

APPROCHE REGIONALE POUR L'ESTIMATION DES PRECIPITATIONS JOURNALIERES EXTREMES DU NORD EST ALGERIEN

TAMARA BENABDESSELAM, HOCINE AMARCHI

Département d'Hydraulique, Université Badji Mokhtar-Annaba, BP 12, 23000 Annaba, Algérie
kondratievat@yahoo.fr

RESUME

L'information sur les valeurs des précipitations extrêmes et leurs fréquences est nécessaire pour les différents projets liés à l'eau. Dans cette étude, l'analyse fréquentielle régionale basée sur les L-moments a été utilisée pour améliorer la qualité d'estimation des quantiles des précipitations journalières maximales annuelles de la région située au Nord Est de l'Algérie représentée par 50 stations de mesure. La région d'étude a été divisée en deux régions statistiquement homogènes en termes des rapports des L-moments. Parmi les différentes distributions testées, (i) la distribution *GEV* (Generalized Extreme Value) a été identifiée comme distribution régionale la plus adéquate pour la modélisation des précipitations de la région 1 caractérisée par un climat méditerranéen et (ii) la distribution lognormale à trois paramètres (*LN3*) pour la région 2, sujette à un climat semi-aride. Les courbes de croissance ont été établies pour chaque région homogène. Ainsi, les quantiles des précipitations dans un site sont déterminés en multipliant la moyenne des précipitations par les quantiles régionaux correspondants. La comparaison des quantiles estimés à partir de l'approche régionale et locale a montré que pour 50,4 % des stations de la région d'étude, le modèle local sous estime les quantiles de fortes périodes de retour.

MOTS CLÉS: Précipitations extrêmes journalières, moments linéaires, analyse fréquentielle régionale, distributions *GEV*, *LN3*, simulation des précipitations.

ABSTRACT

Information on precipitation magnitudes and their frequencies is needed in various projects related to water. In this study, the regional frequency analysis based on L-moments was used to improve the quality of estimating of the annual maximum daily precipitations of the north eastern area of Algeria represented by 50 measuring stations. The investigated area was divided into two regions statistically homogeneous in terms of L-moments ratios. Among the different tested distributions, (i) the Generalized Extreme Value (*GEV*) distribution has been identified as the most appropriate regional distribution for modeling precipitation in the northern region (region 1) characterized by a Mediterranean climate and (ii) the generalized lognormal distribution (*LN3*) for region 2 of southern part subject to semi-arid climate. Growth curves, derived from the regional distributions, were established for each homogeneous region. The growth curves shapes difference reflects the difference between the two homogeneous regions precipitation patterns. Therefore, to estimate the different return period's precipitation quantiles in a given site of the region, the mean precipitation of the site has to be multiplied by the corresponding regional quantile. Comparison of estimated quantiles given by the regional and local approach showed that for 50.4 % of stations, at-site model underestimates the quantiles having high return periods.

KEYWORDS: maximal annual daily precipitation, L-moments, regional frequency analysis, *GEV*, *LN3* distributions, set generation

1 INTRODUCTION

La connaissance des quantiles des précipitations de fréquences rares est nécessaire pour la conception des ouvrages hydrauliques tels que les ouvrages de protection contre les inondations, les réseaux d'assainissement pluvial et dans de nombreuses applications d'ingénierie. L'estimation de ces fréquences est difficile car les événements extrêmes par définition sont rares et les données disponibles localement proviennent souvent des observations de courte durée, insuffisantes pour pouvoir estimer les quantiles extrêmes de manière fiable. Pour

résoudre ce problème, on a recours à l'analyse régionale. Dans le concept de régionalisation, introduit par Dalrymple (1960), « trading space for time », on utilise les données de plusieurs sites pour estimer les quantiles de la variable à chaque site d'une région bien définie. Les estimations des quantiles obtenues à partir de l'échantillon régional sont considérées comme plus précises (Hosking et Wallis, 1997).

Initialement, les méthodes d'analyse fréquentielle ont été élaborées pour l'estimation des crues. Les méthodes de régionalisation des crues introduites par Dalrymple en 1960

ont été continuellement développées depuis, par des chercheurs. GREHYS (1996a), Ouarda et al. (1999 ; 2008) ont présenté et comparé les différentes méthodes de régionalisation des crues.

L'application de l'analyse fréquentielle des crues s'est étendue ensuite à la régionalisation des précipitations et a été à la base de nombreux travaux sur ces dernières. Alila (1999) a développé un modèle hiérarchique régional fréquentiel pour la régionalisation des précipitations de courte durée au Canada. En France, les travaux de Djerboua (2001) et de Mora et al. (2005) ont porté sur l'estimation régionale des précipitations journalières. Nguyen et al. (2002) ont proposé deux méthodes alternatives d'estimation des précipitations extrêmes de différentes durées. Kysely et Picek (2007) ont utilisé une méthode basée sur les L-moments pour l'estimation régionale des précipitations. L'analyse fréquentielle régionale basée sur la méthode de « index de variable » et L-moments a été utilisée par Norbiato et al. (2007) pour analyser les précipitations maximales annuelles de courte durée en Italie. Gaal et al. (2008) ont appliqué la méthode des régions d'influence dans l'analyse fréquentielle des précipitations en Slovaquie. Gellens (2002) a combiné l'approche régionale et la procédure d'extension des données dans l'estimation des précipitations extrêmes en Belgique.

Dans la bibliographie sur la régionalisation des précipitations présentée par St-Hillaire et al. (2003), on remarque que les étapes communes à la plupart des méthodes d'analyse régionale fréquentielle sont généralement : la détermination des régions statistiquement homogènes, l'identification d'une fonction de distribution régionale et finalement la détermination des paramètres de cette distribution et l'estimation des quantiles des précipitations.

