

Article

« Amélioration des performances d'un modèle stochastique de génération de hyétogrammes horaires : application au pourtour méditerranéen français »

P. Arnaud, J. Lavabre et J. M. Masson

Revue des sciences de l'eau / Journal of Water Science, vol. 12, n° 2, 1999, p. 251-271.

Pour citer cet article, utiliser l'information suivante :

URI: <http://id.erudit.org/iderudit/705351ar>

DOI: 10.7202/705351ar

Note : les règles d'écriture des références bibliographiques peuvent varier selon les différents domaines du savoir.

Ce document est protégé par la loi sur le droit d'auteur. L'utilisation des services d'Érudit (y compris la reproduction) est assujettie à sa politique d'utilisation que vous pouvez consulter à l'URI <https://apropos.erudit.org/fr/usagers/politique-dutilisation/>

Érudit est un consortium interuniversitaire sans but lucratif composé de l'Université de Montréal, l'Université Laval et l'Université du Québec à Montréal. Il a pour mission la promotion et la valorisation de la recherche. Érudit offre des services d'édition numérique de documents scientifiques depuis 1998.

Pour communiquer avec les responsables d'Érudit : info@erudit.org

Amélioration des performances d'un modèle stochastique de génération de hyétogrammes horaires : application au pourtour méditerranéen français

Results improvement of stochastic model for generating
hourly hyetographs, applied to the french seaboard

P. ARNAUD^{1*}, J. LAVABRE¹ et J.-M. MASSON²

Reçu le 23 janvier 1998, accepté le 13 janvier 1999**.

SUMMARY

A stochastic model for generating hourly hyetographs has been recently developed, in the Cemagref of Aix-en-Provence, to be coupled with a rainfall runoff conversion modelling. Thus, by simulation of very long periods (1000 years for example), we obtain a large number of hourly hyetographs and flood scenarios that are statistically studied and used in flood predetermination problems. The rainfall model studied is based on the theory that rainfall can be linked to a random and intermittent process whose evolution is described by stochastic laws. It is also based on the hypothesis of independence between variables describing hyetographs and on the hypothesis of the stationary nature of the phenomenon studied. Generating a rainfall time series involves two steps: descriptive study of the phenomenon (nine independent variables are chosen to describe the phenomenon and these variables are defined by a theoretical law of probability fitted to the observations) and creation of a rainfall time series using descriptive variables generated randomly from their law of probability. Initially developed on the Réal Collobrier watershed data, the model has been applied to fifty raingauges located on the Mediterranean French seaboard. The extension of the model applying area has shown heterogeneity in the results. Therefore, modifications have been made to the model to improve its performances. Among these modifications, three of them have presented notable improvements.

A study of the sensitivity of the parameters has been made. Parameters of shape variables and of some other variables had only a slight influence on depth of generated rainfalls. But, the law of mean rainfall intensities clearly differentiates the stations. Then, a theoretical probability distribution for the storm intensity variable, less sensitive to the sampling problems, has been

1. Cemagref, Le Tholonet, B.P. 31, 13612 Aix-en-Provence, France.

2. UMR5569-GBE, Université Montpellier II, Place Eugène Bataillon, 34095 Montpellier cedex 5, France.

* Correspondance. e-mail : patrick.arnaud@cemagref.fr

** Les commentaires seront reçus jusqu'au 30 décembre 1999.

searched. An exponential distribution is fitted to the value smaller than four times the mean of the variable. A slope breakage was then introduced to generate all the values beyond this limit. The breakage at the value four times the mean of the variable and modelling this breakage were based on a study of so-called "regional" distributions of the storm intensity variable. These distributions were designed by clustering the variable's homogenized values for all 50 studied stations.

A second modification has been made to develop new model for the observed dependence between two variables (duration and intensity of the storm). The study of this dependence has been considered directly based on the cumulative frequency of the two variables. Then, an additional parameter was defined to model the dependence between the probabilities of the two variables. This parameter characterises the cumulative frequency curve of the sum of the probabilities of the two variables. This point, neglected during a long time, has been very important in the improvement of the model.

Finally, the modelling of storm persistence in a same rainfall episode has been studied to generate some high 24 hours maximum rainfalls. Persistence modelling is entirely justified by the fact that "ordinary storms" cluster together around the "main storm" (the "main storm" is the greatest storm of an episode and the "ordinary storms" are the other storms of the episode). When the study of this phenomenon is extended, it can be observed that there is a certain positive dependency between occurrence probability of the "main storm" and occurrence probability of storms which come before or after it. Two combined effects occur: within one rainy episode, the strongest "ordinary storms" are preferentially clustered together around the "main storm", and considering the number of "ordinary storms" throughout all the episodes, the strongest storms close to the "main storm" are preferentially associated with the strongest "main storms" and vice versa. This modification improves the performances of the altitude raingauges, which are characterised by high daily rainfall accumulations.

The different modifications added to the initial model, give very important improvements on the calibration of the fifty raingauges studied on the French Mediterranean seaboard. Its aptitude to generate rains observed in Mediterranean climate, strongly variables, consolidates us in the idea of its application on a zone much larger. The generation of hyetographs makes it possible to use the maximum the temporal information of the rain. Thus, we obtain a reliable tool, validated on a large area, for simulating hyetographs and hourly flood scenarios at all frequencies, and used instead of a unique design storm and design flood. The approach allows a new cumulative probability curve extrapolation, which seems sometimes greater than an exponential behaviour. Moreover, the study of many events per year, with many occurrences of the different variables of the model, increase the analysed sample size and seems to make the method more reliable than a statistical approach simply based, for example, on the fitting of annual maximum values.

Key-words: stochastic model, hourly rainfall, french mediterranean seaboard, dependence of variables.

RÉSUMÉ

Depuis quelques années, un modèle stochastique de génération de hyétoigrammes horaires est développé au groupement d'Aix-en-Provence du Cemagref, pour être couplé à une modélisation de la pluie en débit, fournissant ainsi une multitude de scénarios de crues analysés statistiquement et utilisés en prédétermination des débits de crues. L'extension de la zone d'application du modèle de pluies horaires au-delà de sa zone de conception, a fait apparaître une hétérogénéité dans les résultats. Ce constat a entraîné certaines modifica-

tions du modèle comme : la recherche d'une loi de probabilité théorique peu sensible aux problèmes d'échantillonnage pour une variable du modèle (intensité d'une averse), la prise en compte originale de la dépendance observée entre deux variables du modèle (durée et intensité d'une averse), et la modélisation de la persistance des averses au sein d'une même période pluvieuse. Ces différentes modifications apportées au modèle initial ont entraîné une très nette amélioration de ses performances sur la cinquantaine de postes pluviographiques du pourtour méditerranéen français. On obtient ainsi un outil beaucoup plus robuste et validé sur une zone étendue, capable de fournir de multiples formes de hyétogrammes, couvrant toute la gamme des fréquences, permettant ainsi de s'affranchir des pluies de projet uniques. On aborde aussi une nouvelle approche du comportement à l'infini des distributions de fréquences des pluies qui semble parfois supérieur à une tendance strictement exponentielle. De plus, l'étude de plusieurs événements par an dont chacun présente plusieurs réalisations des différentes variables du modèle augmente la taille des échantillons analysés, semblant rendre la méthode plus rapidement fiable qu'une approche statistique classique basée par exemple sur l'ajustement de valeurs maximales annuelles.

