

# À propos de la distribution statistique des cumuls pluviométriques annuels. Faut-il en finir avec la normalité? Statistical distribution of total annual rainfall. Is the Gaussian distribution inadequate?

H. Benjoudi and P. Hubert

Volume 11, Number 4, 1998

URI: <https://id.erudit.org/iderudit/705324ar>

DOI: <https://doi.org/10.7202/705324ar>

[See table of contents](#)

Publisher(s)

Université du Québec - INRS-Eau, Terre et Environnement (INRS-ETE)

ISSN

0992-7158 (print)

1718-8598 (digital)

[Explore this journal](#)

Cite this article

Benjoudi, H. & Hubert, P. (1998). À propos de la distribution statistique des cumuls pluviométriques annuels. Faut-il en finir avec la normalité? *Revue des sciences de l'eau / Journal of Water Science*, 11(4), 617–630.  
<https://doi.org/10.7202/705324ar>

Article abstract

Up to now, annual rainfall accumulation have generally been modelled according to the Laplace-Gauss probability distribution. After a brief survey of the arguments for using this distribution to describe annual rainfall, which are mainly to consider annual rainfall accumulation as the sum of many independent individual rainfall events of similar magnitude identically distributed, we question its capacity to take into account the various characteristics of the rainfall events, in particular their magnitude, their number and their possible correlation or persistence. We have studied an alternative model based on the multifractal theory, well suited to model phenomena where matter and/or energy concentrate on a more and more sparse domain as the observation scale is decreasing. Some elements of the multifractal theory are briefly described. The main feature of the probabilistic model based on the assumption of a multifractal behavior is that for large enough accumulations, the probability distribution tail would have at any time scale an algebraic rather than an exponential behavior as it is the case for the Laplace-Gauss distribution. It is of value to note that such an algebraic behavior corresponds to a probability of occurrence decreasing much more slowly than in the case of an exponential one. It is also important to note that, unlike exponential distribution laws, all statistical moments of algebraic distribution laws are not defined, those of order greater than the exponent of the algebraic law diverging. This fact is of main importance in relation with sampling, as we cannot a priori be sure of the convergence of all statistical estimators and that the convergence of these estimators is likely to much more slow than that of exponential law estimators.

The application part of our study concerned 87 annual rainfall series spanning from 44 to 266 years with a mean of 116 years, gathered within the UNESCO FRIEND-AMHY project (Flow Regimes from International Experimental and Network Data - Alpine and Mediterranean Hydrology). For each series we have drawn on a log-log diagram the curve of the empirical probability of exceeding a given rainfall value (derived by the Weibull formula) versus the rainfall value. On such a diagram an algebraic behavior should be represented by a straight line the slope of which is the exponent characterizing the algebraic law tail. Qualitatively the results tend to argue in favor of an algebraic rather than an exponential behavior of the probability distribution of high annual rainfall (empirical return periods of more than 20 years) but the exponent values are quite scattered, ranging from about 2 for the lowest values to more than 10. We then have concentrated our study on the basis of the 71 stations with series more than 90 years long. A diagram of the exponent value versus the corresponding series length suggest that there may be a slow convergence of exponent values towards a common value close to 3.8 as the length of the series increases (This observation should be related to a possible divergence of high order moments which is the cause of poor estimation from small size samples). This exponent would then be universal, related to the physical processes from which rainfall originates rather than to a geographical location. You have to recall here that the multifractal framework we used has been primarily designed to take into account the scaling symmetries of the Navier and Stokes equations. It is likely that the unknown partial differential equations governing rainfall processes share some properties with those governing atmospheric turbulence and it would not be too surprising that we can catch this way some physical feature of rainfall process.

These preliminary results, if confirmed by further researches, would have considerable theoretical and practical consequences. The algebraic behavior of heavy rainfall distribution, which is supposed, according to the multifractal theory, to arise at all time scales with the same exponent, should lead to reject the numerous empirical and ad-hoc distribution functions which are at present used mainly for practical purposes. These probability distribution functions have poor theoretical and/or physical basis and quite all of them exhibit more or less an exponential behavior. It is likely that a huge amount of efforts has been devoted for decades to the sophistication of fitting methods, using low and mean variate realizations, while large and extreme values were often understated and underused as extreme values are generally said to belong to another population or, equivalently, to be generated by an other random process than normal ones (outliers...). The new generation of distribution functions to design would be less empirical: it should take explicitly into account time scale invariance and should be able to reconcile into an unique model all the realizations of the rainfall accumulation variate, whatever their magnitudes are.

In order to size up the difference between algebraic and exponential statistical models one can note, for example, that the return period of the annual rainfall of four cities (Padova, Marseilles, Rome and Gibraltar) which were estimated at 1,000 years with the fitting of a Gaussian distribution, could be as low as 60 to 100 years with this new algebraic model. The return period is so divided by a factor of 10! All the same annual rainfall accumulations are said to have a rather "soft" behavior, the summation of numerous individual events being supposed to smooth their more "wild" behavior. This is probably not true and one can find in annual and even pluriannual rainfall accumulations the trace of extreme individual events. Whatever the time scale under consideration, it is easy to imagine the consequences of such a dramatic modification of return periods on engineering design. Such a revisiting of estimated return periods should be extended to hydrological events. As an example, the recent flooding of the river Oder in Germany, Poland, and Czech Republic (July 1997) for which preliminary information suggests a return period greater than 10,000 years (Gazowsky, personal communication), might not be so extreme if our new concepts are shown to be valid.

# À propos de la distribution statistique des cumuls pluviométriques annuels Faut-il en finir avec la normalité ?

Statistical distribution of total annual rainfall  
Is the Gaussian distribution inadequate ?

H. BENDJOUDI<sup>1</sup> \*, P. HUBERT<sup>2</sup>

---

Reçu le 17 novembre 1997, accepté le 4 mai 1998\*\*.