L'analyse fréquentielle régionale basée sur les L-moments proposée par Hosking et Wallis (1997) est appliquée pour l'estimation des quantiles des précipitations journalières maximales annuelles en tout site de la région étudiée. Après la description de la méthode utilisée et de la région d'étude et des données, les résultats d'application des étapes de l'analyse fréquentielle régionale sont présentés et discutés dans la Section 4 ; celles-ci sont la formation des régions homogènes, l'identification de la distribution fréquentielle régionale et l'estimation des paramètres et quantiles de la distribution ajustée. Une conclusion est faite finalement

2 METHODE DE REGIONALISATION

2.1 Méthode d'analyse fréquentielle régionale

La méthode d'analyse fréquentielle régionale des précipitations extrêmes appliquée est basée sur les L-moments et associée à la procédure de la méthode « indice de crue » (Darlymple, 1960) appliquée aux données hydrologiques. La procédure utilisée est une procédure de simple scaling, les distributions de fréquences des sites d'une région homogène sont identiques à l'exception d'un

facteur d'échelle spécifique au site. Généralement, le facteur d'échelle est la moyenne de la population au site (Hosking et Wallis, 1993). Par conséquent, les quantiles de fréquence F au site i d'une région homogène de N sites peuvent être déterminés comme suit : $Q_i(F) = \mu_i q(F)$ où μ_i est le facteur d'échelle ou la moyenne au site i . Les quantiles régionaux $q(F)$ forment la « courbe de croissance régionale » définie par une distribution régionale de la variable réduite $y_{ij} = x_{ij} / \bar{x}_i$ où x_{ij} représentent les précipitations journalières maximales annuelles, \bar{x}_i leur moyenne en chaque site et $j=1,2,\dots, n_i$. n_i est l'effectif du site i .

Les paramètres de la distribution régionale sont estimés à partir de l'ensemble des statistiques locales de la région homogène. L'estimation des statistiques locales est obtenue par la méthode des L-moments.

2.2 L-moments

La théorie des L-moments a été développée par Hosking (1990). Les L-moments ont pour but de rendre les ajustements statistiques plus robustes contrairement aux moments traditionnels d'ordre élevé dont les valeurs sont très sensibles à la variabilité d'échantillonnage. Les L-moments sont analogues aux moments traditionnels mais ils sont estimés à partir de combinaisons linéaires des données ordonnées.

Pour un échantillon ordonné $x_1, x_2, x_3, \dots, x_n$ où $x_{1:n} < x_{2:n} < \dots < x_{n:n}$ les moments de probabilité pondérés (PWM) sont estimés par :

$$\begin{aligned} \beta_0 &= n^{-1} \sum_{j=1}^n x_{j:n} \\ \beta_1 &= n^{-1} \sum_{j=2}^n \frac{j-1}{n-1} x_{j:n} \\ \beta_2 &= n^{-1} \sum_{j=3}^n \frac{(j-1)(j-2)}{(n-1)(n-2)} x_{j:n} \\ \beta_3 &= n^{-1} \sum_{j=4}^n \frac{(j-1)(j-2)(j-3)}{(n-1)(n-2)(n-3)} x_{j:n} \end{aligned} \tag{1}$$

Ainsi on peut estimer les L-moments suivants en utilisant les PWM :

$$\begin{aligned}
 l_1 &= \beta_0 \\
 l_2 &= 2\beta_1 - \beta_0 \\
 l_3 &= 6\beta_2 - 6\beta_1 + \beta_0 \\
 l_4 &= 20\beta_3 - 30\beta_2 + 12\beta_1 - \beta_0
 \end{aligned} \quad (2)$$

Le premier L-moment l_1 est égal à la moyenne de la distribution et l_2 est un paramètre d'échelle (L-écart-type).

De plus, les rapports des L-moments suivants ont été introduits :

L-coefficient de variation, $L-Cv$,

$$t = l_2 / l_1 \quad (3)$$

L-coefficient d'asymétrie,

$$L-Cs, t_3 = l_3 / l_2 \quad (4)$$

L-coefficient d'aplatissement,

$$L-Ck, t_4 = l_4 / l_2 \quad (5)$$

2.3 Test d'homogénéité

Pour tester l'homogénéité d'une région (groupe de stations) en termes de rapports des L-moments le test d'homogénéité statistique proposé par Hosking et Wallis (1993) est utilisé. Dans ce test, les paramètres représentatifs d'une région sont les moyennes pondérées des statistiques des L-moments. Ainsi, pour une région de N stations chacune de taille n_i , les rapports des L-moments et les L-moments régionaux sont calculés de manière suivante :

$$\bar{t}_r = \frac{\sum_{i=1}^N n_i t_r^{(i)}}{\sum_{i=1}^N n_i} \quad r \leq 3, \quad (6)$$

où $t_r^{(i)}$ sont les rapports des L-moments d'ordre r à la station i .

Le test de simulation de Monte Carlo est utilisé afin de tester l'homogénéité d'une région. Un grand nombre des données de régions sont générées à partir de la distribution Kappa à quatre paramètres. La distribution Kappa est ajustée en utilisant les rapports des L-moments moyens pondérés.

Chacune des simulations doit refléter la configuration de la base de données de la région considérée. Plus précisément, au cours d'une simulation le nombre de sites et le nombre d'observations à chaque site doivent être reproduits. Pour chaque région générée, les trois mesures suivantes de la variabilité intersites des rapports des L-moments sont calculées comme suit (Hosking et Wallis (1997)) :

La variance pondérée du $L-Cv$:

$$V_1 = \frac{\sum_{i=1}^N n_i (t^{(i)} - \bar{t})^2}{\sum_{i=1}^N n_i} \quad (7)$$

L'écart-type pondéré pour les rapports des L-moments t et t_3 :

$$V_2 = \frac{\sum_{i=1}^N n_i \sqrt{(t^{(i)} - \bar{t})^2 + (t_3^{(i)} - \bar{t}_3)^2}}{\sum_{i=1}^N n_i} \quad (8)$$