Mots clés : modèle stochastique, pluie horaire, pourtour méditerranéen français, variables dépendantes.

1 – INTRODUCTION

Afin d'aborder la prédétermination des débits de crues, un modèle stochastique de génération de hyétogrammes a été développé au pas de temps horaire. Couplé à une modélisation de la pluie en débit, il permet d'obtenir des scénarios de crues dont les caractéristiques et les conséquences peuvent être analysées statistiquement.

La construction de multiples hyétogrammes, sur de très longues périodes de simulation, permet une appréhension originale des problèmes de prédétermination classique : nouvelle approche du comportement asymptotique des distributions de probabilité des cumuls de pluies (et des débits) sur une durée donnée, élaboration de pluies de projet plus réalistes, etc.

Initialement mis au point sur un nombre restreint de postes pluviographiques, le modèle a ensuite été calé sur une zone plus étendue que sa zone de conception. L'extension de la zone d'étude a alors fait apparaître des biais dans les performances du modèle initialement développé. C'est pourquoi, afin de rendre le modèle performant et robuste sur une zone étendue (le pourtour méditerranéen français), des modifications de sa structure et la prise en compte de nouvelles caractéristiques omises jusqu'à présent, ont été nécessaires. On présente ici, les modifications majeures ayant entraîné une amélioration des performances du modèle.

2 – PRÉSENTATION DU MODÈLE INITIAL

Il existe différentes sortes de modèles étudiant la pluie. On trouve des modèles à base physique qui tentent de modéliser la physique atmosphérique (comme le modèle de GEORGAKAKOS et BRAS (1984)). La complexité de la modélisation des mécanismes à l'origine des précipitations, l'importance de l'initialisation et leur grand nombre de paramètres rendent ces modèles difficilement utilisables en hydrologie opérationnelle.

La pluie est plutôt étudiée par des approches statistiques. L'approche empirique qui consiste à ajuster des lois de probabilité sur des variables pluviométriques observées, est l'approche la plus ancienne et la plus répandue. Un choix, parmi différentes lois théoriques et méthodes d'ajustement, permet de retenir une loi de probabilité qui s'ajuste bien aux valeurs observées d'une variable quelconque (généralement la hauteur de pluie observée sur une durée donnée). Issue de la détermination de différents quantiles de variables pluviométriques, l'élaboration de pluies de projet associées à une probabilité d'occurrence est alors effectuée. L'inconvénient de cette approche est le non-respect de l'information temporelle des hyétogrammes et la difficulté d'associer la probabilité d'occurrence d'une variable pluviométrique unique à la probabilité d'occurrence d'un hyétogramme.

Enfin, on trouve une approche, que l'on pourrait qualifier de « conceptuelle », qui vise à générer des séries chronologiques de pluie. La modélisation est alors basée sur une étude statistique de variables descriptives nécessaires à la construction des chroniques de pluie. Différentes approches sont alors possibles pour simuler la pluie au pas de temps horaire, parmi lesquelles on retrouve :

- les modèles de désagrégation de la pluie journalière (WOOLHISER et OSBORN (1985), ECONOPOULY *et al* (1990), KOUTSOYIANNIS (1994)) effectuant le passage de la pluie journalière à la pluie à pas de temps plus fins ;

- les modèles à processus d'agrégation qui ont émergé suite aux travaux de LECAM (1961). Ces modèles s'articulent autour de la modélisation de deux phénomènes aléatoires qui sont, l'occurrence et le positionnement de cellules pluvieuses, et la caractérisation du volume de ces cellules (par leur durée et leur intensité) (ACREMAN (1990), COWPERTWAIT (1991), ONOF *et al.* (1995), RODRIGUEZ-ITURBE *et al.* (1987), WAYMIRE et GUPTA (1981)) ;

- les modèles basés sur l'approche directe, (LEBEL (1984), TOURASSE (1981)), qui s'attache à une description géométrique des hyétogrammes horaires par l'intermédiaire de variables aléatoires.

Le modèle stochastique développé au Cemagref (ARNAUD (1997), CERNESSON (1993)) est un modèle basé sur l'approche directe. Il part du principe que la pluie peut être assimilée à un processus aléatoire et intermittent (succession d'états secs et pluvieux) dont l'évolution est décrite par des lois de nature stochastique. Il repose en outre sur l'hypothèse d'indépendance des variables décrivant les hyétogrammes, ainsi que sur l'hypothèse de stationnarité du phénomène étudié. La génération du signal de pluie comporte deux étapes (*figure 1*).

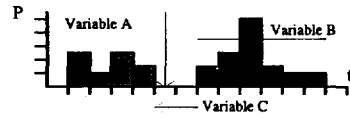
La première étape constitue l'étude descriptive du phénomène. Elle amène au choix des variables qui vont le caractériser. Par hypothèse, chaque variable descriptive devra être indépendante des autres, et sera définie par une loi de probabilité théorique ajustée sur les observations.

1ère ETAPE : ETUDE DESCRIPTIVE

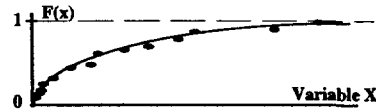
DEFINIR des variables décrivant la pluie indépendantes les unes des autres.

Sur les observations, EXTRAIRE les valeurs des variables descriptives.

AJUSTER la distribution de probabilité de chaque variable descriptive par une loi théorique adaptée.



Variable A	A1	A2	A3	...
Variable B	B1	B2	B3	...
Variable C	C1	C2	C3	...
...



2ème ETAPE : RESTITUTION DE LA PLUIE

GENERER les variables par un tirage aléatoire dans leur loi de distribution.

RECONSTRUIRE les hyétogrammes à partir des variables générées.

VALIDER sur des variables tests non utilisées dans la modélisation

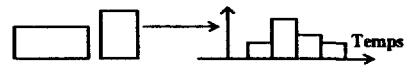
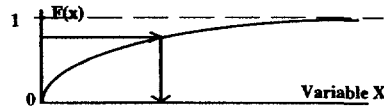


Figure 1 Principe du modèle stochastique de simulation des hyétogrammes.
Principle of the stochastic model for generating hourly hyetographs.

La seconde étape est la construction de la chronique de pluie, à partir des variables descriptives générées de façon aléatoire à partir de leur loi de probabilité. De manière générale, si une variable aléatoire X peut-être représentée par une fonction de répartition $F(x)$, la génération de variables suivant la loi $F(x)$ se fait en tirant de façon aléatoire une valeur de $F(x)$ entre 0 et 1, puis en cherchant la valeur correspondante de x , en inversant la fonction de répartition.

Notons que le modèle fonctionne en mode événementiel, c'est-à-dire qu'il analyse puis simule des événements indépendants, en distinguant deux saisons : l'été de juin à novembre, l'hiver de décembre à mai.

