# Estimation des lois des précipitations extrêmes à partir de données journalières complètes

*Extreme precipitation laws estimated by daily data*

**Abdelatif Djerboua[1], Daniel Duband[2], Philippe Bois[1]**

*This study summarises a method for the estimation of extreme rainfall and it describes a number of probabilistic laws, used for data adjustment. The main issue deals with the choice of the law that best estimates the heaviness of statistical parameters for extreme rainfall. The results of this study show that the reliability of the method depends mainly on the type of data.*
*For this study, the daily rainfall was collected from 45 raingauge stations, which are covering 6 catchments in France and Italy. The average rainfall is calculated for each catchment. From this database, series of daily successive, weekly maximum, 15-daily maximum and monthly maximum rainfall data were constructed.*
*Three probabilistic laws have been selected for the adjustment of the rainfall : Simple exponential, Sum of two exponentials and Gumbel. The sum of two exponentials was applied for daily successive rainfall, which represents the reference of the estimate. For the same data, two empirical formulas were used. The Gumbel law is applied for the 15-daily maximum and monthly maximum rainfall data. However, these two laws are not suitable for the weekly maximum data. For these data, the Gumbel law adjusted by including several thresholds. The use of simple exponential, sum of two exponentials and empirical formulas to describe these data is less accurate. The analysis of the results is made for two time scales (monthly, seasonal) at the local (over the raingauge station) and basin scale.*

## I ■ INTRODUCTION

L'aléa pluviométrique et l'aléa hydrologique associé s'évaluent classiquement en terme de probabilité au non dépassement, ce qui fournit un outil d'aide à la décision, notamment en terme de protection ou d'aménagement. On cherche donc, à partir de relevés historiques, à ajuster sur les données des lois de probabilité ayant souvent une base théorique. Les lois calées seront utilisées en interpolation dans le domaine des fréquences courantes, mais aussi en extrapolation pour les fréquences rares. Se posent alors deux problèmes : le choix de la famille de lois et le calcul des paramètres de la loi retenue.

Ce travail concerne essentiellement l'étude des lois de pluies extrêmes à pas de temps journalier, mesure la plus courante. Une méthode classique, rapide et peu onéreuse en données, consiste à étudier les maxima annuels journaliers. Il est beaucoup plus intéressant et plus robuste de travailler sur des maxima observés sur des périodes plus courtes, le mois ou la semaine ou même d'utiliser toute l'information disponible.

Cette recherche a été effectuée dans le cadre d'un projet européen (Interreg II) de comparaison de méthode de prédétermination de crues dans des bassins alpins en France et en Italie. Nous disposions de 45 séries pluviométriques complètes. Avec les séries pluviométriques, nous avons constitué des échantillons de maxima hebdomadaires, et par quinzaine et mensuels. Pour chaque type de séries, nous avons sélectionné les lois les plus adaptées pour leur utilisation ultérieure dans la prédétermination des crues, notamment par la méthode du Gradex (Bouvard *et al.* [1]), en nous basant sur une expérience acquise dans d'autres régions voisines d'un point de vue climatologique ; puis nous avons comparé les différents ajustements.

Un des paramètres les plus importants de ces lois est le gradient d'augmentation des valeurs avec la probabilité au non dépassement. Ce paramètre est déterminant pour l'extrapolation des valeurs à des fréquences faibles au dépassement et sert de base à une méthode de prédétermination des crues. Ceci explique le soin apporté à son estimation.

1. UMR 5564 (CNRS-INPG-UJF-IRD), Laboratoire d'Etudes des Transferts en Hydrologie et Environnement, BP 53, 38041 Grenoble Cedex.
Tél. : 04 76 82 50 54, Fax : 04 76 82 50 01. e-mail : philippe.bois@hmg.inpg.fr ; abdellatif.djerboua@hmg.inpg.fr
2. Société Hydrotechnique de France, 25, rue des Favorites F-75015 Paris. e-mail : d.duband@shf.asso.fr

## II ■ RAPPELS SUR LA MÉTHODE DU GRADEX DE PRÉDÉTERMINATION DES CRUES ET IMPORTANCE DE LA NOTION DE GRADEX

Une méthode très utilisée en France, notamment par Electricité de France, pour la prédétermination probabiliste des crues de temps de retour élevés, supérieurs à 100 ans est la méthode dite du Gradex (Bouvard *et al.* [1]). Dans une première étape, on doit définir un pas de temps $D$ caractéristique des crues du bassin versant ; c'est la durée moyenne de passage de 80 % du volume de chaque crue importante, de l'ordre de 24 heures pour de nombreux bassins de quelques milliers de km$^2$, plus petit pour des petits bassins et de quelques jours pour des grands bassins français.

On étudie alors la loi des pluies extrêmes spatiales cumulées sur cette durée $D$. Pour les valeurs élevées, on a constaté, sur un très grand nombre de bassins, que, sur un papier à échelle de Gumbel, c'est-à-dire comportant en abscisse $-\mathrm{Log}(-\mathrm{Log}(F(x))$ où $F(x)$ est la probabilité au non dépassement de $x$, que les points sont quasiment alignés ; ceci revient à dire que l'on peut considérer que la branche supérieure de la distribution est du type exponentiel.

On extrapole alors les volumes de crue sur $D$, exprimés dans la même unité que les cumuls de pluie sur $D$, parallèlement à la droite des pluies, à partir d'un temps de retour situé entre 10 et 30 ans. D'un point de vue hydrologique, on fait l'hypothèse que la loi conditionnelle de l'infiltration tend vers une forme limite indépendante de l'intensité de la pluie, pour les fortes pluies. Ceci n'est évidemment pas vérifiable sur tous les bassins.

On conçoit alors aisément que le paramètre important de la distribution des cumuls de pluies extrêmes sur D est la pente de la droite $\frac{\Delta P}{\Delta u}$ où $\Delta P$ est l'accroissement de pluie correspondant à un accroissement $\Delta u$ de la variable réduite de Gumbel $u = -\mathrm{Log}(-\mathrm{Log}(F(x)))$. Cette pente est appelée gradex (valeur du gradient de l'exponentielle).

Pour avoir les débits de pointe, on multiplie le débit moyen correspondant relatif à la durée $D$ et au temps de retour $T$ par un coefficient de forme $k$ moyen obtenu à partir des crues importantes observées ; ce coefficient est de l'ordre de 1,5. Notons que $D$ n'a pas à être défini avec précision ; d'ailleurs, cette durée $D$ diffère d'une crue à l'autre. En effet, si l'on refait les mêmes calculs avec des durées comprises entre $3/4{*}D$ ou $2{*}D$, les résultats seront évidemment différents sur les volumes, mais voisins sur les pointes de crue car le coefficient $k$ sera modifié dans un rapport voisin de celui des volumes de crue. C'est ce qui explique que l'on utilise bien souvent le pas de temps de 24 heures pour des bassins de superficie de l'ordre de plusieurs centaines de km$^2$, ce qui permet d'utiliser les données de pluviomètres.