L'écart-type pondéré pour les rapports des L-moments t_3 et t_4 :

$$V_3 = \frac{\sum_{i=1}^N n_i \sqrt{(t_3^{(i)} - \bar{t}_3)^2 + (t_4^{(i)} - \bar{t}_4)^2}}{\sum_{i=1}^N n_i} \quad (9)$$

où $t^{(i)}$, $t_3^{(i)}$ et $t_4^{(i)}$ désignent respectivement les $L-Cv$, $L-Cs$ et $L-Ck$ du site i ; \bar{t} , \bar{t}_3 et \bar{t}_4 désignent respectivement les $L-Cv$, $L-Cs$ et $L-Ck$ régionaux calculées selon l'équation (6) et N est le nombre de sites. Si on désigne par V l'une de ces trois valeurs V_1 , V_2 et V_3 , le critère d'homogénéité d'une région est calculé comme suit :

$$H_V = \frac{V_{obs} - \mu_v}{\sigma_v}, \quad (10)$$

où V_{obs} est la valeur observée de V et μ_v et σ_v sont respectivement la moyenne et l'écart-type de V obtenu par simulations. La variable H mesure la dispersion des observations à celle des simulations. Selon Hosking et Wallis (1993), une région peut être considérée comme étant acceptablement homogène si $H < 1$, probablement hétérogène si $1 \leq H < 2$ et définitivement hétérogène si $H \geq 2$.

2.3 Identification de la distribution régionale

Parmi les différentes distributions, la distribution Gumbel est celle qui est le plus souvent utilisée en Algérie dans l'analyse fréquentielle des précipitations extrêmes dans un site donné. Cette distribution a été utilisée par Mebarki (2005) dans l'analyse fréquentielle ponctuelle des précipitations journalières maximales annuelles de l'est Algérien.

Cette loi à deux paramètres est aussi largement utilisée dans différents contextes climatiques. Le comportement asymptotique de la distribution Gumbel est cependant remis en question par Koutsoyiannis (2004), qui confirme qu'elle a pour effet de sous estimer les valeurs des précipitations de fréquences élevées par rapport à la distribution *GEV* (*EV2*) (*Generalised Extreme Value, type 2*). Alila (1999) émet certaines réserves quant à l'application de la loi Gumbel dans un contexte régional. Dans son étude de la régionalisation des précipitations de courte durée au Canada, les différentes distributions ont été testées. Parmi ces distributions la distribution *GEV* a été identifiée comme étant la distribution régionale la plus appropriée. La distribution *GEV* est la plus utilisée tant pour l'analyse fréquentielle régionale des précipitations que pour celle des crues. Par exemple, Overeem et al. (2007) l'ont utilisée pour la régionalisation des précipitations de courte durée pour l'ensemble de la Hollande. Djerboua (2001), Versiani et al. (1999), Cannarozzo et al. (1995) ont choisi comme modèle statistique régional des précipitations journalières maximales annuelles la distribution à deux composantes *TCEV* (*Two Component Extreme Value*). Pour déterminer les courbes de croissance régionales des précipitations de courte durée Sveinsson et al. (2002) ont utilisé une approche régionale fréquentielle basée sur la méthode d'indice de crue en prenant en considération les différentes distributions : log-normale à trois paramètres (*LN3*), *GEV*, log-normale (*LN*) et Pearson type 3 (*P3*).

Dans cette étude, nous faisons l'hypothèse que les distributions *GEV*, *LN3*, *P3* et *GLO* (*Generalized Logistic*) s'ajustent aux séries des valeurs réduites des précipitations maximales annuelles journalières de la région d'étude. La pertinence de l'ajustement de chacune de ces distributions à trois paramètres est évaluée en termes de différence entre le L-aplatissement théorique de la distribution ajustée et le L-aplatissement régional. La signification de cette différence est estimée par la statistique *Z* (Hosking et Wallis, 1993) :

$$Z^{DIST} = \frac{\tau_4^{DIST} - \bar{t}_4 + \beta_4}{\sigma_{t_4}^-}, \quad (11)$$

où \bar{t}_4 est le L-aplatissement régional observé, τ_4^{DIST} est le L-aplatissement théorique de la distribution (*DIST*) estimé à partir du L-asymétrie régional observé. β_4 et $\sigma_{t_4}^-$ sont respectivement le biais et l'écart-type de \bar{t}_4 obtenus par simulations d'une région homogène avec la distribution *Kappa*. La statistique Z^{DIST} est basée sur la normalité asymptotique et l'ajustement est satisfait au niveau de 90% si $Z^{DIST} \leq 1.64$.

2.4 Estimation des paramètres et quantiles de la distribution régionale

Pour estimer les paramètres de la distribution régionale les trois premiers L-moments régionaux sont utilisés. La courbe de croissance régionale sera établie sur la base des paramètres de la distribution régionale en appliquant la moyenne comme facteur d'échelle. Dans cette approche les L-coefficients d'asymétrie et de variation régionaux sont supposés constants. Ainsi, pour estimer les précipitations associées aux différentes périodes de retour en un site donné d'une région homogène, il faut multiplier les valeurs du quantile régional correspondant aux mêmes périodes de retour par la moyenne des précipitations journalières maximales.

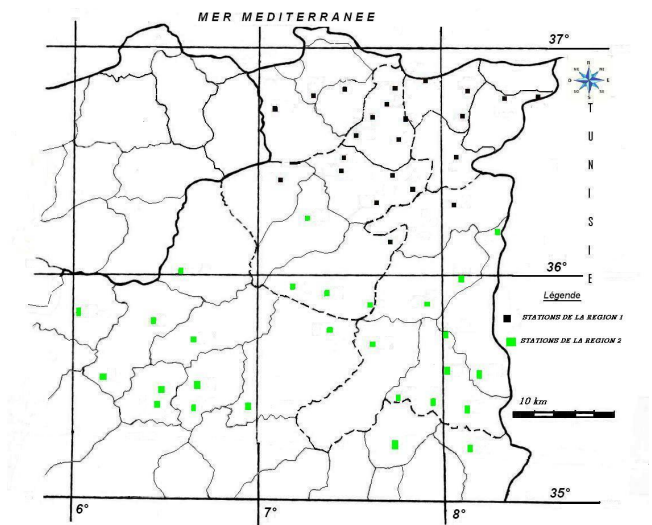


Figure1: Région d'étude et localisation des stations. Les bassins versants des deux grands oueds, ceux de la Seybouse et de la Medjerda sont délimités par une ligne discontinue

3 REGION D'ETUDE ET DONNEES

La région d'étude est située au Nord Est de l'Algérie et couvre les bassins versants de deux grands oueds, celui de la Seybouse et celui de la Medjerda (Fig.1). Les séries des précipitations journalières maximales annuelles des 50 stations de mesure ont été retenues pour cette étude. Les observations s'étendent, en général, sur la période de 1970 à 2007. La taille moyenne des échantillons est égale à 38 ans. La période des observations pour les différentes stations varie de 20 à 49 années. La région d'étude est caractérisée par un climat méditerranéen dans la partie nord et par un climat semi-aride dans la partie sud. Les précipitations moyennes annuelles varient de 804 à 233 mm et la moyenne des précipitations journalières maximales annuelles de 26.11 à 61.24 mm. Leurs gradients diminuent du nord vers le sud de la région.