L'analyse porte alors sur des événements pluvieux sélectionnés et définis à partir de critères journaliers : événements bornés par des pluies journalières inférieures à 4 mm et présentant une pluie journalière d'au moins 20 mm. À partir de ces événements sélectionnés au pas de temps journalier, les hyétogrammes observés au pas de temps horaires, vont être caractérisés par neuf variables descriptives (CERNESSON *et al.* (1996)) : le nombre moyen d'événements par an (NE), l'origine de l'événement dans la journée (TSE), le nombre de périodes pluvieuses dans un événement (NG), le nombre d'averses dans une période pluvieuse (NA) et pour chaque averse sa durée (DA), son intensité moyenne (HMA), la durée sèche (DIA) qui la sépare de l'averse suivante (nulle pour deux averses groupées), la position relative du maximum (RPX) et le rapport entre l'intensité horaire maximale et moyenne (RX).

Notons que l'on distingue les intensités moyennes et les durées de deux types d'averses : l'averse principale qui apporte la plus grande quantité d'eau (produit

HMA \times DA maximum) au cours d'un événement (notation HMAP et DAP), les autres averses étant considérées comme ordinaires (notation HMAO et DAO).

L'étude d'une cinquantaine de postes pluviographiques a permis de déterminer les lois de probabilité théoriques qui caractérisent le mieux la distribution des différentes variables descriptives :

- une loi de Poisson est utilisée pour les variables NE, DAP et DAO,
- une loi géométrique est utilisée pour les variables NG, NA, TSE et DIA,
- une loi normale tronquée entre 0 et 1 est utilisée pour la variable RPX,
- une loi de type exponentiel est utilisée pour les variables RX, HMAP et HMAO.

Le calage de ces lois de probabilité porte alors sur 19 paramètres déterminés pour chaque saison. Les hyétogrammes horaires sont construits averse après averse à partir des variables descriptives générées par un tirage aléatoire dans leur loi de distribution, suivant un ordre précis. Rappelons que l'on fait, jusque là, l'hypothèse d'indépendance des différentes variables descriptives, ce qui permet leur tirage indépendamment les unes des autres au fur et à mesure des besoins.

Les performances du modèle sont jugées sur sa capacité à reproduire des variables « tests » représentatives de la pluviométrie et n'ayant pas servi à l'élaboration des hyétogrammes. La validation porte alors sur la restitution des pluies maximales en 1, 2, 3, 4, 6, 12, 24, 48 et 72 heures (PM1, PM2...), ainsi que sur la durée et la hauteur totale des événements (DTOT, HTOT). L'écart entre les valeurs des variables « tests » observées et calculées, est jugé sur leurs moyennes et leurs écarts-types, indicateurs du respect de leur distribution de probabilité. La bonne restitution des distributions de fréquence de ces variables « tests », nous garantit le respect de la structure temporelle des hyétogrammes observés.

3 – EXTENSION DE LA ZONE D'APPLICATION DU MODÈLE INITIAL

Initialement, le modèle a été développé sur les postes pluviographiques du Bassin Versant de Recherche Expérimentale (BVRE) du Réal Collobrier par CERNESSON (1993). Par la suite, pour étudier la régionalisation des paramètres, le modèle a été calé localement sur 24 postes pluviographiques dans un premier temps, par CERNESSON (1993) et RIBERO (1992), puis sur une cinquantaine de postes du pourtour méditerranéen français (ARNAUD (1997)) (figure 2).

Hors de sa zone de conception le modèle a fourni des résultats assez hétérogènes (figure 3).

On présente sur cette figure la distribution de fréquence des écarts relatifs ($\text{écart}(\%) = 100 \times (\text{calculé} - \text{observé})/\text{observé}$) calculés sur les moyennes et les écarts-types de différentes variables « tests », entre leurs valeurs observées et l'ensemble de leurs valeurs simulées sur 1 000 ans par le modèle initial, appliqué sur les 50 postes pluviographiques (50 écarts par saison et par variable « test »).

La tendance générale du modèle était de surestimer les moyennes et les écarts-types des variables « tests ». Les variables « tests » les plus surestimées étaient les pluies maximales de 4 à 24 h, avec une surestimation plus forte des écarts-types que des moyennes. Cette surestimation concernait surtout les pluies

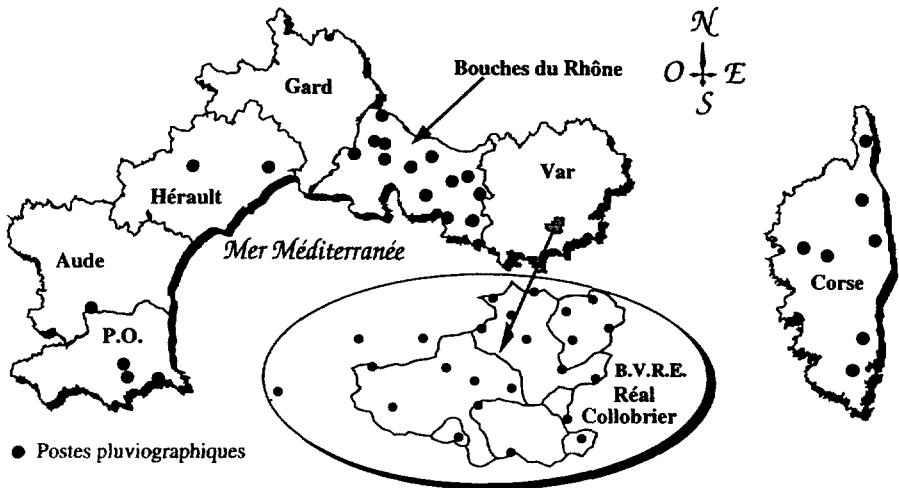


Figure 2 Localisation des 50 postes pluviographiques ayant servi au calage du modèle.
Localization of the 50 rainfall recorders used to calibrate the parameters of the model.

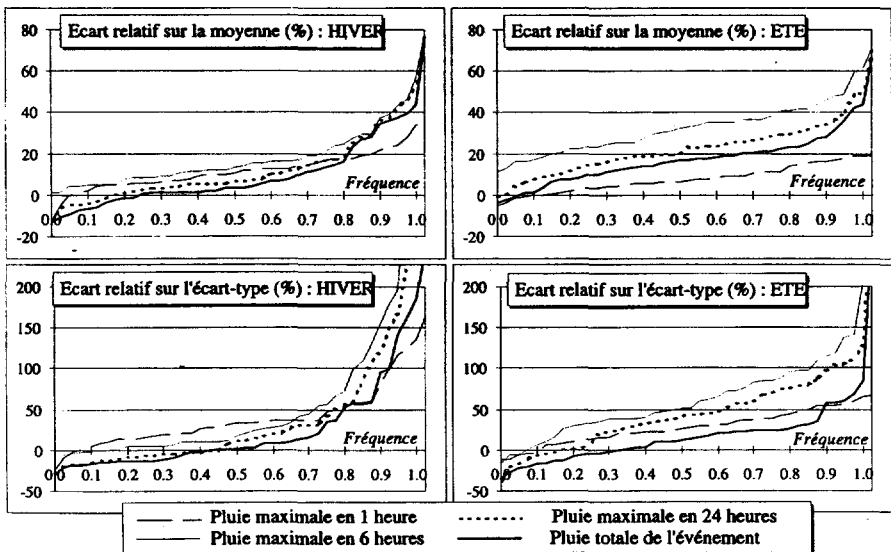


Figure 3 Distributions de fréquence des écarts relatifs calculés sur les moyennes et sur les écarts-types des variables « tests » observées et issues d'une simulation de 1 000 ans par le modèle initial.
Distribution of relative differences between observed and simulated (1000 years, by initial model) mean and standard deviation of « test » variables.

des postes les moins pluvieux, en particulier ceux des Bouches du Rhône, alors que les postes du BVRE du Réal Collobrier, postes de développement du modèle, présentaient les résultats les plus satisfaisants. Les résultats du modèle appliqué sur une zone plus large, n'étaient donc pas satisfaisants sur l'ensemble des postes pluviographiques.