La *figure 1* illustre les principes de la méthode.

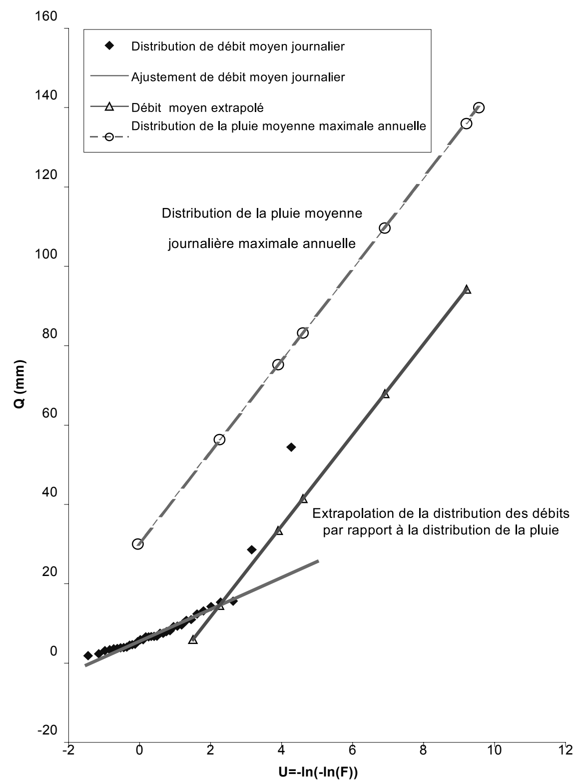


**Figure 1 : Extrapolation de la distribution des débits moyens journaliers max. annuels par la distribution des pluies journalières max. annuelles.**

## III ■ DONNÉES PLUVIOMÉTRIQUES UTILISÉES DANS CETTE ÉTUDE

Nous avons collecté les données pluviométriques des stations des 6 bassins versants de 582 à 2 000 km$^2$, fournies dans le cadre du projet Interreg II 1994-99.

Les bassins d'étude sont l'Arc à Bramans (bassin français), la Stura di Lanzo, Orco, la Dora Riparia, Toce et Sésia (bassins italiens). Cinquante séries de stations ont été fournies par EDF, Météo France et le service hydrologique de la région Piémont. Après une critique des données, on a retenu 45 stations pour la période commune 1951-1986 (car les données italiennes ne sont disponibles que jusqu'en 1986), soit une période de 36 ans qui nous semble suffisante pour l'estimation du gradex des pluies des bassins cibles de l'étude Interreg *(fig. 2)*.

Pour chaque jour et chaque bassin, on a évalué la pluie moyenne sur le bassin à partir des pluviomètres situés dans le bassin correspondant. Selon Lebel (1984 [6]), le réseau est satisfaisant si les distances entre stations sont inférieures au tiers ou au quart de la portée du phénomène. Cette portée dépend du pas de temps, elle est plus faible pour les phénomènes orageux que pour les phénomènes frontaux. L'ordre de grandeur de cette portée est de $d = 20\sqrt{\Delta t}$, $d$ en km, $\Delta t$ en heures, soit un espacement maximum de 32 km entre stations pour le pas de temps de 24 heures, ce qui est le cas. Le *tableau 1* donne les caractéristiques des bassins étudiés.

*Tableau 1 : Caractéristiques des bassins de l'étude Interreg.*

| Région | Bassin | Surface ($km^2$) | Nombre pluvio |
|---|---|---|---|
| | Sésia | 600 | 9 |
| | Toce | 2 000 | 9 |
| Région Piemont | Orco | 800 | 5 |
| | Dora Riparia | 700 | 7 |
| | Stura Di Lanzo | 582 | 10 |
| France | Arc à Bramans | 635 | 5 |

## IV ■ PRÉSENTATION DES LOIS D'AJUSTEMENTS UTILISÉES

Compte tenu que cette étude était essentiellement destinée à l'estimation de débits de crues par la méthode du Gradex [1], les lois d'ajustement utilisées sont très liées à cette méthode (lois de type exponentiel). Et le point le plus important est l'estimation de la pente d'accroissement du volume des pluies extrêmes en fonction de la probabilité au non dépassement, en variable de Gumbel.

Signalons qu'il existe d'autres lois et d'autres recherches en cours sur les valeurs extrêmes de précipitations. Notamment, certains chercheurs pensent que la distribution exponentielle sous-estime les pluies de temps de retour très élevés, et qu'il faut lui préférer plutôt des lois algébriques.

### ● IV.1 Loi simple exponentielle

La loi à décroissance exponentielle la plus simple est une simple exponentielle :

$$F(x) = 1 - \exp(-x/a) \text{ avec } F(x) = \text{Probabilité } [X <= x]$$

Elle s'applique sur des échantillons qui ont une fréquence des valeurs nulles égale à zéro. Pour les cas où la fréquence des valeurs nulles est différente de zéro, on peut utiliser une loi plus générale :

$$F(x) = 1 - \alpha \exp(-x/a) \text{ avec } F(x) = \text{Probabilité } [X <= x]$$

Il existe plusieurs méthodes d'estimation des paramètres d'une loi de probabilité. La méthode la plus simple dans notre cas est la méthode des moments qui donne :

$$\alpha = 1 - F(0) : \text{où } F(0) \text{ est la fréquence des valeurs nulles}$$

si $\mu$ est la moyenne sur la population, qu'on estime par la moyenne calculée sur l'échantillon $a = \mu/\alpha$ est le *gradex*, qui exprime le gradient des valeurs extrêmes.

Dans le cas ou la fréquence des valeurs nulles est égale à zéro, on trouve $\alpha$ égal à 1 et on retrouve le premier cas de notre loi.

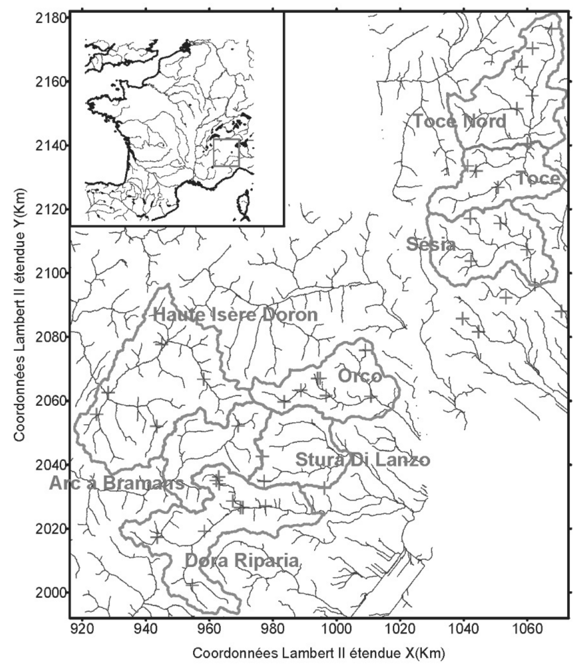


**Figure 2 : Présentation des bassins cibles de l'étude Interreg.**

### ● IV.2 Loi somme de deux exponentielles

C'est l'une des rares lois qui donne une excellente approximation de la distribution empirique des séries de pluies (y compris les valeurs nulles) à différents pas de temps, allant de 2 heures jusqu'à 5 jours. On ajuste à la distribution empirique la fonction de répartition suivante (D. Duband, 1967 [3], 1985 [4]) et (Smith et Schreiber, 1974 [10]) :

$$F(x) = 1 - \alpha \exp(-x/a) - \beta \exp(-x/c)$$

Cette loi a 4 paramètres $\alpha$, $\beta$, $a$ et $c$ :
$\alpha$, $\beta$, sans dimension sont deux paramètres de forme,
$\alpha + \beta = \theta = 1 - F(0)$ : fréquence des précipitations non nulles, $F(0)$ est estimée expérimentalement à partir de l'échantillon.