## SUMMARY

*Up to now, annual rainfall accumulation have generally been modelled according to the Laplace-Gauss probability distribution. After a brief survey of the arguments for using this distribution to describe annual rainfall, which are mainly to consider annual rainfall accumulation as the sum of many independent individual rainfall events of similar magnitude identically distributed, we question its capacity to take into account the various characteristics of the rainfall events, in particular their magnitude, their number and their possible correlation or persistence. We have studied an alternative model based on the multifractal theory, well suited to model phenomena where matter and/or energy concentrate on a more and more sparse domain as the observation scale is decreasing. Some elements of the multifractal theory are briefly described. The main feature of the probabilistic model based on the assumption of a multifractal behavior is that for large enough accumulations, the probability distribution tail would have at any time scale an algebraic rather than an exponential behavior as it is the case for the Laplace-Gauss distribution. It is of value to note that such an algebraic behavior corresponds to a probability of occurrence decreasing much more slowly than in the case of an exponential one. It is also important to note that, unlike exponential distribution laws, all statistical moments of algebraic distribution laws are not defined, those of order greater than the exponent of the algebraic law diverging. This fact is of main importance in relation with sampling, as we cannot a priori be sure of the convergence of all statistical estimators and that the convergence of these estimators is likely to much more slow than that of exponential law estimators.*

- 
1. UMR 7619, Laboratoire de Géologie Appliquée, Université Pierre et Marie Curie, case 123, 4, place Jussieu, 75252 Paris Cedex 05, France.  
bdj@biogeodis.jussieu.fr
  2. UMR 7619, Centre d'Informatique Géologique, École des Mines de Paris, 35, rue St Honoré, 77305 Fontainebleau Cedex, France.  
hubert@cig.ensmp.fr

\* Correspondance.

\*\* Les commentaires seront reçus jusqu'au 31 juillet 1999.

The application part of our study concerned 87 annual rainfall series spanning from 44 to 266 years with a mean of 116 years, gathered within the UNESCO FRIEND-AMHY project (Flow Regimes from International Experimental and Network Data - Alpine and Mediterranean Hydrology). For each series we have drawn on a log-log diagram the curve of the empirical probability of exceeding a given rainfall value (derived by the Weibull formula) versus the rainfall value. On such a diagram an algebraic behavior should be represented by a straight line the slope of which is the exponent characterizing the algebraic law tail... Qualitatively the results tend to argue in favor of an algebraic rather than an exponential behavior of the probability distribution of high annual rainfall (empirical return periods of more than 20 years) but the exponent values are quite scattered, ranging from about 2 for the lowest values to more than 10. We then have concentrated our study on the basis of the 71 stations with series more than 90 years long. A diagram of the exponent value versus the corresponding series length suggest that there may be a slow convergence of exponent values towards a common value close to 3.8 as the length of the series increases (this observation should be related to a possible divergence of high order moments which is the cause of poor estimation from small size samples). This exponent would then be universal, related to the physical processes from which rainfall originates rather than to a geographical location. You have to recall here that the multifractal framework we used has been primarily designed to take into account the scaling symmetries of the Navier and Stokes equations. It is likely that the unknown partial differential equations governing rainfall processes share some properties with those governing atmospheric turbulence and it would not be too surprising that we can catch this way some physical feature of rainfall process.

These preliminary results, if confirmed by further researches, would have considerable theoretical and practical consequences. The algebraic behavior of heavy rainfall distribution, which is supposed, according to the multifractal theory, to arise at all time scales with the same exponent, should lead to reject the numerous empirical and *ad hoc* distribution functions which are at present used mainly for practical purposes. These probability distribution functions have poor theoretical and/or physical basis and quite all of them exhibit more or less an exponential behavior. It is likely that a huge amount of efforts has been devoted for decades to the sophistication of fitting methods, using low and mean variate realizations, while large and extreme values were often understated and underused as extreme values are generally said to belong to another population or, equivalently, to be generated by an other random process than normal ones (outliers...). The new generation of distribution functions to design would be less empirical: it should take explicitly into account time scale invariance and should be able to reconcile into a unique model all the realizations of the rainfall accumulation variate, whatever their magnitudes are.

In order to size up the difference between algebraic and exponential statistical models one can note, for example, that the return period of the annual rainfall of four cities (Padova, Marseilles, Rome and Gibraltar) which were estimated at 1,000 years with the fitting of a Gaussian distribution, could be as low as 60 to 100 years with this new algebraic model. The return period is so divided by a factor of 10! All the same annual rainfall accumulations are said to have a rather "soft" behavior, the summation of numerous individual events being supposed to smooth their more "wild" behavior. This is probably not true and one can find in annual and even pluriannual rainfall accumulations the trace of extreme individual events. Whatever the time scale under consideration, it is easy to imagine the consequences of such a dramatic modification of return periods on engineering design. Such a revisiting of estimated return periods should be extended to hydrological events. As an example, the recent flooding of the river Oder in Germany, Poland, and Czech Republic (July 1997) for which preliminary information suggests a return period greater than 10,000

years (Gazowsky, personal communication), might not be so extreme if our new concepts are shown to be valid.

**Key-words:** *annual rainfalls, extreme events, multifractals, normal process, algebraic model.*

## RÉSUMÉ

Il est communément admis que la distribution statistique des précipitations cumulées annuelles suit une loi de Laplace-Gauss. Les écarts entre cette loi et les distributions empiriques sont cependant un fait d'expérience : au-delà d'une probabilité au non dépassement correspondant à une période de retour d'une vingtaine d'années et pour les valeurs les plus fortes de pluie, l'ajustement n'est plus acceptable. Ce décrochage par rapport à la loi normale est mieux mis en évidence par l'étude des longues séries pluviométriques, plus riches en événements extrêmes. Pour étudier le comportement statistique de ces derniers, il est fait appel à un formalisme multifractal qui permet de mettre en évidence que, contrairement à ce qui est généralement admis, la décroissance de la probabilité au dépassement est de nature hyperbolique plutôt qu'exponentielle. Les probabilités des événements catastrophiques sont donc plus importantes que l'on ne le croyait jusqu'ici, ce qui peut avoir des conséquences particulièrement importantes. Cette approche appliquée à un ensemble de séries pluviométriques de longue durée permet de cerner le paramètre caractérisant la décroissance de la probabilité au dépassement. Les résultats obtenus jusqu'ici laissent à penser que ce paramètre pourrait être universel.