L'extension de la zone d'application du modèle a montré les difficultés du modèle à s'adapter à de nouveaux postes. Il a alors été nécessaire d'apporter des modifications au modèle initial, afin de diminuer et d'homogénéiser ses erreurs et le rendre applicable au-delà de sa zone de conception.

4 – MODIFICATIONS APPORTÉES AU MODÈLE

Depuis la version du modèle présentée par CERNESSON *et al.* (1996), différentes modifications ont été effectuées, et tout d'abord pour respecter au mieux les distributions des variables descriptives du modèle. En effet, la génération des variables par inversion de leur fonction de répartition au fur et à mesure des besoins, conduit parfois à des hyétoigrammes simulés ne respectant pas les définitions de l'analyse initiale (par exemple le minimum relatif entre deux averses successives appartient à la première averse, etc.). Le non-respect de ces différentes contraintes a été résolu dans un premier temps par le rejet des hyétoigrammes générés. Ce procédé très pragmatique ne respecte pas la rigueur des techniques d'acceptation-rejet de simulation, et son emploi abusif a conduit à simuler des échantillons n'appartenant plus à la population mère initiale. La suppression directe de certaines contraintes peu importantes, et les modifications dans l'ordre des tirages aléatoires ont permis d'éliminer les biais introduits sur la génération des variables utilisées par le modèle, bien que la restitution des variables « tests » n'ait pas été sensiblement améliorée (ARNAUD, 1997). L'intérêt de cette première modification a été de respecter au mieux les variables descriptives utilisées par le modèle en évitant les biais qui auraient pu minimiser l'effet des modifications effectuées par la suite.

Trois modifications supplémentaires ont été réalisées, entraînant de nettes améliorations des performances du modèle. La première, appuyée sur l'étude de la sensibilité du modèle à ses différents paramètres, a été de trouver une loi théorique moins sensible à l'échantillonnage, pour la variable intensité. La seconde, a été de prendre en compte la liaison observée entre les intensités et les durées des averses. Enfin, la persistance des averses au sein d'un même événement a été examinée.

4.1 Recherche d'une loi des intensités moins sensible à l'échantillonnage

Les variables « intensité moyenne des averses » (HMAP et HMAO) présentent une distribution de fréquence de type exponentiel. Plusieurs lois théoriques ont été testées pour ajuster la distribution de ces variables : la loi exponentielle simple à 2 paramètres, la loi de Weibull à 3 paramètres (EV3), la loi de Jenkinson (GEV) et une loi à 4 paramètres combinant deux branches exponentielles.

On observe, sans surprise, que les lois théoriques ayant un nombre de paramètres élevé donnent l'ajustement le plus fidèle à la distribution observée. Cependant cet ajustement peut être fortement influencé par la présence de valeurs extrêmes se détachant de la distribution empirique (« outliers » des anglo-saxons, appelés « horsains » par MASSON (1991)). Plus la loi théorique aura de paramètres, plus elle sera sensible aux problèmes d'échantillonnage et à la présence de horsains. Afin d'éviter ces problèmes, nous avons retenu une loi faiblement paramétrée.

Parmi les lois testées, c'est la loi exponentielle simple, exprimée par (1), qui reste la moins sensible à l'échantillonnage du fait de ses deux paramètres θ et v :

$$P(X \leq x) = F(x) = 1 - e^{-\theta(x-v)} \quad \text{si } x > v \quad (1)$$

L'estimation de ses paramètres par la méthode des moments est basée sur le calcul de la moyenne μ_x et de l'écart-type σ_x de l'échantillon des observations (2) (DAGNELIE (1973)).

$$\theta = \frac{1}{\sigma_x} \quad \text{et} \quad v = \mu_x - \sigma_x \quad (2)$$

Ces paramètres sont tout de même soumis à l'influence des valeurs extrêmes. Par exemple, sur le poste pluviographique de Felce (Corse), la valeur maximale observée de la variable HMAP est de 53,2 mm/h (pendant 1 h). Si on omet cette valeur, la valeur maximale devient alors 17,8 mm/h (pendant 2 h), faisant passer l'écart-type de la variable de 5,5 mm/h à 3,3 mm/h. Les paramètres de la loi exponentielle, fonction de l'écart-type, vont alors aussi être fortement modifiés.

Il existe des tests statistiques basés sur la distribution d'échantillonnage empirique du coefficient de dissymétrie, qui permettent d'estimer la présence d'horsains (ROSSI *et al.* (1984)). Dans le cas des variables HMAP et HMAO, les valeurs observées sont bien sûr le résultat d'une mesure de précipitation, mais aussi le résultat du découpage temporel d'une averse (initialement discrétisée au pas de temps de 5 min) en tranches horaires. Pour reprendre l'exemple de l'intensité de 53,2 mm/h en une heure à la station de Felce, la même averse qui aurait duré 1 h 05, aurait été comptée sur deux pas de temps, divisant alors son intensité par 2. C'est pourquoi, on admet sans faire de tests sur les horsains, que certaines valeurs les plus extrêmes ne sont dues qu'au hasard du découpage horaire des averses et n'appartiennent pas à la population considérée.

Pour étudier cette population des intensités et en particulier leurs valeurs extrêmes, une analyse régionale a été effectuée sur la totalité des valeurs prises sur les 50 postes, afin d'augmenter l'effectif de l'échantillon analysé. Une homogénéisation des données est alors réalisée en divisant les valeurs de chaque série par sa moyenne. On obtient ainsi des séries de même moyenne, égale à 1.

On fait aussi l'hypothèse que les variables réduites par la moyenne ont le même écart-type (c'est le coefficient de variation des variables initiales), ce qui est vérifié si l'on ne tient pas compte des valeurs extrêmes. On trace alors la distribution de fréquence « régionale » des intensités « homogénéisées » sur un système d'axes qui représente la loi exponentielle par une droite (figure 4).