Pour éviter d'utiliser des Moments d'ordre supérieur à l'ordre 2, donc peu robustes, on fait une hypothèse sur $\alpha$ : $\alpha = 1/CV^2 = \mu^2/\sigma^2$, avec $\sigma$ écart-type.

Cette hypothèse proposée en 1967, a été vérifiée par D. Duband (1985 [4]) en utilisant un ajustement de moindres carrés sur la partie supérieure de la distribution :

$$F(x) = 1 - \alpha \exp(-x/a) \text{ pour } F(x) > 0{,}95$$

Une transformation logarithmique simple est appliquée sur cette formule, d'où on obtient :

$$\mathrm{Ln}(1 - F(x)) = \ln(\alpha) - x/a$$

Par un ajustement de moindres carrés, on en déduit les valeurs de $\alpha$ et $a$. On constate ainsi $\alpha = m^2/s^2.m$, $s$ sont respectivement la moyenne et l'écart-type estimés.

Le paramètre $\beta$ sera déduit de la relation $F(0) = 1 - \alpha\text{-}\beta$. Donc, il nous reste à déterminer les valeurs de $a$ et $c$. Quand on égalise les deux premiers moments théoriques de la loi avec les deux premiers moments empiriques, on trouve deux équations à deux inconnues :

$$\text{Moyenne : } \mu = \alpha \times a + \beta \times c = m$$

$$\text{Variance : } \sigma^2 = 2\alpha \times a^2 + 2\beta \times c^2 - \mu^2 = s^2$$

Le système se résume à une équation du second degré en fonction de $a$ : $\alpha\theta a^2 - 2\alpha\mu a - \beta K + \mu^2 = 0$.

Cette équation a des solutions en fonction de la valeur de $\Delta$ : $\Delta = \alpha\beta(\theta K - \mu^2)$ avec $K = (\sigma^2 + \mu^2)/2$.

On peut écrire $\Delta$ d'une autre façon :

$$\Delta = \alpha\beta K(\theta - \mu^2/K) = \alpha\beta K(1 - F(0) - 2/(1 + CV^2))$$

$\alpha$ et $K$ sont positifs quelles que soient les valeurs de l'échantillon ; il reste à étudier le signe du produit :

$$\beta(1 - F(0) - 2/(1 + CV^2))$$

Cette loi a ses limites d'application. Puisque dans nos hypothèses, l'on considère que $\beta$ varie entre 0 et 1, la loi dépendra essentiellement des valeurs du coefficient de variation et de la fréquence des valeurs nulles. On présente dans la *figure 3* le domaine d'application de cette loi (D. Gouy, 1994 [5]).

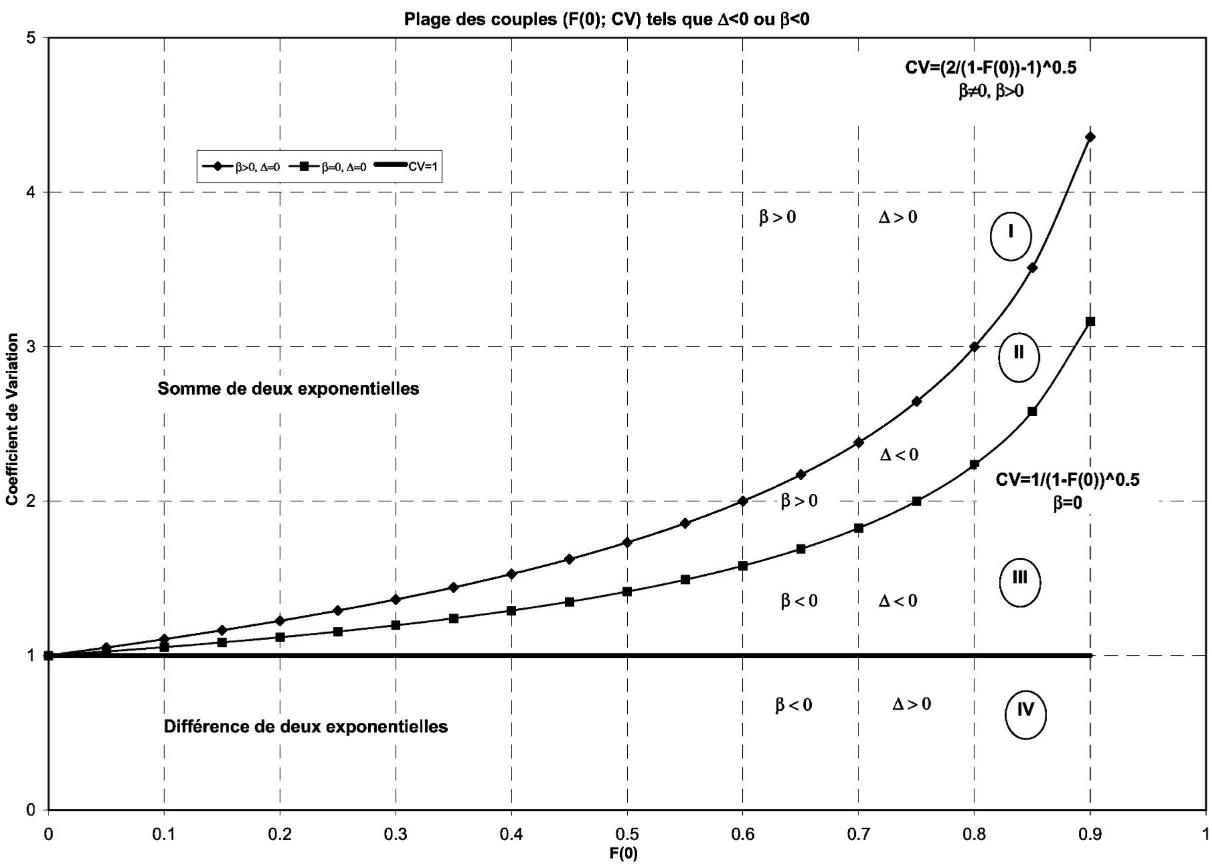


**Figure 3 : Présentation du domaine d'application de la loi somme de deux exponentielles.**

Le domaine d'application est divisé en quatre zones principales délimitées par des courbes :