**Mots clés :** *précipitations, événements extrêmes, loi normale, loi hyperbolique, multifractals, invariance d'échelle.*

## 1 – INTRODUCTION : LES LIMITES DE LA LOI NORMALE

La distribution statistique des variables hydrologiques a fait et continue de faire l'objet d'une littérature tellement abondante que nous ne nous risquons pas ici à vouloir en faire une revue exhaustive. Ce sujet fait toujours l'objet de nombreuses controverses et polémiques souvent virulentes. Il existe pourtant dans ce domaine un îlot de large consensus, celui des précipitations annuelles dont la plupart des auteurs s'accordent à dire qu'elles suivent une loi normale (Laplace-Gauss)

C'est ce qu'on peut lire dans REMENIERAS (1976) qui se réfère en particulier à un article d'ARLERY (1950) relatif à une série pluviométrique mesurée à Paris et qui conclut que « la série des modules pluviométriques annuels de la période de 150 ans comprise entre 1800 et 1949 peut être raisonnablement bien représentée par une loi de Laplace-Gauss ». C'est aussi ce qu'affirme YEVJEVICH (1972) qui s'appuie sur une étude de MARKOVIC (1965) : « It was found that the normal function could fit relatively well empirical distributions of annual precipitation of a large number of 1614 precipitation gauging stations in western North America ». On pourrait multiplier les exemples sous toutes les latitudes.

Une citation tirée d'un manuel technique de l'Agence de l'eau Seine-Normandie (AFBSN, 1972) donne les fondements « théoriques » de ce consensus :

« ..., les précipitations annuelles peuvent être considérées comme la somme des précipitations journalières, qui sont des variables nombreuses et très peu dépendantes entre elles et qui peuvent suivre une même loi de probabilité. Dans ces conditions, la loi de Laplace-Gauss décrit correctement la répartition des précipitations annuelles autour de leur moyenne ». ROCHE (1963), TUCCI (1993) et de nombreux autres auteurs donnent une justification analogue à l'usage très général de la loi normale pour modéliser les précipitations annuelles.

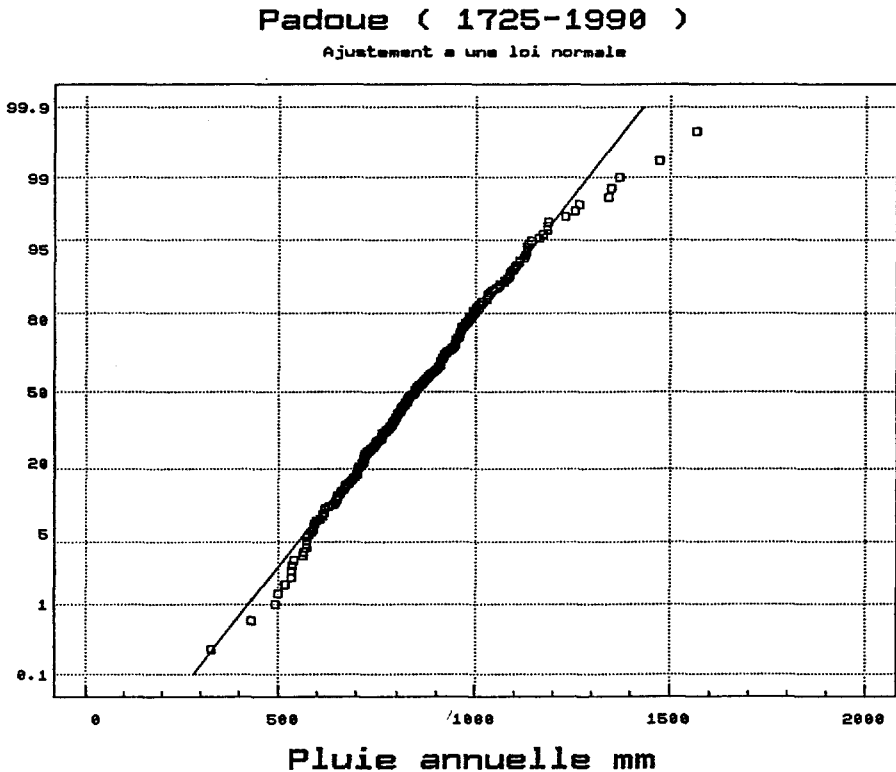
C'est en fait le théorème central limite (Laplace-Liapounoff) qui est invoqué ici. Ce théorème démontre que toute somme de  $n$  variables aléatoires indépendantes, de variance finie et identiquement distribuées, est asymptotiquement distribuée normalement (HAAN, 1977 ; KOTTEGODA, 1980). Ce théorème peut être généralisé à une somme de  $n$  variables aléatoires quelconques pour autant « qu'aucune variable ou groupe de variables n'occupe dans la somme une place prépondérante par rapport aux autres, ou encore, que toutes les variables initiales soient du même ordre de grandeur » (DAGNELIE, 1970).

C'est ici que le bât blesse et que l'on sollicite sans doute exagérément le théorème central limite. Tout d'abord, est-il légitime de considérer que les quelques dizaines, au maximum, d'averses, de pluies journalières ou d'épisodes pluvieux observables au cours d'une année, constituent « un grand nombre » d'événements ? Ensuite, ces événements sont-ils tous de même nature, ont-ils tous la même loi de probabilité et sont-ils indépendants ? Mais, surtout, il est difficile de soutenir que ces événements sont du même ordre de grandeur alors qu'ils varient entre quelques dixièmes de millimètres et plusieurs centaines de millimètres, soit un rapport de l'ordre de 1 à 1 000. On sait d'ailleurs le poids que peut avoir une simple averse dans un bilan pluviométrique annuel.