La fréquence F de non-dépassement est calculée par la formule de Hazen, $\mu(\text{HMAP})$ et $\mu(\text{HMAO})$ représentent les moyennes locales des variables HMAP et HMAO. On observe ainsi la tendance exponentielle de la distribution des intensi-

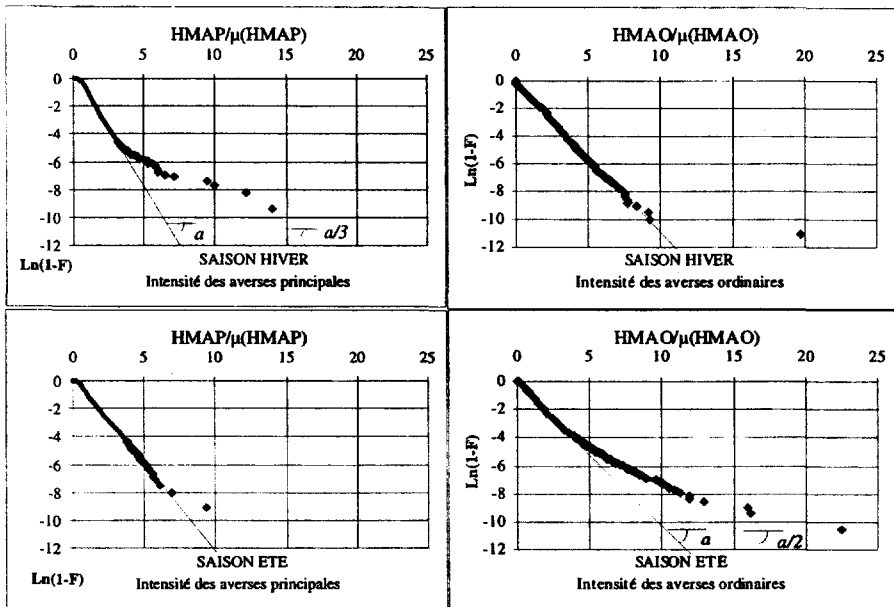


Figure 4 Distributions de fréquence « régionale » des « intensités moyennes » des averses de l'ensemble des 50 postes, réduites par leur moyenne locale.

Regional distribution of mean hourly intensity (all 50 stations considered together) reduced by its local mean.

tés moyennes d'averses sur l'ensemble des postes étudiés. Parmi les quatre cas de figure qui se présentent, deux présentent une rupture de pente de la distribution qui apparaît à partir des valeurs supérieures à 4 fois la moyenne de la série. À partir de ces résultats, on choisit de décrire les variables HMAP et HMAO par une loi exponentielle qui sera calée sur les valeurs inférieures à 4 fois la moyenne de la série, mais présentant une rupture de pente pour toutes les valeurs supérieures à cette borne. La valeur de la seconde pente sera déterminée par le rapport observé sur les distributions dites « régionales » et reporté sur le tableau 1 qui synthétise l'information obtenue par la figure 4.

Tableau 1 Caractéristiques des lois « régionales » des intensités moyennes des averses.

Table 1 Characteristics of the regional distributions of mean hourly intensity.

Variable - Saison	Abscisse de la rupture de pente	Rapport entre les 2 pentes	Nombre de valeurs dans la distribution	Pourcentage des intensités > 4 μ
HMAP - HIVER	Rupture de pente à 4 μ	3	5 700	0,45 %
HMAP - ÉTÉ	Pas de rupture de pente	1	4 423	0,95 %
HMAO - HIVER	Pas de rupture de pente	1	31 472	0,95 %
HMAO - ÉTÉ	Rupture de pente à 4 μ	2	18 888	1,95 %

L'intérêt de cette étude est de caler la loi exponentielle sur un échantillon moins sensible à la présence de valeurs exceptionnellement élevées, tout en les prenant en compte de façon « régionale » par la modélisation de la rupture de pente. Cette façon de procéder permet d'augmenter la robustesse de l'ajustement des intensités moyennes horaires, et donc du modèle.

Lors de la simulation, la nouvelle loi des intensités apporte une amélioration sensible sur la restitution des moyennes et des écarts-types des variables « tests », en particulier sur les postes présentant des valeurs extrêmes d'intensité moyenne d'averse (comparaison des graphiques de droite de la figure 3 et des graphiques de gauche de la figure 7). On observe, de façon générale, une amélioration croissante du modèle avec l'utilisation de lois de moins en moins sensibles aux problèmes d'échantillonnage des variables HMA.

Une remarque peut être faite sur la signification des ruptures de pente observées. La première rupture de pente observée est celle de la variable HMAP de la saison « hiver », provoquée par 26 valeurs sur l'ensemble des 50 postes. On constate que si l'on analyse ces intensités avec les intensités des averses principales de la saison « été » (dont la moyenne est globalement 2 fois supérieure à celle d'« hiver »), elles perdent tout caractère exceptionnel et se mêlent à la distribution « d'été » sans créer de rupture de pente. De même, la seconde rupture de pente observée pour les intensités des averses ordinaires en « été », est provoquée par 368 valeurs (sur les 18 888) qui perdent aussi leur caractère exceptionnel si on les analyse avec les intensités des averses principales de la même saison (dont la moyenne est globalement 3 fois supérieures à celle des ordinaires). Il est résulte que ces ruptures de pentes observées peuvent être dues soit au découpage saisonnier strict (la typologie des pluies d'été pouvant avoir été observée en *hiver*), soit au choix d'une seule averse principale par événement (une seconde averse de forte intensité ne pouvant pas être retenue aussi comme principale). La modélisation des ruptures de pente peut alors servir à prendre en compte les incertitudes liées à l'analyse des événements.

4.2 Prise en compte de la liaison Intensité-Durée

L'hypothèse de base des modèles stochastiques suppose que toutes les variables descriptives sont indépendantes les unes des autres. Cette hypothèse a été préalablement vérifiée par le calcul de coefficients de corrélation entre les différentes variables (ANDRE (1991), CERNESSON (1993), TOURASSE (1981)). Du fait de la faiblesse des coefficients de corrélation linéaires (r^2 souvent inférieur à 0,2), l'hypothèse d'indépendance des variables descriptives du modèle n'a pas, jusqu'à présent, été remise en cause.

Cependant une faible liaison, peu linéaire, a été observée entre les variables durée (DA) et intensité moyenne (HMA) des averses. Cette liaison a été modélisée par VUILLET (1993) par une « courbe enveloppe » imposant des contraintes sur des valeurs ne pouvant être générées ensemble (graphique a de la figure 5). La modification apportée n'a pas entraîné d'amélioration du modèle.

La modélisation d'une courbe enveloppe reste une manière peu précise pour circonscrire le domaine de deux variables. En effet, l'occupation du domaine de définition des deux variables étudiées va dépendre de la fonction de répartition de chacune d'entre elles. Pour s'affranchir de leur fonction de répartition, il faut alors étudier l'indépendance ou la liaison entre deux variables en terme de proba-

bilité d'apparition comme le suggère le théorème de Bayes : deux variables sont considérées comme indépendantes si la probabilité d'apparition de l'une ne dépend pas de la probabilité d'apparition de l'autre. Générer alors les deux variables indépendamment et utiliser la courbe enveloppe pour rejeter des tirages incompatibles est un procédé source d'artefact. C'est pourquoi il nous est apparu original de mettre en œuvre une approche directement fondée sur la dépendance de la probabilité d'apparition des variables.