- La courbe $CV = \sqrt{\dfrac{1+F(0)}{1-F(0)}}$ présente une loi simple exponentielle avec une fréquence de jours de pluies $\alpha + \beta$ et un gradex $a = c$.
- La courbe $CV = \dfrac{1}{\sqrt{1-F(0)}}$ donne une loi simple exponentielle avec $\beta = 0$ et de paramètres $\alpha$ et $a$.
- La zone I présente le domaine d'application de la loi somme de deux exponentielles, qu'on peut définir par l'inégalité $CV > \sqrt{\dfrac{1+F(0)}{1-F(0)}}$ . Dans ce cas, $\Delta$ est positif et $\beta$ varie entre 0 et 1, ce qui vérifie nos hypothèses de départ.
- Pour la zone II, $\Delta$ est négatif et $\beta$ varie entre 0 et 1. Elle est définie par la double inégalité : $\sqrt{\dfrac{1+F(0)}{1-F(0)}} > CV > \dfrac{1}{\sqrt{1-F(0)}}$ . Dans ce cas de figure, il n'y a pas de solution.
- De même pour la zone III, il n'y a pas de solution. Dans ce cas $\beta$ devient négatif (ce qui ne vérifie pas les hypothèses

de la loi : $0 < \beta < 1$). Elle correspond à la zone $\frac{1}{\sqrt{1 - F(0)}} > CV > 1$.

• $CV < 1$ définit la zone IV, qui représente une loi de différence de deux exponentielles ; il y a une solution malgré une valeur de β négative. Cette loi ne représente pas le type de nos distributions.

Pour les valeurs de β entre 0 et 1 et Δ positif, la méthode des moments fournit une estimation simple des paramètres $a$ et $c$ :

$$\begin{cases} a = \dfrac{\alpha\mu + \sqrt{\alpha\beta(K\theta - \mu^2)}}{\alpha\theta} \\ \\ c = \dfrac{\mu - \alpha a}{\beta} \end{cases} \quad \text{avec } K = \frac{\mu^2 + \sigma^2}{2}$$

$\mu$, σ sont respectivement la moyenne et l'écart-type théorique.

Sur 25 stations de la région Cévennes, Slimani (1985 [9]) a fait une étude comparative sur l'estimation du gradex par 3 méthodes différentes :

• Ajustement d'une loi de Gumbel par la méthode des moments (1) et par le maximum de vraisemblance (2) sur les maxima mensuels d'automne.

• Ajustement d'une loi somme de deux exponentielles sur les pluies consécutives de la saison d'automne (3).

Il a constaté que pour les pas de temps inférieurs à 6 heures, il y a une sous-estimation systématique du gradex calculé par somme de deux exponentielles par rapport aux deux autres méthodes. Pour mieux évaluer la sous-estimation du gradex, il a regardé l'influence de l'autocorrélation des pluies successives.

Une dizaine de stations a été l'objet de cette étude. Slimani n'a pu conclure si l'autocorrélation est la cause de cette sous-estimation du gradex ; il a donc revu l'estimation du paramètre α et propose une correction à faire, par laquelle il faut multiplier α :

$$C_k(t, F(0)) = \frac{1}{\left(1 - \log\left(\frac{t}{24}\right)\right)^{F(0)/2}}$$

Ce coefficient dépend du pas de temps choisi ($t$ inférieur ou égal à 24 heures) et de la fréquence des valeurs nulles ($F(0)$). L'estimation des autres paramètres ne change pas.

Un problème qui se pose lorsqu'on travaille sur des pluies journalières successives est l'autocorrélation des pluies. D. Duband (1967 [3]) a travaillé sur des données du Massif Central et des Cévennes, avec une quinzaine de stations de 20 à 50 ans d'observations. Il a trouvé que le coefficient d'autocorrélation varie selon la saison :

• D'octobre à avril, le coefficient d'autocorrélation pour le centre de la France est de l'ordre de 0,3 (saison d'hiver).

• De juin à septembre, le coefficient d'autocorrélation est plus faible, il est de l'ordre de 0,13 (saison d'été).

• Pour les mois de transition en mai et octobre, la durée d'observation était très courte et ne permet pas de faire un calcul robuste sur l'autocorrélation (mois de transition entre les 2 régimes).

En comparaison avec la loi Gamma Incomplète, D. Duband a trouvé que les deux lois sont pratiquement confondues pour $F$ compris entre 0,01 et 0,99.

### IV.3 Loi de Gumbel

C'est une loi à décroissance exponentielle, elle s'ajuste, pour des durées de 1 heure à quelques jours, sur les maxima de 10 ou 15 jours, maxima mensuels, maxima saisonniers ou maxima annuels. Son avantage est d'étirer l'échelle de distribution au-delà de la probabilité 0,90. C'est une fonction double exponentielle : $F(x) = \exp(-\exp(-(x - x_0)/a))$,
$a$ : est le gradex, qui signifie le gradient de l'exponentiel (paramètre d'échelle),
$x_0$ : est le paramètre de position (le mode),
$a$ et $x_0$ ont la même dimension que la variable,
$u = -\ln(-\ln(F(x))) = (x - x_0)/a$ est appelée variable réduite de Gumbel.

La distribution d'une variable de Gumbel sur un papier de Gumbel ($u$ en abscisse et $x$ en ordonnées) est une droite de pente $a$ (gradex) et d'ordonnée à l'origine le mode.

On obtient ces deux paramètres par la méthode des moments :

$$A = (\sqrt{(6)}/\pi) \times \sigma \text{ et } x_0 = \mu - 0{,}57721 \times a$$

0,57721 est la constante d'Euler ; la moyenne de la variable réduite de Gumbel $(x - x_0)/a$ lui est égale.
$\pi/\sqrt{(6)}$ est l'écart-type de la variable réduite de Gumbel.
$\mu$, σ sont respectivement la moyenne et l'écart-type théorique de la population.

### IV.4 Méthode pseudo-empirique

D. Duband propose de formuler le gradex $g$ en fonction des moments de premier et second ordre de la distribution empirique. Il donne une formulation générale de type :

$$g = k.(CV)^p.\sigma$$

où $CV = \sigma/\mu$ est le coefficient de variation, σ est l'écart-type théorique, μ moyenne théorique.

D. Duband a proposé en 1967 un estimateur déterminé à partir des paramètres de la loi Gamma Incomplète :

$$g_2 = s.CV^{0{,}75}$$

En 1985, D. Duband [4] a élaboré un autre estimateur construit à partir de l'utilisation de la loi somme de deux exponentielles :

$$g_1 = 0{,}78.s.CV^{0{,}93}$$

Ces estimateurs ont été calés sur des séries pluviométriques de la région des Cévennes et pour un pas de temps de l'ordre de la journée.