Les écarts entre la loi normale et les distributions empiriques sont cependant un fait d'expérience qui conduit certains analystes à rechercher le secours d'autres lois (Galton par exemple) pour retrouver un peu d'asymétrie. D'autres se contenteront, plus radicalement, d'écarter des ajustements les horsains (valeurs « aberrantes » ou « outliers ») réputés relever d'une autre logique quand ils ne sont pas tout simplement considérés comme des erreurs. Cela est pour le moins choquant car les modèles statistiques ont en général des objectifs pragmatiques comme, par exemple, l'estimation d'événements de récurrence rare. À quoi bon caler ces modèles si les événements extrêmes qui sont parmi les plus intéressants dans cette perspective, en sont exclus ?

Les longues séries sont bien entendu celles qui permettent le mieux de mettre à l'épreuve les modèles, car on y trouvera, en principe, davantage d'événements rares dont la probabilité d'apparition est faible. Nous avons reporté sur la figure 1, les probabilités empiriques au non-dépassement d'un seuil pour les valeurs de la série pluviométrique de Padoue (Padova) en Italie. Les probabilités empiriques ont été estimées par la formule de Weibull ( $p_k = k/(n + 1)$ ),  $k$  rang et  $n$  taille de l'échantillon), le choix d'une autre formule (Cunnane, Hazen ou Chegodayev par exemple) ne changeant pas fondamentalement les résultats du fait de l'échelle (papier normal) du graphique. Avec une durée de 266 ans cette série, compilée et critiquée par CAMUFFO (1984), est l'une des plus longues dont nous disposons. Sur le même graphique nous avons reporté la probabilité au non dépassement résultant de l'ajustement à une loi normale des données empiriques (Droite de Henry).

L'ajustement est satisfaisant... sauf pour les valeurs extrêmes. Pour les valeurs les plus fortes, le décrochage de la loi normale se produit pour une proba-



**Figure 1** Ajustement à une loi normale de la série pluviométrique de Padoue (Italie).  
*Normal fitting of the rainfall serie of Padova (Italy).*

bilité au non-dépassement de l'ordre de 0,95 (soit une période de retour d'une vingtaine d'années). Au-delà de cette valeur, les points représentatifs se situent systématiquement et très nettement en dessous de la droite représentant l'ajustement normal, mettant en défaut l'adéquation de cette loi, précisément dans la gamme de valeurs qui nous semble revêtir un très grand intérêt opérationnel et scientifique.

### Approche multifractale

Pour tenter de trouver une formulation probabiliste où les valeurs extrêmes des précipitations annuelles trouveraient réellement leur place, nous sommes partis d'une heuristique inspirée par la théorie des multifractals dont nous dirons d'abord quelques mots.

La géométrie fractale (MANDELBROT, 1982) a renouvelé les approches de nombreux domaines des sciences de la nature en y introduisant explicitement la notion d'échelle. Ainsi le caractère fractal de l'occurrence des pluies du Sahel africain a été mis en évidence par HUBERT et CARBONNEL (1988, 1989). Cependant, l'occurrence est définie par référence à un seuil d'intensité de la précipitation, et la dimension fractale de l'ensemble observé décroît lorsque le seuil de

référence augmente (HUBERT *et al*, 1995). En effet, les divers champs (piézométrie, pluviométrie, etc.) qui intéressent l'hydrologue ne peuvent pas être réduits à des êtres géométriques caractérisés par la présence ou l'absence d'un phénomène, ce qui conduirait à négliger la distribution dans l'espace et/ou le temps de l'intensité des phénomènes étudiés. L'approche multifractale est ambitieuse car elle vise à une prise en compte simultanée des échelles (de temps et/ou d'espace) et des intensités, et se révèle utile dans l'étude des phénomènes atmosphériques en général et l'étude des pluies en particulier, par exemple pour l'interprétation des courbes intensité-durée-fréquence (BENDJOUDI *et al*, 1997).

Très généralement, les multifractals pourront être associés à des processus en cascade concentrant la matière et/ou l'énergie dans des domaines spatio-temporels de plus en plus ténus. Les modèles de multifractals universels dont nous évoquerons ici quelques conséquences ont d'abord été conçus comme des modèles phénoménologiques de la turbulence, spécifiquement construits pour reproduire les principales propriétés des équations de Navier-Stokes (invariance d'échelle, conservation d'un flux et dynamique locale). Ces propriétés devraient se retrouver dans les équations, encore inconnues, régissant les phénomènes atmosphériques et en particulier les précipitations (SCHERTZER et LOVEJOY, 1987a).

Schématiquement, dans un processus de cascade multiplicative, à chaque niveau, chaque tourbillon se divise en sous tourbillons en redistribuant tout ou partie de son flux ou de son énergie à ces derniers. Un facteur aléatoire détermine la fraction du flux transmise d'un gros tourbillon à un plus petit. À une échelle d'observation  $\lambda$  donnée ( $\lambda$  est le rapport entre l'échelle externe et l'échelle d'observation) la probabilité pour que le flux  $\varepsilon_\lambda$  dépasse un seuil donné  $\lambda^\gamma$  est donnée par (SCHERTZER et LOVEJOY, 1991) :

$$\text{Prob}[\varepsilon_\lambda > \lambda^\gamma] = A_\lambda \lambda^{-C(\gamma)}$$

où  $\gamma$  est un ordre de singularité et  $C(\gamma)$  une fonction de codimension caractérisant la probabilité d'occurrence des singularités d'ordre supérieur à  $\gamma$ .  $A_\lambda$  étant un préfacteur variant peu avec  $\lambda$ .