Le graphique *b* de la figure 5 reprend les valeurs du graphique *a* de la figure 5 après leur mise en fréquence. Il représente donc la probabilité d'apparition de la durée d'une averse en fonction de la probabilité d'apparition de son intensité. Sur cet exemple, on observe la dépendance des deux variables par la répartition des points autour d'une des diagonales, qui suggère une liaison négative entre la durée et l'intensité. Les valeurs de forte probabilité d'une variable sont préférentiellement associées aux valeurs de faible probabilité de l'autre variable, et inversement. Il reste à déterminer une loi permettant de quantifier puis de modéliser la liaison entre les fréquences de deux variables.

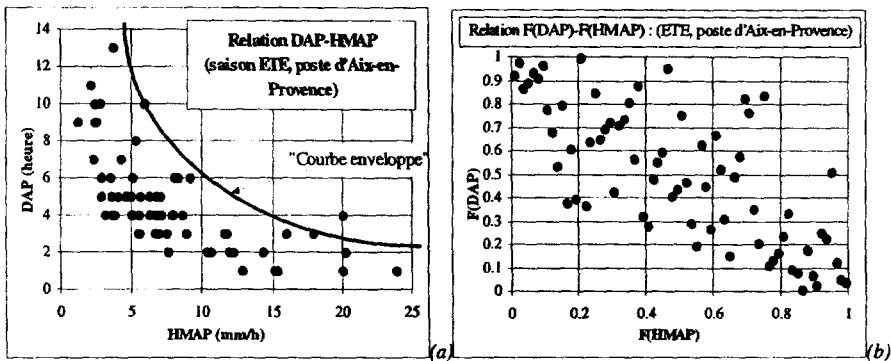


Figure 5 Mise en évidence de la dépendance des variables DAP (durée) et HMAP (intensité moyenne) du poste d'Aix-en-P., sur les valeurs brutes et sur leur probabilité d'apparition.

Observation of the dependence between duration and mean hourly intensity variables, on their values and on their cumulative frequencies.

On part de la loi de la somme de deux variables aléatoires indépendantes : soit deux variables aléatoires X et Y , continues dans l'intervalle $]-\infty ; +\infty[$, de densités de probabilité respectives f_1 et f_2 , soit une variable définie par $Z = X + Y$, alors Z est une variable aléatoire dont la densité de probabilité est définie par :

$$f(z) = \int_{-\infty}^{+\infty} f_1(x)f_2(z-x)dx = \int_{-\infty}^{+\infty} f_1(z-y)f_2(y)dy \tag{3}$$

Si l'on prend l'exemple simple de la somme de deux variables uniformes indépendantes, variant dans l'intervalle $]0 ; 1[$, on trouve la loi de la densité de probabilité de Z et sa fonction de répartition par intégration (DAGNELIE (1973)) :

$$f(z) = \begin{cases} z & \text{si } z \in]0 ; 1[\\ 2 - z & \text{si } z \in]1 ; 2[\end{cases}$$

$$F(z) = \int_D f(z)dz = \begin{cases} \frac{z^2}{2} & \text{si } z \in]0 ; 1] \\ 2z - \frac{z^2}{2} - 1 & \text{si } z \in]1 ; 2] \end{cases} \quad (4)$$

Si deux variables aléatoires (par exemple *DAP* et *HMAP*) sont indépendantes, leurs probabilités de non-dépassement ($F(DAP)$ et $F(HMAP)$) suivent une loi uniforme et sont également indépendantes. Ainsi la somme de $F(DAP)+F(HMAP)$ doit suivre la loi de probabilité $F(z)$.

Pour vérifier l'indépendance de deux variables aléatoires, on trace alors la distribution de la somme de leur probabilité d'apparition respective. La figure 6 reprend l'exemple de la saison « été » sur le poste d'Aix-en-Provence. On s'aperçoit de l'écart entre la distribution observée et la loi théorique ($n = 2$) représentant l'indépendance des deux variables. On cherche alors à ajuster la distribution observée par une loi permettant de prendre en compte la liaison observée. Cette loi doit être définie entre $]0 ; 2[$, prendre des valeurs entre $]0 ; 1[$ et avoir une intégrale égale à 1.

On introduit alors dans la formulation (4) un paramètre n qui permet d'ajuster une courbe théorique sur la distribution observée de la variable Z , et qui caractérisera la dépendance entre les variables X et Y :

$$F(z) = \begin{cases} \frac{z^n}{2} & \text{si } z \in]0 ; 1] \\ 1 - \frac{(2-z)^n}{2} & \text{si } z \in]1 ; 2] \end{cases}$$

soit
$$\begin{cases} z = [2 \cdot F(z)]^{\frac{1}{n}} & \text{si } F(z) \in]0 ; 0,5] \\ z = 2 - [2 - 2 \cdot F(z)]^{\frac{1}{n}} & \text{si } F(z) \in]0,5 ; 1] \end{cases} \quad (5)$$

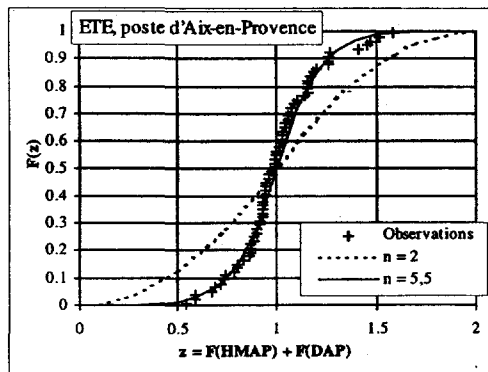


Figure 6 Ajustement de la fonction de répartition de la somme des probabilités d'apparition des durées et des intensités moyennes d'averse, caractérisant le degré de dépendance des deux variables.
Cumulative distribution function of the sum of duration and mean hourly intensity cumulative probabilities (characteristic of the dependence of the two variables).

Pour $n = 2$ on retrouve la loi de la somme de deux variables aléatoires uniformes indépendantes (4). Pour des valeurs de n inférieures à 2, on aura une liaison positive, c'est-à-dire que les fortes valeurs d'une variable seront préférentiellement associées aux fortes valeurs de la seconde variable. Pour des valeurs de n supérieures à 2 on aura par contre une liaison négative, c'est-à-dire que l'on aura préférentiellement l'association de valeurs fortes d'une variable avec les valeurs faibles de l'autre variable (exemple du poste d'Aix-en-Provence où l'ajustement optimal sur la distribution observée est obtenu pour $n = 5,5$). L'optimisation du paramètre n est réalisée de façon à minimiser le carré des écarts entre les valeurs des distributions théorique et observée.