## V ■ STABILITÉ DE LA LOI SOMME DE DEUX EXPONENTIELLES

Dans une première partie, nous nous intéressons plutôt à la robustesse de la loi somme de deux exponentielles. Pour cela, nous avons estimé qu'une comparaison est nécessaire entre les moments d'ordre 3 et 4, théoriques et empiriques. Le principe de base de l'ajustement d'une loi somme de deux exponentielles est d'utiliser trois moments pour l'estimation des trois paramètres restants de la loi (la fréquence des valeurs nulles est estimée directement sur l'échantillon, le paramètre $\beta$ est fonction de $F(0)$ et il ne nous reste qu'à estimer $\alpha$, $a$ et $c$). Nous savons que les moments d'ordre 3 et 4 sont très sensibles à l'effet d'échantillonnage ; cela a conduit à introduire une estimation expérimentale de $\alpha$ pour éviter l'utilisation du moment d'ordre 3. Nous voulons tester le cas où l'on a utilisé le moment 3 pour l'estimation des paramètres et voir si cela a une grande influence sur notre ajustement. Cette étude a été faite par D. Duband (1985 [4]) sur les régions des Alpes, Pyrénées, Massif Central et le bassin du Tarn. Il a utilisé 13 stations pour les mois de septembre à novembre de 1965 à 1983 pour le bassin du Tarn, et 22 stations pour les autres régions, en utilisant des données journalières de 1951 à 1980 pour les mois de septembre et octobre et de novembre à mars. Donc 56 séries journalières des différentes saisons ont été utilisées dans le but de comparer les moments théoriques et empiriques. A partir de la loi, nous estimons les moments centrés théoriques d'ordre 3 et 4 par les formules suivantes :

$$\mu_3 = 6\alpha a^3 + 6\beta c^3 - 3\mu\sigma^2 - \mu^3$$

$$\mu_4 = 24\alpha a^4 + 24\beta c^4 - 4\mu\mu_3 - 6\mu^2\sigma^2 - \mu^4$$

$\alpha$, $\beta$, $a$ et $c$ sont les quatre paramètres de la loi,

$\mu$, $\sigma$, $\mu_3$ et $\mu_4$ : sont respectivement la moyenne, l'*écart*-type et les moments centrés d'ordre 3 et 4.

Les moments empiriques centrés calculés sur un échantillon des données journalières de $n$ valeurs sont donnés par les équations suivantes :

$$m_3' = \frac{n^2}{(n-1)(n-2)}\, m_3$$

$$m_4' = \frac{n^3 \cdot m_4 - 3(n-1)(2n-3)s^2}{(n-1)(n^2 - 3n - 2 + 3)}$$

$m_3'$ et $m_4'$ sont les moments centrés non biaisés. $m_3$, $m_4$ sont les moments centrés empiriques biaisés :

$$m_j = \frac{1}{n}\sum (x_i - \bar{x})^j \qquad j = 3{,}4$$

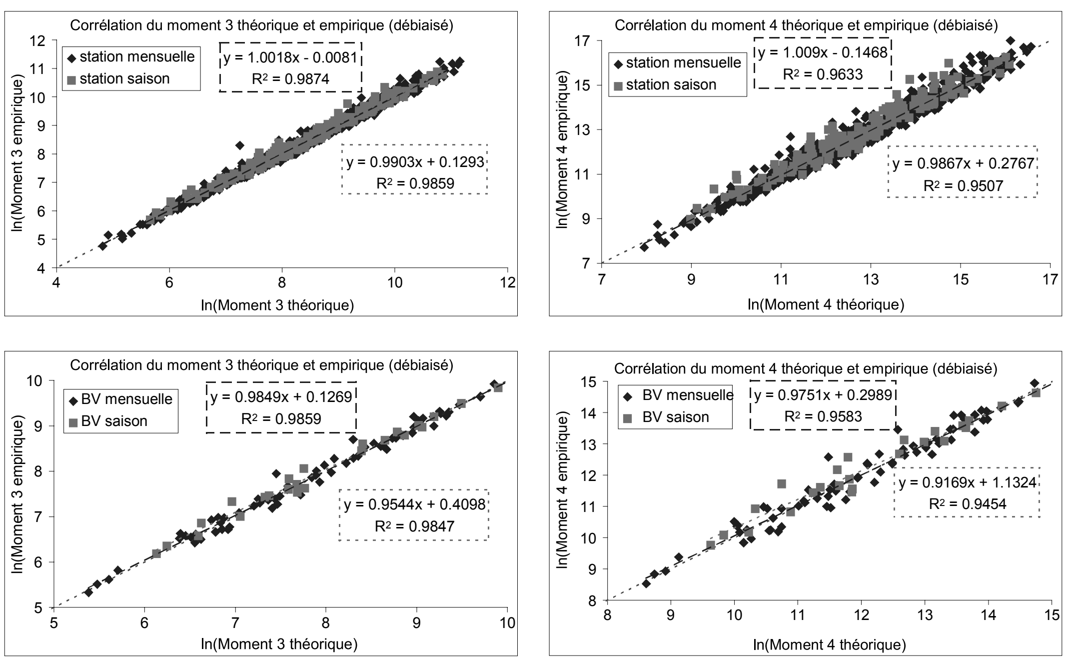


**Figure 4 : Corrélation entre les logarithmes des moments centrés théorique et empirique d'ordre 3 et 4. Moments calculés sur des données journalières (45 stations et 6 bassins versants).**

Dans le cadre de cette étude, nous avons comparé des moments des données ponctuelles et spatiales, mensuelles et saisonnières. Pour les données journalières de chaque mois, nous disposons de 540 séries ponctuelles (45 stations et 12 mois) et 72 séries des lames d'eau spatiales (6 bassins et 12 mois). Les données saisonnières présentent 180 échantillons des données ponctuelles (45 stations et 4 saisons) et 24 séries de lames d'eau spatiales (6 bassins et 4 saisons).

On a calculé les corrélations entre moments empiriques et théoriques sur les valeurs brutes et sur les valeurs transformées en logarithmes. Nous obtenons des corrélations très fortes pour toutes les séries. On notera des corrélations supérieures à 0,97 ($R^2 > 0{,}94$) pour des moments d'ordre 3 et 4. Tous ces résultats montrent la robustesse et la stabilité de la loi somme de deux exponentielles. Pour des pluies ponctuelles ou spatiales les corrélations sont bonnes ; de plus, les droites de corrélation sont très proches de la bissectrice *(fig. 4)*.