De façon équivalente, cette équation peut être réécrite en termes de moments :

$$\langle \varepsilon_\lambda^q \rangle = B_\lambda \lambda^{K(q)}$$

où  $q$  est un ordre de moment,  $K(q)$  une fonction de structure et  $B_\lambda$  un préfacteur variant peu avec  $\lambda$ .

Dans l'étude des cascades multifractales, on distingue soigneusement leurs propriétés nues et leurs propriétés habillées. Les propriétés nues correspondent à la construction d'une cascade jusqu'à une certaine échelle d'homogénéité liée au phénomène. Au contraire, les propriétés habillées sont obtenues par intégration à l'échelle d'observation d'une cascade totalement développée. Par définition, les observations pluviométriques horaires, journalières et bien sûr annuelles, sont des propriétés habillées puisqu'elles résultent de l'intégration (au cours du temps) de processus se développant jusqu'à une certaine échelle temporelle beaucoup plus fine. Un des résultats les plus significatifs des recherches sur les multifractals concerne le comportement des moments statistiques d'ordre suffisamment élevé de ces propriétés habillées, lesquels peuvent tendre vers l'infini. Cette divergence qui peut apparaître dans certains cas de figure (MANDELBROT,

1974 ; SCHERTZER et LOVEJOY, 1987b), est équivalente à une chute algébrique (hyperbolique) de la distribution des intensités au-delà d'un certain seuil, chute lente qui correspond en fait à une variabilité maximale. La divergence des moments a de nombreuses conséquences théoriques et expérimentales puisque la loi des grands nombres ne s'applique plus, d'où perte d'ergodicité, divergence des estimateurs statistiques usuels et dépendance des estimations à la taille de l'échantillon.

On peut montrer (SCHERTZER et LOVEJOY, 1993) qu'il peut exister une valeur critique de l'ordre des moments  $q$ ,  $q = q_D$ , telle que les moments statistiques divergent dès que  $q > q_D$ .

Un ordre de singularité critique  $\gamma_D$  correspond à l'ordre critique  $q_D$  relatif aux moments et on démontre que pour les valeurs de l'ordre de singularité  $\gamma$  supérieures à  $\gamma_D$ , la fonction de codimension  $C(\gamma)$  devient linéaire, sa pente étant égale à  $q_D$ . Des analogies thermodynamiques ont conduit à qualifier de « transition de phase » le passage au-delà de l'ordre de singularité critique  $\gamma_D$  (SCHERTZER et LOVEJOY, 1992).

$$C(\gamma) = q_D(\gamma - \gamma_D) + C(\gamma_D) \quad \text{si } \gamma > \gamma_D$$

Dans ces conditions, pour des valeurs suffisamment grandes de  $\gamma$ , la distribution de probabilité du flux  $\varepsilon_\lambda$  peut se réécrire :

$$\text{Prob}[\varepsilon_\lambda \geq \lambda^\gamma] = A'_\lambda \lambda^{-q_D \gamma}$$

Si nous définissons un seuil  $\varepsilon = \lambda^\gamma$ , il vient alors :

$$\text{Prob}[\varepsilon_\lambda \geq \varepsilon] = A'_\lambda \varepsilon^{-q_D} \quad (1)$$

Cette équation signifie que, quelle que soit l'échelle  $\lambda$  considérée, la probabilité au dépassement d'un seuil suffisamment grand est une fonction algébrique décroissante de ce seuil, et il est important de souligner que l'exposant de cette relation algébrique est indépendant de l'échelle considérée.

Il faut noter que le comportement de telles distributions diffère notablement de celui des lois classiques qui présupposent généralement des lois de nature exponentielle dont la décroissance est infiniment plus rapide.

### Application à de longues séries pluviométriques

Revenant aux cumuls annuels des précipitations, nous avons voulu voir dans quelle mesure le formalisme précédent pouvaient expliquer le comportement de leurs valeurs extrêmes. C'est sur ces bases que nous avons étudié les queues des distributions de précipitations annuelles et notamment voulu voir si une hypothèse de décroissance algébrique de la probabilité de dépassement d'un seuil était recevable.

Nous nous sommes intéressés à 87 séries pluviométriques annuelles recueillies dans le cadre du projet FRIEND-AMHY du Programme Hydrologique International de l'Unesco (BENDJOURI et HUBERT, 1997). La longueur des 87 séries recueillies dans les pays de la bordure méditerranéenne et de l'arc alpin va de 44 à 266 ans avec une valeur moyenne de 116 ans.



L'hypothèse d'un comportement algébrique de la probabilité au dépassement d'un seuil se traduit mathématiquement par l'équation (1) qui, pour un seuil  $h$  de la précipitation  $H$ , peut s'écrire :

$$\text{Prob}[H > h] = Ah^{-qD} \quad (2)$$

mais on ne peut éventuellement s'attendre à un tel comportement que pour une valeur suffisamment grande du seuil  $h$ , ou de façon équivalente, pour des probabilités au dépassement suffisamment petites.

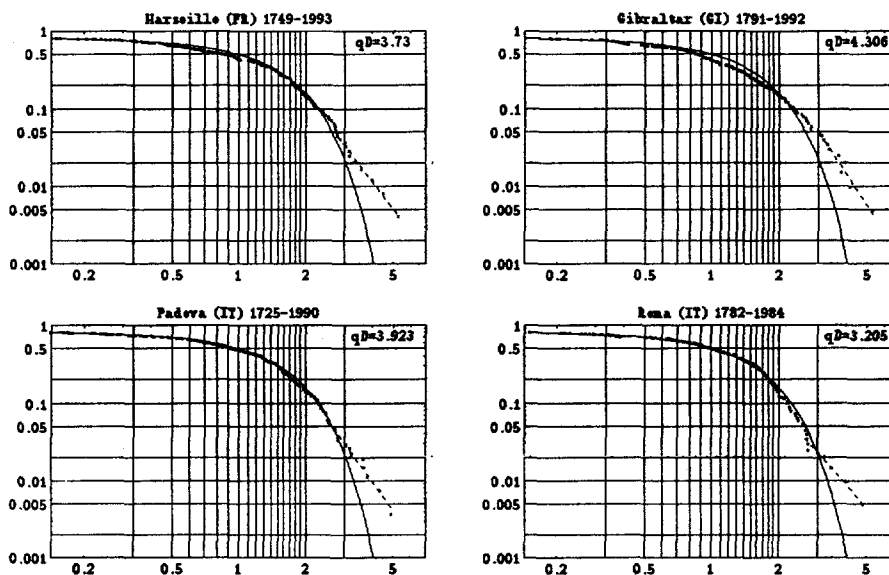
Les séries ayant été normées (moyenne = 1) et réduites (écart type = 1) nous avons reporté pour chacune d'entre elles, sur un diagramme log-log, les points représentatifs des probabilités empiriques au dépassement, estimées par la formule de Weibull. Nous avons superposé à ces points la courbe représentant la probabilité au dépassement résultant de l'ajustement à une loi normale des données de la série.