4 RESULTATS ET DISCUSSION

La première étape de l'analyse fréquentielle régionale a été l'estimation de l'homogénéité des 50 stations de la région. Pour évaluer le degré d'homogénéité de la région, les 50 régions de données ont été générées en utilisant la distribution Kappa. Selon les valeurs de la mesure d'hétérogénéité H obtenues (Tableau 1), la région est probablement hétérogène en termes de $L-Cv$ ($H_{V_1} = 1.1$) et homogène en termes de $L-Cs$ et $L-Ck$. Etant donné que le H_{V_1} est légèrement supérieur à 1, l'étape suivante sera la définition des régions homogènes. La classification ascendante hiérarchique (CAH) des caractéristiques des stations (longitude, latitude, altitude, précipitations moyennes annuelles et la moyenne annuelle de nombres de jours pluvieux) a été utilisée comme outil auxiliaire pour la formation des groupes de stations. Les k-clusters ($k=2, \dots, 4$) obtenus par l'application de deux méthodes d'agrégation (méthode de « lien moyen » et méthode de Ward) ont été étudiées. Pour pouvoir mieux interpréter les résultats de CAH et les utiliser dans la formation des régions homogènes la méthode Silhouettes (Rousseeuw, 1987) a été appliquée. En s'appuyant sur cette méthode, l'homogénéité statistique des 10 différents partitionnements des stations a été testée en termes de Hosking-Wallis test. Finalement, le partitionnement des stations a abouti à la formation de deux régions homogènes (Fig.1) en termes de $L-Cv$, $L-Cs$ et $L-Ck$ (Tableau 1). D'après les résultats du tableau 1 on constate la présence des valeurs négatives de H_{v_2} et H_{v_3} . Selon Hosking et Wallis (1997), cela indique que la variabilité des statistiques des séries est moindre que ce qui serait attendu d'une région homogène avec des distributions de fréquences indépendantes du site. Il s'agit généralement d'une indication d'une large corrélation croisée entre les distributions de fréquences des sites.

Sur la figure 1, on observe une cohérence géographique et climatique dans le groupement des stations. Les stations de la région 1 sont situées dans la partie nord de la zone d'étude caractérisée par un climat méditerranéen. Les précipitations moyennes annuelles de cette région varient de 804 à 442 mm. Les stations de la région 2 sont situées dans la partie sud sujette d'un climat semi-aride. Les précipitations moyennes annuelles de cette région varient de 512 à 233 mm. Les statistiques régionales sont présentées dans le tableau 2

Tableau 1: Résultats du test d'homogénéité des différentes régions

Régions	H_{V_1}	H_{V_2}	H_{V_3}
Région	1.10	-0.53	-0.89
Région 1	0,79	-0,80	-0,91
Région 2	0,93	-0,32	-0,78

Tableau 2: Statistiques régionales

Régions	N	\bar{l}_1	\bar{l}_2	\bar{t}	\bar{t}_3	\bar{t}_4
Région	50	43.83	9.59	0.219	0.190	0.154
Région 1	23	51,61	10,98	0,211	0,174	0,154
Région 2	27	36,48	8,29	0.227	0,204	0,151

N : nombre de stations; \bar{l}_1 et \bar{l}_2 : la moyenne et L-écart-type régionaux pondérés des précipitations (exprimés en mm)

Tableau 3: L-coefficient d'aplatissement théorique et Z-statistique des différentes distributions.

Distribution	Région 1		Région 2	
	τ_4^{DIST}	Z	τ_4^{DIST}	Z
GEV	0,152	-0,24	0,175	0.87
LN3	0,146	-0,57	0,163	0.32
P3	0,126	-1,87	0,128	-1,69
GLO	0,192	2,29	0,209	2.79

Les 2 régions étant homogènes, pour identifier la distribution régionale de chaque région parmi les distributions GEV , $P3$, $LN3$ et GLO , la Z^{DIST} statistique a été calculée pour chaque distribution. Les valeurs de cette dernière obtenues par la réalisation de 500 simulations en utilisant la distribution Kappa et les valeurs de L-aplatissement théorique de chaque distribution ajustées sont données dans le tableau 3. Parmi les quatre distributions considérées, les deux distributions GEV et $LN3$ sont les distributions dont l'ajustement aux données régionales de deux régions est le plus satisfaisant. De plus, la valeur absolue de la statistique Z de la distribution GEV est inférieure à celle de $LN3$ pour la région homogène 1 et supérieur pour la région 2. Pour faire le choix définitif des distributions régionales, le diagramme des rapports des L-moments a été utilisé. Les rapports des L-moments (t_3, t_4) des échantillons et leurs moyennes régionales pondérées ont été portés sur le diagramme (Fig.2). Les points définis par les valeurs des moyennes régionales des $L-Cs$ et $L-Ck$ (Fig.2) sont proches de la distribution GEV pour la région 1 et de $LN3$ pour la région 2. Ainsi en se basant sur le diagramme des rapports des L-moments et sur les valeurs de la Z-statistique, les distributions GEV et $LN3$ sont identifiées comme les distributions les plus robustes respectivement pour les régions 1 et 2

