

- Bernier J. (1993). Sur les utilisations des L-Moments en hydrologie statistique. Communication personnelle en vue d'une publication.
- Bobée B. et P. F. Rasmussen (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.
- Efron B. (1979). Bootstrap methods: another look at the jackknife. *Annals of Statistics*, 7: 1-26.
- El Adlouni S., St-Hilaire A., Bobée B. et T. B.M.J. Ouarda (2002). Développement d'outils pour le choix de lois statistiques représentatives des variables climatologiques et hydrologiques. Rapport d'étape présenté dans le cadre du projet Ouranos.
- Greenwood J.A., Landwehr J.M., Matalas N.C. and J. R. Wallis (1979). Probability-weighted moments : Definition and relation to parameters of several distributions expressible in inverse form. *Water Resources Research*, 15: 1049-1054.
- Hosking J.R.M. (1990). L-Moments: Analysis and estimation of distributions using linear combinations of order statistics. *Journal of Royal Statistical Society*, 52: 105-124.
- Hosking J.R.M. and J.R. Wallis (1995). A comparison of unbiased and plotting-position estimators of L Moments. *Water Resources Research*, 31: 2019-2025.
- Hosking J. R.M. and J.R. Wallis (1997). Regional frequency analysis: an approach based on L-moments. Cambridge University Press, Cambridge, UK.
- Klemeš V.(2000). Tall tales about tails of hydrological distributions / II. *Journal of Hydrologic Engineering*, 5(3): 232-239.
- Landwehr J.M., Matalas N.C. and J.R. Wallis (1979). Probability-weighted moments compared with some traditional techniques in estimating Gumbel parameters and quantiles. *Water Resources Research*, 15: 1055-1064.
- Onoz B. and M. Bayazit (1995). Best fit distributions of largest available flood samples. *Journal of Hydrology*, 167: 195-208.