De façon quasi systématique on note que les fortes valeurs des précipitations annuelles s'écartent notablement de l'ajustement gaussien. À partir d'une probabilité au dépassement se situant autour de 0,05 cet ajustement semble minorer considérablement les probabilités d'occurrence des précipitations. Nous inspirant de l'allure de ces graphiques nous avons systématiquement ajusté une droite aux points dont la probabilité empirique au dépassement est inférieure à 0,05. Les graphiques montrent, et le calcul du coefficient de corrélation confirme, que l'ajustement à une droite des probabilités de dépassement d'un seuil en fonction de ce seuil, en axes log-log, est acceptable pour les valeurs extrêmes. Ce résultat apporte donc des arguments en faveur de l'hypothèse d'un comportement hyperbolique, plutôt qu'exponentiel, de la distribution de probabilité des fortes précipitations annuelles. La figure 2 montre quatre exemples de ces tracés pour les quatre plus longues séries disponibles dans notre base : Padoue, Marseille, Rome et Gibraltar dont la durée est supérieure à 200 ans.

### Étude de l'exposant de la décroissance algébrique de la distribution de probabilité

L'équation (2) traduite en axes log-log, est celle d'une droite de pente  $[-qD]$ . Les pentes des droites ajustées aux queues de distribution permettent une estimation du paramètre  $qD$  pour chacune de nos séries. On trouvera dans le tableau 1 la liste des stations avec pour chacune d'elles, son nom, son pays d'origine (AFNOR, 1994), sa longueur en années, la valeur du paramètre  $qD$  ajusté, le coefficient de corrélation de cet ajustement, le nombre de points utilisés pour l'ajustement ainsi que ses coefficients d'asymétrie et d'aplatissement. Seules les 71 stations assez longues (au-delà de 90 ans) pour que le paramètre  $qD$  soit estimé à partir de quatre points au moins, ont été retenues.

On remarque une très importante dispersion des valeurs estimées du paramètre  $qD$  selon les stations. Ces valeurs s'échelonnent de 2 à 23 avec une moyenne globale voisine de 6. Il nous a semblé intéressant de regarder s'il existait un lien entre la valeur du paramètre  $qD$  et la longueur de la série ayant servi à l'estimer. La figure 3 montre cette relation. On y retrouve la grande dispersion des valeurs de  $qD$  mais on constate que la dispersion des estimations est d'autant plus faible que la longueur des séries augmente. Les points représentatifs semblent être encadrés par deux courbes enveloppes virtuelles, respectivement



**Figure 2** Probabilités empiriques et ajustement à une loi normale des séries de Padoue, Marseille, Rome et Gibraltar. Les valeurs empiriques sont représentées par des points et le meilleur ajustement gaussien est représenté par un trait plein. En pointillé est représenté l'ajustement à une droite de chaque queue de distribution.

*Empirical probabilities and fitting to the normal law for Padova, Marseilles, Rome and Gibraltar series. marks represent empirical values and full line shows the better gaussian fitting. Break line represents the fitting to a straight line of each distribution tail.*

majorante et minorante, et convergeant vers une même limite. Nous avons estimé cette limite à  $3,8 \pm 0,5$  en nous fondant sur la moyenne et la dispersion des  $qD$  relatifs aux quatre séries les plus longues (Padoue, Marseille, Rome et Gibraltar).