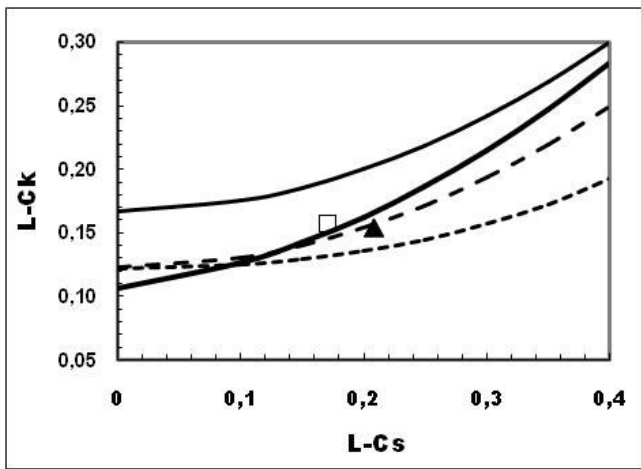
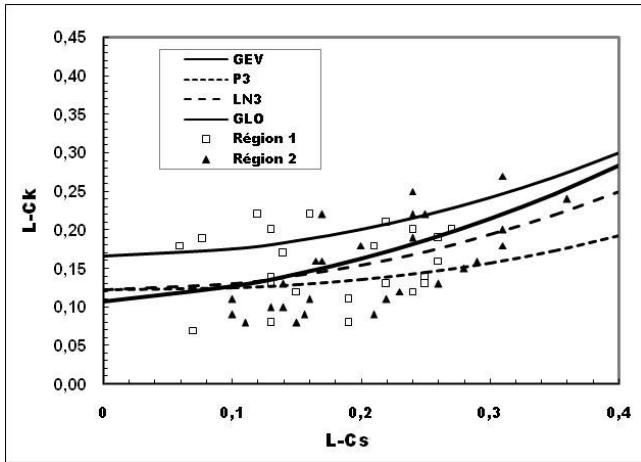


Figure2:Diagramme des rapports des L-moments. Les valeurs des rapports des L-moments (t_4 / t_3) des stations (moyennes régionales) sont présentées à gauche (droite) de la figure

Tableau 4: Paramètres des distributions régionales GEV et LN3

Régions		ξ	α	$\sigma (k)$
Région 1	GEV	0,826	0,302	-0,010
Région 2	LN3	0,918	0,373	-0,422

La fonction quantile de la distribution GEV est la suivante :

$$x(F) = \xi + \frac{\alpha}{k} \left\{ 1 - [-\ln(F(x))]^k \right\} \text{ pour } k \neq 0 \quad (12)$$

La fonction cumulative de la distribution LN3 est la suivante :

$$F(x) = \frac{1}{\sigma \sqrt{2\pi}} \int_0^u \frac{1}{u} e^{-\frac{\ln^2(u)}{2\sigma^2}} du, \quad (13)$$

où $u = \frac{x - \xi}{\alpha}$ et les ξ , α et $\sigma (k)$ sont respectivement les

paramètres de position, d'échelle et de forme. Pour un échantillon fini, ces paramètres sont définis à partir des L-moments l_1 , l_2 et t_3 par les équations suivantes :

Pour la distribution GEV :

$$\xi = l_1 - \frac{\alpha}{k} [1 - \Gamma(1 + k)] \quad (14)$$

$$\alpha = \frac{l_2 k}{(1 - 2^{-k}) \Gamma(1 + k)} \quad (15)$$

$$k = 7.859 \left(\frac{2}{3+t_3} - \frac{\ln 2}{\ln 3} \right) + 2.9554 \left(\frac{2}{3+t_3} - \frac{\ln 2}{\ln 3} \right)^2 \quad (16)$$

Pour la distribution LN3 :

$$\xi = l_1 - \frac{\alpha}{\sigma} \left(1 - e^{\sigma^2 / 2} \right) \quad (17)$$

$$\alpha = \frac{l_2 \sigma \exp(-\sigma^2 / 2)}{\text{erf}(\sigma / 2)} \quad (18)$$

$$\sigma = -t_3 \frac{A_0 + A_1 t_3^2 + A_2 t_3^4 + A_3 t_3^6}{1 + B_1 t_3^2 + B_2 t_3^4 + B_3 t_3^6}, \quad (19)$$

où $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$ représente la fonction Gamma ; erf

est la fonction d'erreur ; A_0, \dots, A_3 et B_1, B_2, B_3 sont les constantes d'approximation.

Les paramètres régionaux de ces distributions (Tableau 4) ont été estimés à partir des rapports des L-moments régionaux comme expliqué dans la deuxième section

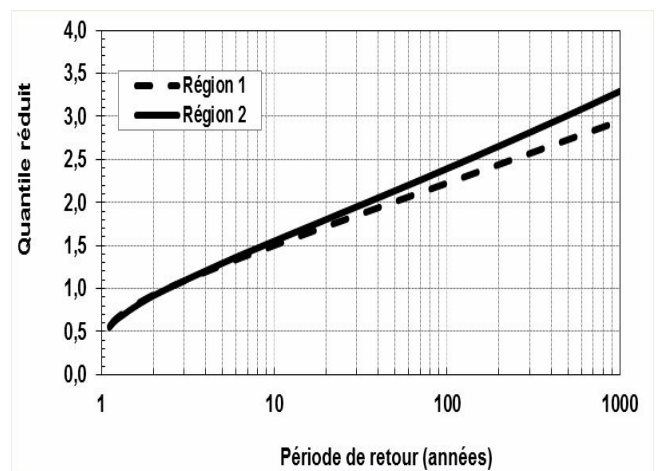


Figure 3: Courbes de croissance

Les courbes de croissance régionales, dérivées des distributions régionales ont été tracées pour les périodes de retour spécifiées T (Fig. 3). Ces courbes de croissance traduisent la variation du quantile régional $q(F)$ en fonction de la probabilité au non dépassement F ou en fonction de la période de retour T , ($T = 1/(1-F)$). On observe à partir de la figure 3 que les valeurs des quantiles régionaux de la région 2 sont plus fortes que celles de la région 1 pour les périodes de retour élevées.