Pour chaque poste pluviographique étudié, pour chaque saison (« été » et « hiver ») et pour chaque type d'averse (principale et ordinaire), la valeur du paramètre n a été calée. Ce nouveau paramètre permet de prendre en compte la dépendance observée entre les valeurs des durées et des intensités des averses. On génère une valeur de z en tirant dans sa loi de répartition $F(z)$ définie par la formule (5). On obtient alors la somme de la probabilité d'apparition des durées et des intensités : $z = F(DA) + F(HMA)$. On tire ensuite de façon aléatoire la probabilité d'apparition de l'une des variables (on a choisi HMA) et on en déduit celle de l'autre variable (DA) à partir de la valeur de z . Si la valeur de $F(DA) = z - F(HMA)$ n'appartient pas à son domaine de définition $]0 ; 1[$, on procède à un tirage de la valeur de $F(HMA)$. La valeur de z n'étant jamais retirée, on est sûr de respecter la liaison. Les distributions des variables HMA et DA sont aussi respectées tant que la liaison à modéliser n'est pas trop importante (majorité des cas). Lorsque la liaison est trop importante, un léger biais peut être introduit dans la distribution de fréquence des variables HMA et DA, dû à la difficulté de trouver des valeurs compatibles pour certaines fréquences.

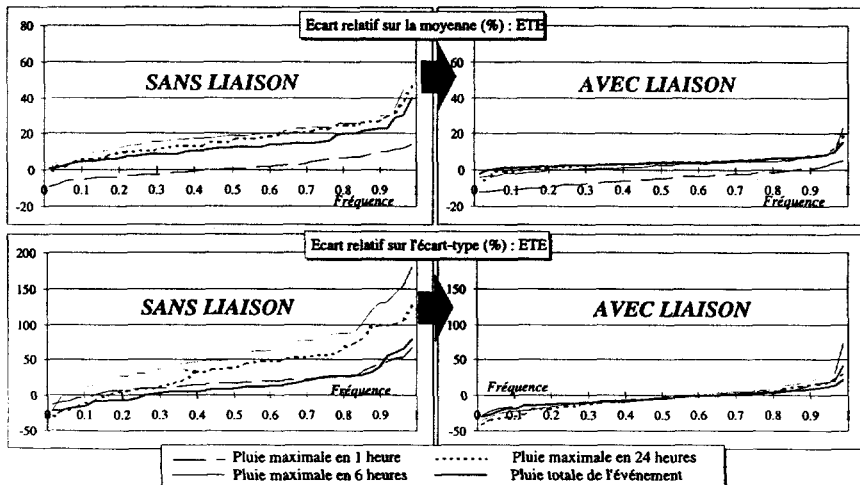


Figure 7

Évolution des écarts relatifs sur la moyenne et l'écart-type des variables « tests », lors de la prise en compte de la liaison entre les durées et les intensités moyennes des averses (saison été).

Evolution of relative difference on the mean and the standard deviation of « test » variables, after modelling the dependence between storms duration and intensity (summer period).

La prise en compte de la liaison entre les durées et les intensités des averses améliore considérablement les performances du modèle (figure 7 pour la saison « été », qui était la saison la moins bien restituée, et pour laquelle la liaison apparaît comme plus forte). L'amélioration est également plus importante pour les pluies maximales de plusieurs heures qui étaient elles aussi moins bien restituées que les pluies en une ou deux heures.

Si la prise en compte de la liaison entre durée et intensité moyenne des averses amène à une diminution des erreurs relatives sur la restitution des variables « tests », elle est surtout intéressante car elle entraîne une homogénéisation des erreurs relatives des différents postes. À ce stade, le modèle restitue correctement les moyennes et les écarts-types des variables « tests » pour les deux saisons.

4.3 Modélisation de la persistance des averses

Le modèle a des difficultés à générer de façon rare mais pas exceptionnelle de très forts cumuls de pluie sur une journée, bien que le cumul de pluie sur la durée totale de l'événement soit correctement généré. En effet, le modèle ne modélise pas la persistance des fortes averses, nécessaires à la génération des forts cumuls journaliers. C'est plutôt par la combinaison d'une forte intensité avec une longue durée pour l'averse principale, qu'il génère les forts cumuls journaliers.

Une analyse qualitative d'événements présentant de forts cumuls journaliers (comme celui observé au Col de Bavella en Corse : près de 800 mm en 24 h), permet de s'apercevoir qu'ils sont principalement dûs à la succession de nombreuses averses regroupées et présentant de fortes intensités. Ce phénomène de persistance des averses a été observé de façon quantitative sur la plupart des événements présentant de forts cumuls de pluies (en particulier pour les postes d'altitude). Elle se traduit par l'observation de deux effets :

- au sein d'un même événement pluvieux, les averses les plus fortes sont regroupées entre elles. La position des averses n'est pas aléatoire, mais laisse apparaître une certaine structure temporelle.
- sur l'ensemble des événements pluvieux, les averses fortes sont préférentiellement observées dans un même événement. Ainsi un fort cumul de pluie sur quelques heures peut être consécutif à une agglomération d'averses d'intensités importantes sans être exceptionnelles.

À partir de ces observations, on a modélisé une persistance par le regroupement des plus fortes averses ordinaires autour de l'averse principale, et en les plaçant au sein de la période pluvieuse ayant le plus grand nombre d'averses. La figure 8 montre l'influence de cette modification sur la distribution de fréquence des pluies maximales en 1 et 24 h simulées sur une période de 1 000 ans, à partir des paramètres du poste de Néoulous (P.O.).

La persistance est alors paramétrée par une fréquence, définie comme la fréquence seuil de la variable HMAP, au-delà de laquelle on modélise le regroupement des averses. Une fréquence de 1 indique que la persistance n'est jamais modélisée, et par exemple, une fréquence de 0,9 indique que la persistance est modélisée chaque fois que la probabilité de non-dépassement de la variable HMAP est supérieure à 0,9. Afin de faciliter une future paramétrisation régionale basée sur l'unique information journalière, la fréquence de cette modélisation est déterminée par une relation simple avec le coefficient de variation de la pluie jour-

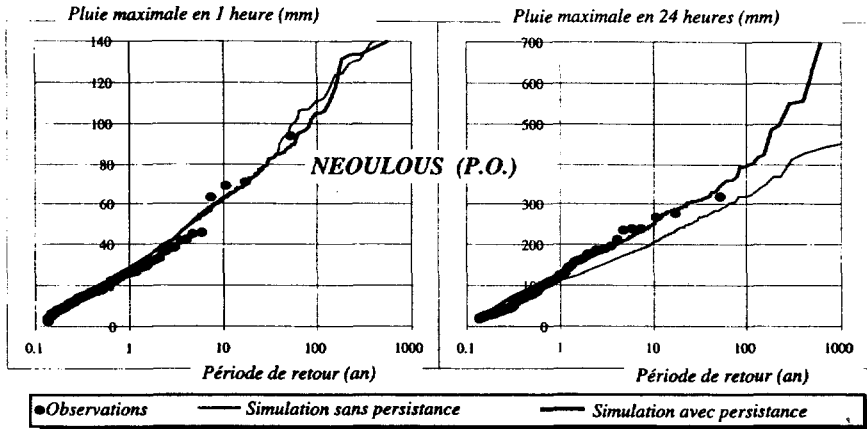


Figure 8 Distributions de fréquence des pluies maximales en 1 et 24 h observées et issues d'une simulation de 1 000 ans, avec et sans la modélisation de la persistance.