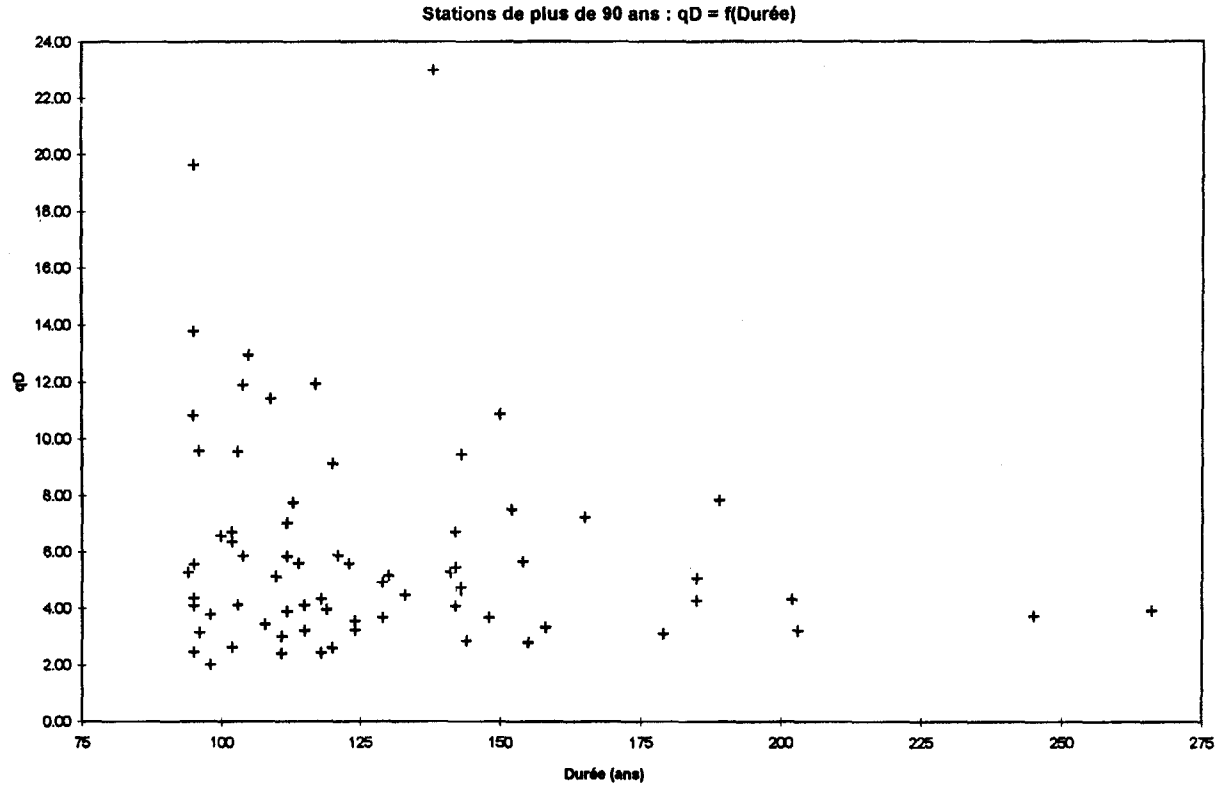
## CONCLUSIONS

L'ensemble des résultats précédents nous amène à la conjecture suivante :

**Pour des valeurs suffisamment élevées (dont la probabilité au dépassement est de l'ordre de 0,05, c'est-à-dire dont la durée de retour est supérieure à 20 ans), la probabilité de distribution des précipitations annuelles est de nature hyperbolique et le paramètre décrivant cette décroissance algébrique de probabilité est universel. La valeur de ce paramètre peut être estimée à  $3,8 \pm 0,5$ .**

Si une telle conjecture devait être confirmée, notamment par l'étude d'autres longues séries recueillies de par le monde, ses conséquences pratiques seraient considérables.





**Figure 3** Variation de l'exposant de la décroissance algébrique ( $qD$ ) en fonction de la longueur de la série pluviométrique.  
*Varying of the algebraic downfall exponent ( $qD$ ) function of the length of the rainfall serie.*

Par exemple, en ce qui concerne l'estimation des événements extrêmes, pour les cumuls annuels de pluie qui nous intéressent ici, l'année millénaire selon la loi normale n'est plus, en gros, que centennale selon une loi à décroissance hyperbolique. On admettra aisément que la différence est de taille et pourrait avoir des conséquences pratiques de très grande importance. Le tableau 2 montre les estimations de ces différences de période de retour pour les stations de Padoue, Marseille, Rome et Gibraltar.

**Tableau 2** Période de retour « réelle » d'une pluie millénaire selon la loi normale.

**Table 2** *Real return period of the millennial rainfall derived from normal law.*

Station	Durée	Période de retour	
		Normale	Empirique
Gibraltar	202	1 000	60
Rome	203	1 000	100
Marseille	245	1 000	100
Padoue	266	1 000	100

La difficulté de l'estimation de la récurrence des événements extrêmes a été remise à l'ordre du jour par les événements catastrophiques de l'automne 1993 et de l'hiver 1994 dans le sud de la France. Les résultats obtenus par CERNESSON (1994) donnent des périodes de retour surestimées par rapport aux travaux récents de NEPPEL (1997).

Mais ces préoccupations ne sont pas limitées à la zone d'extraction des données exploitées dans la présente étude. À titre d'exemple : les premières informations recueillies sur les pluies diluviennes de juillet 1997 en Allemagne, Pologne et République Tchèque (crues de l'Oder), permettraient d'estimer, de façon préliminaire, la période de retour à plus de 10 000 ans (Gazowski, communication personnelle) par les ajustements classiques, alors que cette période de retour pourrait se révéler moins extrême si le modèle que nous proposons s'applique effectivement.

Par ailleurs l'examen des valeurs du coefficient d'asymétrie reportées dans le tableau 1 montre une nette prédominance des valeurs positives (61 sur 71) ce qui ne milite pas en faveur de l'hypothèse de normalité. On pourrait alors se poser la question de savoir ce que donnerait un ajustement avec des lois à dissymétrie positive. Cependant les conclusions données plus haut ne s'appliquent pas seulement à la loi normale mais à toute loi à décroissance exponentielle, qui ne peut envelopper une décroissance algébrique.

Plutôt que de peaufiner des méthodes de calage des lois traditionnelles qui sont pour la plupart à décroissance exponentielle, il semble donc nécessaire pour les hydrologues statisticiens d'élargir leur champ d'investigation à des lois susceptibles de reproduire un comportement hyperbolique comme la loi de Halphen de type B-1 (MORLAT, 1956 ; PERRON *et al*, 1996) ou la loi de Pareto (VAN MONFORT ET WITTER, 1986).

De plus la possible universalité du paramètre  $qD$  ouvre aussi d'immenses perspectives. Nous avons déjà signalé qu'il sera nécessaire d'étendre le champ

géographique des investigations : les premiers traitements effectués sur quelques séries pluviométriques d'Europe du nord aboutissent à des résultats analogues à ceux de la présente étude.

## REMERCIEMENTS

Nous tenons à remercier G. de Marsily, P. Givone, G. Oberlin pour leurs critiques pertinentes et leurs suggestions, D. Schertzer pour son soutien théorique, ainsi que Z. Gazowski pour les données de la crue de l'Oder et tous les collaborateurs du thème « longues séries hydropluviométriques » du projet FRIEND-AMHY pour avoir fourni les données de ce travail.