## VI ■ COMPARAISON DES AJUSTEMENTS

Un des buts de notre travail est de trouver la meilleure estimation possible des risques à partir des séries, lorsqu'on ne dispose pas de l'information journalière complète. En général, l'information pluviométrique la plus facile à obtenir est constituée des maxima annuels ou mensuels. Parfois le dépouillement des données n'est pas toujours complet : par exemple, il arrive que certaines stations ne sont dépouillées que sur les maxima hebdomadaires (on dispose ainsi de 90 pluviographes de la région Rhône-Alpes dépouillées en maxima hebdomadaires), la longueur de ces séries variant entre 10 et 30 ans. Pour pouvoir utiliser toute l'information disponible, il fallait trouver une loi qui s'adapte bien avec les séries des maxima hebdomadaires. Une étude de Pitiot *et al*. (1994, [7]) montre que l'ajustement des maxima hebdomadaires de 12 heures sur la station Valcivière (Massif Central) est comparable à celui de toutes les données successives de 12 heures. Dans le but d'approfondir ce travail, nous avons étudié les pluies journalières ponctuelles et spatiales des stations présentées dans le paragraphe 2. En première partie, nous avons effectué des ajustements des différentes lois sur les max hebdomadaires (reconstitués à partir des données journalières). Puis, nous avons gardé les lois qui ont donné des résultats proches et ne posent aucun problème d'ajustement (divergence, résultats différents,...). Nous avons utilisé 4 lois principales selon le type des données :

*Données journalières :* nous avons choisi pour les données journalières deux types d'estimations :

— la loi somme de deux exponentielles, dont le gradex estimé est une valeur de référence utilisée dans la comparaison avec les autres méthodes,

— une méthode pseudo-empirique, où une formulation du gradex est fixée de type $g = k.(CV)P.\sigma$. On donne deux formules : la première méthode est $g_1 = 0{,}78.s.CV^{0{,}93}$ appelée méthode M1, la deuxième utilise la formule $g_2 = s.CV^{0{,}75}$ dite méthode M2.

*Données max. des 15 jours et max. mensuels :* on a utilisé la loi de Gumbel, appliquée directement sur les séries.

*Données max. hebdomadaires :* c'est ce type de données qui pose problème, d'où l'intérêt du travail présenté. Dix estimations différentes du gradex sont comparées au gradex de référence, qu'on trouve par les trois lois suivantes :

1. *Loi de Gumbel censurée :* on a appliqué quatre seuils différents de 2.5, 5, 10 et 15 mm ce qui nous a semblé suffisant. Le gradex est déterminé par une corrélation simple entre les valeurs supérieures à un seuil fixe et les variables réduites de Gumbel correspondantes (empiriques).

2. *Loi simple exponentielle tronquée à zéro :* le gradex dans ce cas est la moyenne des valeurs strictement positives. Les deux méthodes pseudo-empiriques qui ont été appliquées sur les données journalières sont utilisées sur les valeurs maxima hebdomadaires strictement positives (la moyenne et l'*écart*-type sont calculés sur les valeurs supérieures à 0). Les deux méthodes sont appelées M3 et M4, elles correspondent respectivement aux formules de $g_1$ et $g_2$.

3. *Combinaison entre somme de deux exponentielles et simple exponentielle tronquée à zéro :* l'application de la loi somme de deux exponentielles ne vérifie pas toujours les conditions nécessaires (conditions sur $\beta$ et $\Delta$, *fig. 3*). On va donc prendre le résultat de la somme de deux exponentielles, quand elle vérifie les conditions, sinon on prend la loi simple exponentielle tronquée à 0. Les méthodes M5 et M6 qui correspondent respectivement aux formules de $g_1$ et $g_2$ seront appliquées selon la loi qui a été ajustée. C'est-à-dire, la moyenne et l'écart-type seront calculés sur toute la série si on a appliqué la somme de deux exponentielles, sinon ils seront calculés sur les valeurs supérieures à 0.

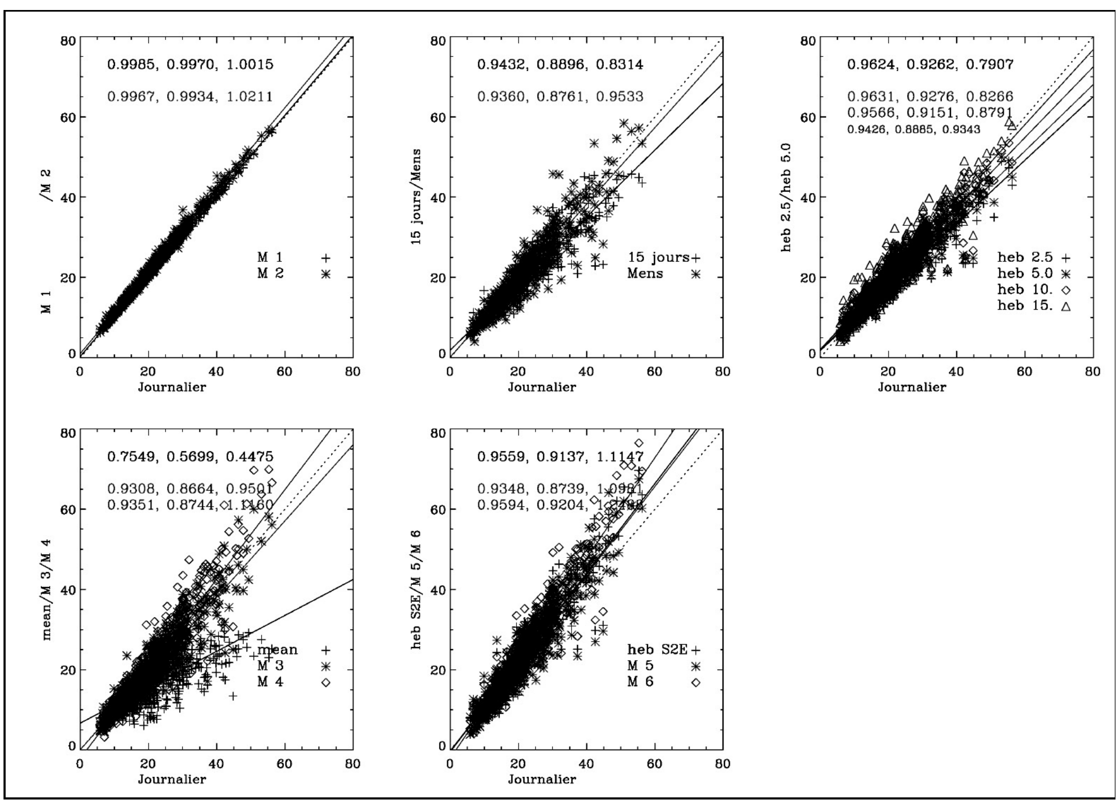


**Figure 5 : Comparaison entre le gradex journalier mensuel calculé par la loi somme de deux exponentielles et le gradex calculé par les autres méthodes.**

Toutes ces méthodes d'ajustement sont appliquées sur des données journalières, des données maxima mensuelles *(fig. 5)* ou saisonnières *(fig. 6)*, sur des données ponctuelles (45 stations) ou spatiales (pluies moyennes sur chacun des 6 bassins versants). Les résultats sont présentés dans les *tableaux 2* et *3*. On a divisé la comparaison en deux parties, la première concerne les ajustements sur les données journalières, max. des 15 jours et max. mensuels qui nécessitent des méthodes d'ajustement sans contrainte. La seconde concerne les méthodes d'ajustements sur les données max. hebdomadaires.