La figure 4 présente la distribution de fréquences empiriques des valeurs réduites des précipitations observées de l'échantillon régional de chaque région homogène. Les fréquences empiriques ont été calculées par la formule de Cunnane $((i-0.4)/(N-0.2))$ où N désigne l'effectif total de l'échantillon et i le rang de l'observation, par ordre décroissant).

Les courbes de croissance régionales, reportée sur la figure, s'ajuste bien à la distribution des valeurs réduites observées correspondantes. Ceci confirme une fois de plus le choix de GEV et $LN3$ comme distributions régionales. On constate également que le comportement asymptotique de la distribution régionale GEV est quasi exponentiel étant donné que son paramètre de forme est proche de zéro ($k=0.01$).

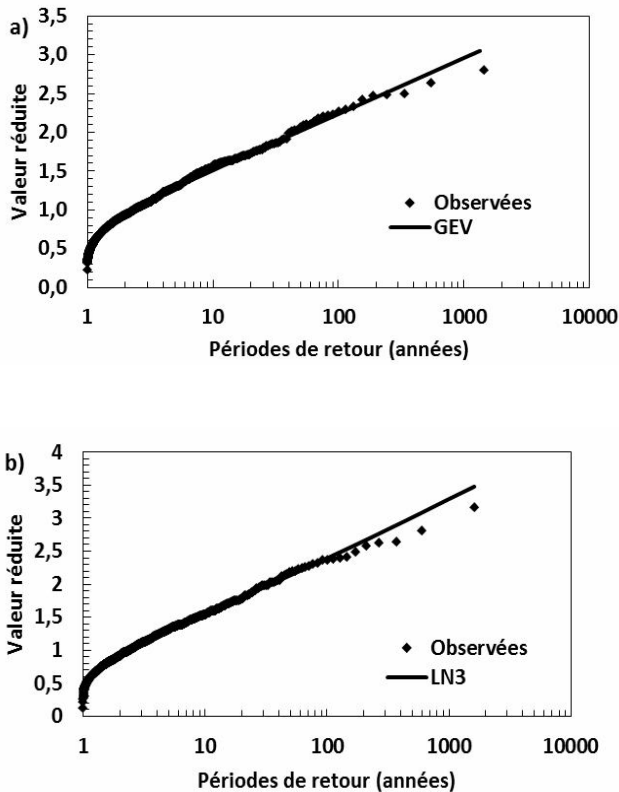


Figure 4: Ajustement des distributions régionales GEV et $LN3$ aux valeurs réduites des précipitations de l'échantillon régional. a) région 1 ; b) région 2

Pour déterminer la précision et quantifier les incertitudes d'estimation des quantiles régionaux l'algorithme de

simulation Monte Carlo (Hosking et Wallis, 1997) a été utilisé. Cet algorithme est basé sur l'application de la procédure de ré-échantillonnage (bootstrap) pour calculer l'intervalle de confiance des quantiles régionaux de période de retour spécifiée (q_T). Pour réaliser la corrélation entre les sites d'une région, une matrice de corrélation a été utilisée avec un coefficient de corrélation pris égal à 0.65 selon Hosking et Wallis (1997). En décomposant la matrice de corrélation (décomposition de Cholesky) on obtient la matrice triangulaire qui sert à générer une matrice de Nn_{max} variables aléatoires corrélées uniformes (N est le nombre des stations et n_{max} est la taille maximale d'un site de la région). Les 10000 (R) réalisations de chaque région ont été effectuées. Pour chaque réalisation les quantiles régionaux spécifiés \hat{q}_T ont été estimés à partir de l'ajustement des échantillons générés à la loi GEV pour la région 1 et à la loi $LN3$ pour la région 2. Ainsi on constitue une série des quantiles simulés $q = \{\hat{q}_{T,1}, \dots, \hat{q}_{T,R}\}$. Dans la série triée les quantiles empiriques $q_{T,\alpha/2}$ et $q_{T,1-\alpha/2}$ de probabilité au non dépassement de respectivement $\alpha/2$ et $1-\alpha/2$ ont été déterminés. On obtient donc l'intervalle de confiance de q_T au niveau $1-\alpha$: $[q_{T,\alpha/2}; q_{T,1-\alpha/2}]$.

Les valeurs de la REQM (racine carrée de l'erreur quadratique moyenne) relative régionale et de l'intervalle de confiance (90%) des deux courbes de croissance estimées pour les différentes périodes de retour sont présentées dans le tableau 5. Les faibles valeurs de la REQM démontrent la fiabilité de la méthode d'analyse fréquentielle régionale appliquée. Les intervalles de confiance (90%) obtenus sont peu larges du fait des distributions régionales quasi exponentielles.

Tableau 5: REQM relative et intervalle de confiance (90%) dans l'estimation des quantiles régionaux $q(T)$ de période de retour T .

T (années)	q(T)	REQM	Intervalle de confiance (90%)	
			Inférieur	Supérieur
Région 1				
5	1.273	0.017	1.225	1.317
10	1.502	0.016	1.427	1.568
50	2.001	0.031	1.841	2.113
100	2.226	0.043	2.003	2.372
200	2.442	0.051	2.137	2.617
1000	2.948	0.063	2.506	3.202
Région 2				
5	1.295	0.010	1.245	1.340
10	1.553	0.018	1.474	1.626
50	2.137	0.022	2.002	2.261
100	2.394	0.031	2.155	2.559
200	2.656	0.042	2.324	2.873
1000	3.292	0.059	2.798	3.585

Pour voir l'efficacité de l'analyse fréquentielle régionale, les valeurs des quantiles estimés à partir de l'analyse régionale et locale sont comparées. Sur la figure 5 sont présentées les valeurs des quantiles centennaux estimés à partir de l'approche régionale et locale (*GEV* et *LN3*). La comparaison de ces valeurs montre que l'ajustement local sous estime les quantiles pour 50.4% des stations avec un écart relatif maximal égal à 32.7 %. Cet écart augmente avec l'accroissement de la période de retour pour attendre la valeur maximale de 51.65 % pour T=1000 ans. On constate, d'après la figure 5, que l'approche régionale réduit la variabilité entre les sites des valeurs des quantiles estimés à partir des modèles locaux d'une région homogène. Par exemple, pour la région 1, constituée de 23 stations, d'après le résultat de l'analyse fréquentielle locale les valeurs des quantiles centennaux varient entre 80 et 159 mm, tandis que dans l'analyse régionale elles varient entre 94 et 137 mm. Pour la région 2, les valeurs des quantiles « locaux » varient entre 59 et 116 mm et celles des quantiles « régionaux » varient de 64 à 106 mm. On constate, également, que les valeurs des quantiles diminuent du nord au sud de la région d'étude (Fig.5).