## RÉFÉRENCES BIBLIOGRAPHIQUES

- AFBSN, 1972. *Les bassins de la Seine et des cours d'eau normands*, Tome 1, Ressources d'eau et données hydrologiques, Fascicule 3, Données climatiques, Agence Financière de Bassin Seine Normandie, Paris, 112 p.
- AFNOR, 1994. *Codes pour la représentation des noms de pays ; norme NF EN 23166*, ISO 3166.
- ARLERY R., 1950. Sur la signification des moyennes de la pluie annuelle à Paris, *Journal scientifique de la Météorologie*, 2, pp 82-99.
- BENDJOUDI H. et HUBERT P., 1997. Longues séries disponibles dans la base Amhy, FRIEND Flow Regimes from International Experimental and Network Data, Projects H-5-5 (IHP IV) and 1.1 (IHP V) Third report : 1994-1997, CEMAGREF éditions, pp. 362-366.
- BENDJOUDI H., HUBERT P., SCHERTZER D. et LOVEJOY S., 1997. Interprétation multifractale des courbes intensité-durée-fréquence des précipitations, *CR Acad. Sci. Paris, Sciences de la vie*, 1997, 325, pp. 323-326.
- CAMUFFO D., 1984. Analysis of the series of precipitations at Padova, Italy, *Climatic Change*, 6, pp. 57-77.
- CERNESSON F., 1994. De la difficulté d'estimer la récurrence des épisodes pluvieux extrêmes, in *Les événements hydropluviométriques intenses récemment observés sur le Sud-Est de la France*. Actes de la journée scientifique du 2 juin 1994, J. Lavabre [Ed.], Cemagref, Aix-en-Provence.
- DAGNELIE P., 1970. *Théorie et méthodes statistiques*, volumes 1 et 2, Duculot, Gembloux, 378 et 451 p.
- HAAN C.T., 1977. *Statistical methods in hydrology*, The Iowa State University Press, Ames, 378 p.
- HUBERT P. et CARBONNEL J.P., 1988. Caractérisation fractale de la variabilité et de l'anisotropie des précipitations tropicales. *CR Acad. Sci. Paris* 2, 307, pp. 909-914.
- HUBERT P. et CARBONNEL J.P., 1989. Dimensions fractales de l'occurrence de pluie en climat Soudano-Sahélien. *Hydro, Continent*, 4, pp. 3-10.
- HUBERT P., FRIGGIT F., CARBONNEL J.P., 1995. Multifractal structure of Rainfall Occurrence in West Africa, Z.W. Kundzewicz [Ed.], *New Uncertainty Concepts in Hydrology and Water Resources*, Cambridge University Press, pp. 109-113.
- HUBERT P. et BENDJOUDI H., 1997. Caractéristiques fractales des séries pluviométriques, FRIEND Flow Regimes from International Experimental and Network

- Data, Projects H-5-5 (IHP IV) and 1.1 (IHP V) Third report : 1994-1997, CEMAGREF éditions, pp. 352-356.
- KOTTEGODA N. T., 1980. *Stochastic Water Resources Technology*, The MacMillan Press Ltd, 384 p.
- MANDELBROT B.B., 1974. Intermittent turbulence in self-similar cascades : divergence of high moments and dimension of the carrier, *J. Fluid Mech.*, 62, 331-350.
- MANDELBROT B.B., 1982. *The fractal geometry of nature*, Freeman, San Francisco, 460 p.
- MARKOVIC R.D., 1965. Probability functions of best fit to distributions of annual precipitation and runoff, Colorado State University Hydrology Paper n° 8, Fort Collins.
- MORLAT G., 1956. Les lois de probabilité de Halphen, *Revue de statistique appliquée*, 4, 21-43.
- NEPPEL L., 1997. Le risque pluvial en région Languedoc : caractérisation de l'aléa climatique, Thèse Université Montpellier II.
- PERRON H., BOBEE B., PERREAULT L. et BRUNEAU P., 1996. Estimation des événements hydrologiques extrêmes (crues et étiages), *Comptes rendus des Rencontres hydrologiques Franco-Roumaines de Tulcea*, Edité par J.P. Carbonnel et al., Unesco, Paris, publication SC-96/WS/19, volume II, pp. 363-372.
- ROCHE M., 1963. *Hydrologie de surface*, Gauthier-Villars, Paris, 429 p.
- REMENIERAS G., 1976. *L'hydrologie de l'ingénieur*, seconde édition, Eyrolles, Paris, 456 p.
- SCHERTZER D. et LOVEJOY S., 1987a. Physically based rain and cloud modeling by anisotropic scaling of multiplicative processes, *Journal of Geophysical Research* D, 92, pp. 9693-9714.
- SCHERTZER D. et LOVEJOY S., 1987b. Singularités anisotropes et divergences des moments en turbulence, *Ann. Sc. Math. du Québec*, 11, 139-181.
- SCHERTZER D. et LOVEJOY S., 1991. Scaling nonlinear variability in geodynamics : Multiple singularities, observables, universality classes ; Non-linear variability in Geophysics, Scaling and fractals, D. Schertzer and S. Lovejoy eds, Kluwer, pp. 41-82.
- SCHERTZER D., LOVEJOY S., 1992. Hard and soft multifractal processes, *Physica A*, 195, 187-194.
- SCHERTZER D., LOVEJOY S., 1993. Nonlinear Variability in Geophysics : Scaling and Multifractal Processes, Lecture Notes. Institut d'études scientifiques de Cargèse.
- TUCCI C.E.M., 1993. *Hidrologia, ciência e aplicação*, Editoria da Universidade, Porto Alegre, 943 p.
- VAN MONFORT M.A.J. et Witter J.V., 1986. The Generalized Pareto distribution applied to rainfall depths, *Hydrological Journal des Sciences Hydrologiques*, 31, 151-162.
- YEVJEVICH V., 1972. *Probability and statistics in Hydrology*, Water Resources Publications, Fort Collins, 302 p.