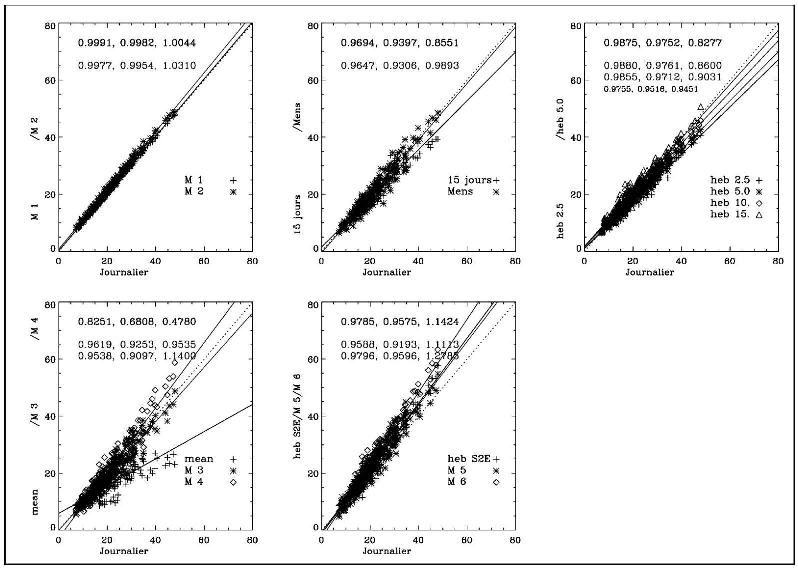


**Figure 6 : Comparaison entre le gradex journalier saisonnier calculé par la loi somme de deux exponentielles et le gradex calculé par les autres méthodes.**

On fait la comparaison des méthodes en fonction de 3 paramètres, qui sont :

— le Coefficient de détermination entre le gradex journalier calculé par la méthode somme de deux exponentielles et le gradex calculé par les différentes méthodes ;

— la Pente : représente la position de notre ajustement par rapport à la bissectrice ;

— la Distance : la distance correspond à la moyenne des écarts au carré entre le gradex journalier calculé par la méthode somme de deux exponentielles et le gradex calculé par les différentes méthodes (distance par rapport à la bissectrice).

*Tableau 2 : Comparaison entre le gradex journalier calculé par la loi somme de deux exponentielles et les formules pseudo-empiriques (M1 et M2), ainsi que le gradex des 15 jours et mensuels ajusté par une loi de Gumbel (45 stations).*

| | | **M1 ($g_1$)** | **M2 ($g_2$)** | **15 jours** | **Mens.** |
|---|---|---|---|---|---|
| | $R^2$ | 0,998 | 0,995 | 0,940 | 0,931 |
| Saison | Pente | 1,004 | 1,031 | 0,855 | 0,989 |
| | Distance | 0,2 | 2,4 | 7,6 | 5,9 |
| | $R^2$ | 0,997 | 0,993 | 0,890 | 0,876 |
| Mensuelle | Pente | 1,002 | 1,021 | 0,831 | 0,953 |
| | Distance | 1,1 | 3,6 | 13,5 | 12,7 |

On remarque que les méthodes pseudo-empiriques M1 et M2 donnent les meilleurs résultats. La corrélation est pratiquement égale à 1, la pente est proche de 1 et on constate une légère différence en distance entre M1 et M2. Ce résultat ne sera appliqué que pour les séries journalières ; il ne sera intéressant que dans le cas où l'on n'arriverait pas à ajuster la somme de deux exponentielles (dans la limite d'application de la loi *fig. 3*).

Les méthodes, simple exponentielle, M3, M4, M5, M6 et la combinaison SE-S2E (simple exponentielle et somme de deux exponentielles) donnent les plus mauvais résultats (surestimation et une grande distance par rapport la bissectrice). Le résultat de la loi de Gumbel censurée dépend du seuil appliqué. Si on prend un seuil faible, on introduit les faibles valeurs dans le calcul, ce qui occasionne une *sous*-estimation du gradex (pente 0,83 et 0,79). Le seuil 10 mm présente, dans notre cas, le meilleur ajustement. En effet, il représente une distance et une pente optimales du calcul. On ne peut sans doute pas généraliser ce résultat, car le seuil dépend de la série elle-même ; si on a un nombre faible de maxima qui dépassent le seuil, le calcul sera biaisé par les fortes valeurs, d'où une surestimation du gradex.

On trouve que l'ajustement des données maxima journalières de 15 jours est comparable à celui des données seuillées à 10 mm. L'estimation du gradex à partir des données maxima journalières mensuelles est moins robuste ; la dispersion d'échantillonnage est plus importante, mais en terme de distance, le résultat n'est pas très mauvais (résultat proche de celui des données hebdomadaires avec un seuil de 2,5).

**Tableau 3 : Comparaison entre le gradex journalier ajusté par la loi somme de deux exponentielles et le gradex hebdomadaire calculé par les différents ajustements (45 stations).**

| | | Heb. Seuil 2.5 | Heb. Seuil 5.0 | Heb. Seuil 10.0 | Heb. Seuil 15.. | Simple expo | M3 ($g_1$) | M4 ($g_2$) | SE-S2E | M5 ($g_1$) | M6 ($g_2$) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Saison | $R^2$ | 0,975 | 0,976 | 0,971 | 0,952 | 0,681 | 0,925 | 0,910 | 0,957 | 0,919 | 0,960 |
| | Pente | 0,828 | 0,860 | 0,903 | 0,945 | 0,478 | 0,953 | 1,140 | 1,142 | 1,111 | 1,278 |
| | Distance | 10,9 | 6,0 | 2,7 | 4,0 | 55,6 | 6,6 | 11,0 | 9,1 | 12,1 | 21,1 |
| Mensuelle | $R^2$ | 0,926 | 0,928 | 0,915 | 0,888 | 0,570 | 0,866 | 0,874 | 0,914 | 0,874 | 0,920 |
| | Pente | 0,791 | 0,827 | 0,879 | 0,934 | 0,447 | 0,950 | 1,115 | 1,114 | 1,098 | 1,249 |
| | Distance | 15,8 | 11,0 | 8,5 | 11,1 | 65,5 | 14,0 | 17,3 | 16,1 | 20,4 | 27,9 |