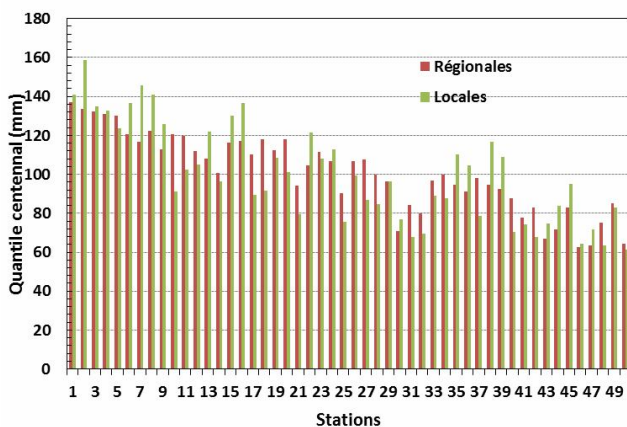


Figure 5: Comparaison des quantiles centennaux estimés par l'approche régionale et locale. Les stations de la région 1 sont numérotées de 1 à 23, celles de la région 2 sont numérotées de 24 à 50

Pour vérifier la cohérence des résultats obtenus avec l'hypothèse d'homogénéité et les lois régionales choisies, la distribution *GEV* et *LN3* ont été reportées sur la distribution en fréquences empiriques des valeurs réduites observées aux trois stations de différentes tailles (Fig.6). On note le bon ajustement des distributions régionales aux échantillons des précipitations observées de trois stations en particulier aux précipitations de période de retour élevée. Donc, les résultats obtenus sont cohérents avec l'hypothèse d'homogénéité selon laquelle les valeurs réduites des précipitations suivent en tous points d'une région homogène la même loi de distribution. Pour la première station de taille n=49, le modèle local *GEV* surestime les quantiles (Fig.6a). Pour les deux autres stations de taille n=34 et n=24, le modèle local *LN3* sous estime les quantiles (Fig.6 b, c). La différence entre les quantiles réduits des

courbes de croissance régionales et locales apparaît à partir de la période de retour de 20 ans.

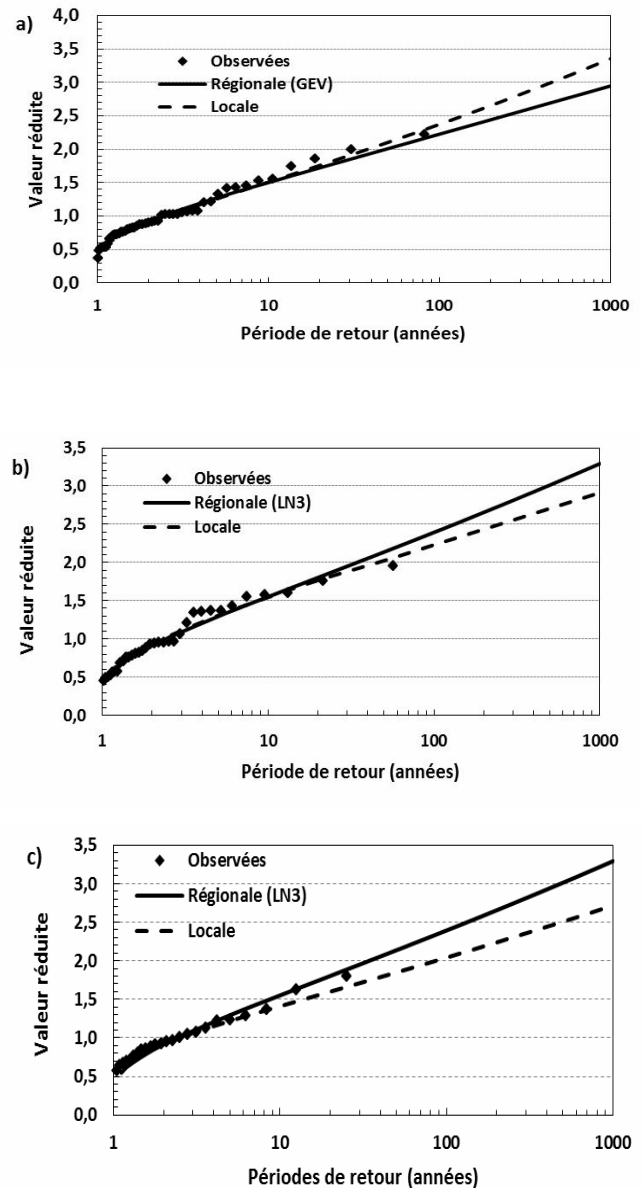


Figure 6: Ajustement des distributions régionales (*GEV*, *LN3*) et locales aux valeurs réduites des précipitations observées des séries de différentes tailles (n) : a) n=49 ; b) n=34 ; c) n=24.

5 CONCLUSIONS

Cette étude a porté sur la régionalisation des précipitations journalières maximales annuelles, dans le but d'obtenir une qualité fiable de l'estimation des quantiles extrêmes dans les stations de la région d'étude. La méthode d'analyse régionale fréquentielle basée sur les L-moments introduite par Hosking et Wallis (1997) a été utilisée. L'application de test d'homogénéité a permis de diviser la zone d'étude en deux régions statiquement homogènes. Le groupement des stations présente une cohérence géographique et

climatique. La première région est constituée des stations situées dans la partie nord de la région d'étude caractérisée par un climat méditerranéen. Les stations de la deuxième région sont situées dans sa partie sud caractérisée par un climat semi-aride.