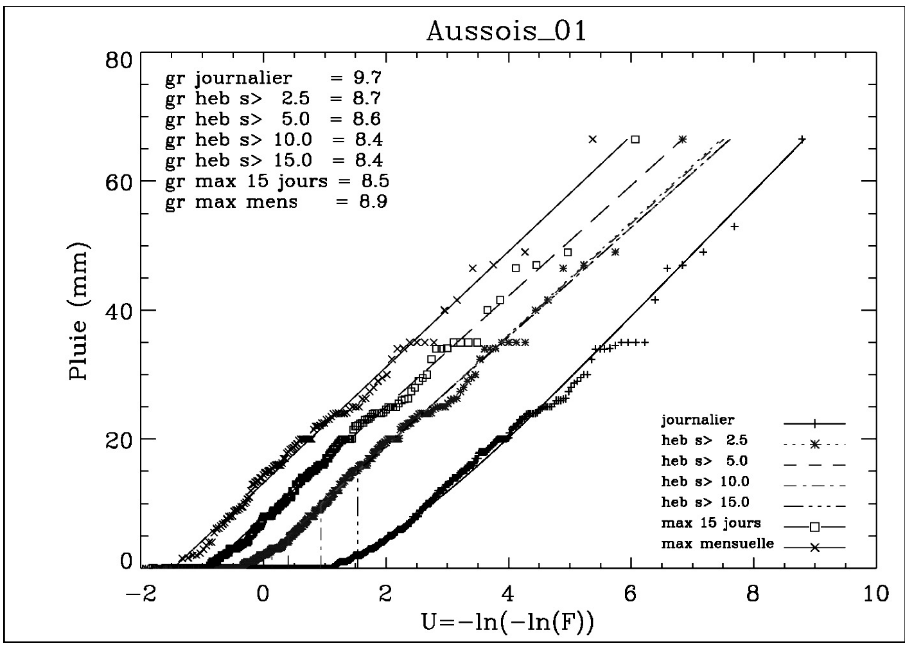


**Figure 7 : Comparaison des différentes méthodes d'ajustements. Station : Aussois, saison d'automne (septembre, octobre et novembre)**

## VII ■ CONCLUSION

La problématique de notre travail était de chercher une méthode d'estimation des risques pour des séries dont l'ajustement par des lois simples n'est pas possible. Rappelons également, que comme le but ultime de cette étude est la prédétermination des crues, notamment par la méthode du Gradex, nous avons privilégié l'estimation du gradex des pluies (pente de la distribution sur un papier de Gumbel) qui utilise essentiellement les distributions exponentielles. Afin de trouver une solution à ce problème, nous avons sélectionné 6 bassins versants et 45 stations pluviométriques. Ensuite, nous avons déterminé les risques ponctuels et spatiaux (mensuels et saisonniers) par différentes méthodes d'estimation et pour chaque type de donnée, soit données journalières, max hebdomadaire et max mensuel. Nous avons utilisé la loi somme de deux exponentielles pour les données journalières (cette loi est utilisée pour les pas de temps compris entre 6 et 24 heures) ; cependant pour les pas de temps inférieurs à 6 heures, il est préférable d'introduire un coefficient de correction dépendant du pas de temps et de $F(0)$, (Slimani 1984 [9]). Si on ne dispose que des max. hebdomadaires, nous avons appliqué une loi de Gumbel censurée pour augmenter la taille de l'échantillon. Le but était d'essayer d'utiliser le maximum d'information disponible (afin de diminuer l'effet d'échantillonnage). Pour les max. des 15 jours, on préfère l'utiliser seulement quand on dispose d'une longue série ; pour les courtes séries, cela nous paraît un peu dangereux de l'appliquer (du fait qu'il y a un pourcentage de valeurs nulles qui biaise l'estimation). Pour les autres séries, il n'y a pas de choix possible, pour les données max. mensuels ou annuels, on ajuste une loi de Gumbel simple (ou d'autres lois adaptées à ce type de données).

De ce travail, on peut conclure que les formulations pseudo-empiriques appliquées aux données journalières donnent les résultats les plus robustes. Cependant, leurs application demande toute l'information journalière, ce qui est une nécessité garantissant la précision et la stabilité du gradex, puisque les données sont généralement fournies sous forme des max hebdomadaires (dans le meilleur des cas), max mensuel ou max annuel. Mais, on peut considérer cette méthode comme une sécurité pour l'application à d'autres lois d'ajustement sur les données journalières.

La loi de Gumbel seuillée est une bonne solution pour estimer les risques extrêmes à partir des données maxima hebdomadaires. Il reste néanmoins que le choix du seuil est un peu délicat. Il faut prendre plusieurs seuils et comparer les ajustements sur le même graphique, puis il faut choisir l'ajustement qui correspond au mieux à la distribution.

Nous remarquons que les ajustements des séries saisonnières fournissent un résultat plus robuste. On le mesure par la dispersion du nuage de corrélation, la pente de la droite de corrélation et la diminution de la distance par rapport à la bissectrice.

## REMERCIEMENTS

Cette étude n'a pu être effectuée que grâce à l'appui du programme franco italien INTERREG Inondations gérée scientifiquement coté français par le pôle Grenoblois d'Etude et de Recherches sur les Risques Naturels.

## BIBLIOGRAPHIE

[1] Bouvard (M.), Garros Berthet *et al.*, 1994. — Les crues de projet des barrages : Méthode du Gradex. *Bulletin du comité français des grands barrages*, 18e Congrès Durban CIGB/ICOLD N° 2, novembre.

[2] Duband (D.), 1967. — Composition des lois de probabilité de précipitations journalières. Symposium International d'Hydrologie USA, Fort Collins, septembre.

[3] Duband (D.), 1967. — Composition des lois de probabilité de précipitations journalières. *Société Hydrotechnique de France*, Commission pour l'étude des débits, Réunion du 9 mai.

[4] Duband (D.), 1985. — Distribution empirique des précipitations journalières. Calcul du gradex. Division Technique Générale. Centre Hydrologique Alpes, 1er août.

[5] Gouy (D.), 1994. — Contribution à la méthodologie du gradex. Particularités des distributions empiriques telles que $Cv < 1$ et $F(0) \approx 0$. Développement de nouveaux estimateurs du gradex, Application au groupement des 49 stations Cévennes Vivarais. Service Ressource en Eau, EDF-DTG, 22 juin.

[6] Lebel (T.), 1984. — Moyenne spatiale de la pluie sur un bassin versant : Estimation optimale, génération stochastique et gradex des valeurs extrêmes. Thèse Docteur, Ingénieur, Université Scientifique et Médicale, INPG, Grenoble, 20 février, p. 326.

[7] Pitiot (M.) & Rodriguez (Y.), 1994. — Comparaison d'un gradex en 12 h à un gradex calculé sur tous les maximums hebdomadaires sur une station du Massif Central (Valcivière). Service Ressource en Eau, EDF-DTG, 17 janvier.

[8] Rider (P.R.), 1960. — The method of moments applied to a mixture of Two exponential distribution. Aeronautical Research Laboratories, Wright-Patterson Air Force Base, 12 August 1962, pp. 143-147.

[9] Slimani (M.), 1985. — Etude des pluies de fréquence rare à faibles pas de temps sur la région Cévennes Vivarais : Estimation, relations avec le relief et cartographie synthétique. Thèse Docteur Ingénieur, LTHE-INPG, Grenoble, 23 septembre.

[10] Smith (R.E.) et Schreiber (H.A.), 1974. — « Point Process of Seasonal Thunderstorm Rainfall ». Water Resources Research, June, Vol. 10, n° 3.

[11] Projet Interreg II France-Italie (1994-99). — Connexion des réseaux de données et mise en commun des connaissances et des expériences pour la gestion des risques d'inondations en région alpines. Région Rhône Alpes et Région Piémont. Rapport + 188 pp.