L'étape importante de l'analyse régionale a été l'identification de la distribution régionale de chaque région homogène. Les distributions à trois paramètres *GEV*, *LN3*, *GLO* et *P3* ont été testées. Parmi ces dernières, selon le test de la *Z*-statistique et le diagramme des rapports des *L*-moments, la distribution *GEV* a été identifiée comme la plus adéquate distribution régionale pour la région 1 et la distribution *LN3* pour la région 2.

La courbe de croissance régionale a été établie pour chaque région homogène sur la base des paramètres de la distribution régionale. Ainsi, pour estimer les précipitations associées aux différentes périodes de retour en un site donné d'une région homogène, il faut multiplier les valeurs du quantile régional correspondant aux mêmes périodes de retour par la moyenne des précipitations journalières maximales.

Pour évaluer la précision et quantifier les incertitudes de l'estimation des quantiles régionaux le test de simulation Monte Carlo a été effectué. Les faibles valeurs de la REQM relative démontrent la fiabilité de la méthode appliquée.

La comparaison des quantiles des précipitations extrêmes estimés à partir de l'approche régionale et locale a montré que pour 50.4% des stations de la région d'étude les modèles locaux sous estime les quantiles de fortes périodes de retour.

REFERENCES

- [1] Alila, Y. (1999) A hierarchical approach for the regionalization of precipitation annual maxima in Canada. *J Geophys Res.* **104**(D24), 31645-31655.
- [2] Cannarozzo, M., D'Asaro, F., & Ferro, V. (1995) Regional and frequency analysis for Sicily using the two component extreme value distribution. *Hydrol. Sci. J.* **40**(1), 19-42.
- [3] Djerboua, A. (2001) Prédétermination des pluies et crues extrêmes dans les Alpes franco-italiennes. Prévision quantitative des pluies journalières par la méthode des analogues. Thèse de doctorat de l'Institut National Polytechnique de Grenoble, France.
- [4] Dalrymple, T. (1960) Flood frequency methods. *US Geol. Survey Water Supply Paper.* **1543A**, 11-51.
- [5] GREHYS (1996a) Presentation and review of some methods for regional flood frequency analysis. *J. Hydrol.* **186**, 63-84.
- [6] Gaal, L., Kysely, J. & Szolgay, J. (2008). Region-of-influence approach to a frequency analysis of heavy precipitation in Slovakia. *Hydrol. Earth Syst. Sci.* **12**, 825-839.
- [7] Gellens, D. (2002). Combining regional approach and data extension procedure for assessing *GEV* distribution of extreme precipitation in Belgium. *J. Hydrol.* **268**, 113-126.
- [8] Hosking, J. R. M. (1990) *L*-moments: analysis and estimation of distribution using linear combination of order statistics. *J. Roy. Statist. Soc.* **B52** (1), 105-124.
- [9] Hosking, J. R. M. & Wallis, J. R. (1993) Some statistics useful in regional frequency analysis. *Water Resour. Res.* **29**(2), 271-281.
- [10] Hosking, J. R. M. & Wallis, J. R. (1997) *Regional frequency analysis: an approach based on L-moments*. Cambridge University Press, Cambridge, UK.
- [11] Koutsoyiannis, D. (2004) Statistics of extremes and estimation of extreme rainfall: II. Empirical investigation of long rainfall records. *Hydrol. Sci. J.* **49**(4), 591-610.
- [12] Kysely, J. & Picek, J. (2007) Regional growth curves and improved design value estimates of extreme precipitation events in the Czech Republic. *Climate Research.* **33**, 243-255.
- [13] Mora, R. D., Bouvier, C., Neppel, L. & Neel, H. (2005) Regional approach for the estimation of low-frequency distribution of daily rainfall in the Lanquedoc-Roussillon region, France. *Hydrol. Sci. J.* **30**(1), 85-109.
- [14] Mebarki, A. (2005) Hydrologie des bassins de l'est Algérien: Ressources en eau, aménagement et environnement. Thèse de doctorat d'état. Université Mentouri de Constantine, Algérie.
- [15] Nguyen, V. T. V., Nguyen, T. D. & Ashkar, F. (2002) Regional frequency analysis of extreme rainfall. *Water Sci. Techn.* **45**(2), 75-81.
- [16] Norbiato, D., Borga, M., Sangati, M. & Zanon, F. (2007). Regional frequency analysis of extreme precipitation in the eastern Italian Alps and the August 29, 2003 flash flood. *J. Hydrol.* **345**, 149-166.
- [17] Ouarda, T. B. M., Long, M., Bobée, B., Bernier, J. & Bois, P. (1999) Synthèse de modèles régionaux d'estimation de crues utilisées en France et au Québec. *Rev. Sci. Eau* **12**(1), 155-182.
- [18] Ouarda, T. B.M., St-Hilaire, A. & Bobée, B. (2008) Synthèse des développements récents en analyse régionale des extrêmes hydrologiques. *Rev. Sci. Eau* **21**(2), 219-232.
- [19] Overeem, A., Buishand, A. & Holleman, I. (2007) Rainfall depth-duration-frequency curves and their uncertainties. *J. Hydrol.* **348**, 124-134.
- [20] Rousseeuw PJ (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J. Computat. Appl. Math.* **20**, 53-65.
- [21] St-Hilaire, A., Ouarda, T. B. M., Lachange, M., Bobée, B., Barbet, M. & Bruneau, P. (2003) La régionalisation des précipitations : une revue bibliographique des développements récents. *Rev. Sci. Eau* **16**, 27-54.

[22] Sveisson, O. G. B., Salas, J. & Duane, G. B. (2002) Regional frequency analysis of extreme precipitation in northern Colorado and the Fort Collins flood of 1997. *J. Hydrol. Engng. ASCE*. **7**(1), 49-63.

[23] Versiani, B. R., De Andrade Pinto, E. J. & Bois, P. (1999) Analyse des pluies extrêmes annuelles sur la région de Minas Gerais (Brésil): modèle de régionalisation TCEV. *IAHS Publ.* **255**, 201-207.