Cumulative distribution functions of maximum rainfalls over 1 and 24 hours, observed and simulated (1000 years), before and after persistence modeling.

nalière maximale. En effet, on sait que les cumuls de pluies importants apparaissent en particulier pour les postes d'altitude où une perturbation peut se trouver bloquée plusieurs heures voire la journée complète. C'est donc sur ces postes, présentant un fort coefficient de variation de la pluie journalière maximale, que la modélisation de la persistance aura un intérêt certain.

Sur les postes concernés, cette modification a une influence nulle sur les pluies maximales de 1 à 4 h ainsi que sur les pluies totales de l'événement. La prise en compte de la persistance des averses n'intéresse, principalement, que les pluies maximales de 6 à 24 h. De plus cette modification n'étant effectuée que dans certains cas (les plus propices à générer de forts cumuls de pluie), elle influence surtout les écarts-types et peu les moyennes qui étaient d'ailleurs bien restituées.

5 – BILAN DES AMÉLIORATIONS APPORTÉES AU MODÈLE ET DISCUSSION

La première modification ayant amélioré la restitution des variables tests est le changement de loi de probabilité pour les intensités moyennes des averses. Le choix d'une loi moins sensible aux problèmes d'échantillonnage, réduit les problèmes de surestimation de la variable par la présence de valeurs extrêmes dont les fréquences semblent surestimées. Ainsi, la loi exponentielle simple calée sur les valeurs inférieures à 4 fois la moyenne de la variable, améliore les performances du modèle.

La modification la plus notable est la prise en compte de la liaison existant entre l'intensité et la durée des averses. En effet, elle améliore considérablement la restitution des moyennes et des écarts-types des variables tests les moins bien restituées. Cette modification est d'autant plus intéressante qu'elle permet d'obtenir une certaine homogénéité des erreurs sur l'ensemble des postes. Ainsi la prise en compte d'une éventuelle dépendance entre les variables durée et intensité semble être la particularité longtemps ignorée du caractère différenciant des postes étudiés.

Enfin, la modélisation des problèmes de persistance au niveau du regroupement des averses pour générer de forts cumuls en 24 h, a permis de réduire la sous-estimation des écarts-types des pluies maximales en 24 h de certains postes d'altitude.

Pour avoir une vision globale de l'évolution du modèle suite aux différentes modifications, la figure 9 présente les moyennes et les écarts-types des pluies maximales en 1 et 24 h, issus du modèle initial et du modèle final (saisons confondues), comparés aux valeurs observées.

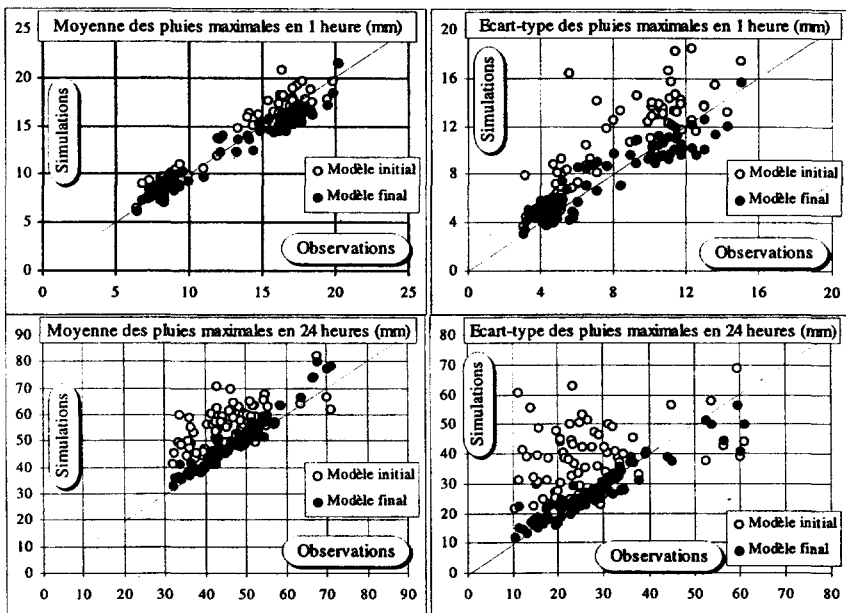


Figure 9 Évolution globale du modèle : moyennes et écarts-types des pluies maximales en 1 et 24 h, issus du modèle initial et du modèle final, comparés aux valeurs observées.

Global evolution of the model : comparison of mean and standard deviation of maximum rainfalls over 1 and 24 hours observed and simulated.

La pluie maximale en 1 heure qui était la variable la mieux restituée par le modèle initial, est la moins touchée par les modifications. Par contre la restitution des pluies maximales sur plusieurs heures a été considérablement améliorée, en particulier les écarts-types qui étaient soumis à de très fortes erreurs. L'extension

du modèle sur une zone plus étendue que sa zone de conception a ainsi été réalisée. Les résultats du modèle sont largement acceptables sur l'ensemble des 50 postes étudiés.

On obtient ainsi un outil beaucoup plus robuste et validé sur une zone étendue au pourtour méditerranéen français, qui permet la génération de multiples formes de hyétoigrammes couvrant toute la gamme des fréquences. De plus, la répartition spatiale des postes, bien qu'hétérogène (car liée à la disponibilité des données), permet de couvrir une large gamme pluviométrique (pluie moyenne annuelle de 500 à 1 500 mm), ainsi qu'un large éventail de typologie de pluie (pluie d'hiver ou pluie d'été, et pluviométrie de plaine ou d'altitude jusqu'à 1 100 m). Son aptitude à générer des pluies observées en climat méditerranéen, fortement variables, nous conforte dans l'idée de son application sur une zone beaucoup plus étendue.

La génération de hyétoigrammes permet de conserver au maximum l'information temporelle des pluies. Elle permet une approche différente de la détermination des courbes Intensité-Durée-Fréquence (*figure 10*) et sûrement plus fiable notamment au niveau des parties supérieures des distributions de fréquences ajustées. En effet, la simulation, qui n'est autre que la génération de hyétoigrammes de mêmes caractéristiques statistiques que les hyétoigrammes observés, enrichit les échantillons sur lesquels peuvent être réalisés les ajustements.

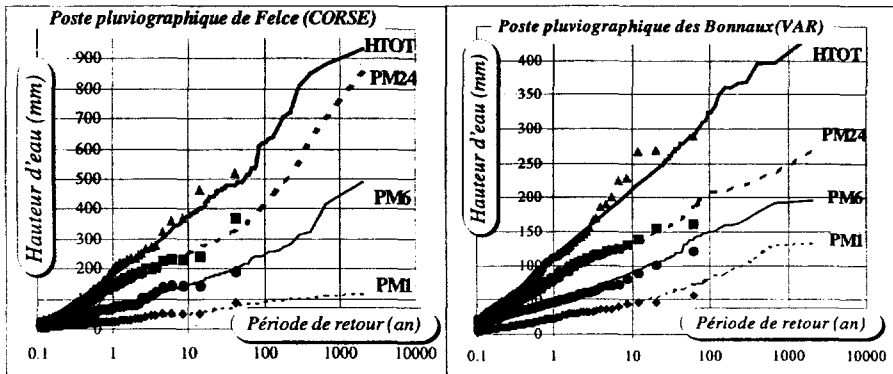


Figure 10 Distributions de fréquences des variables « tests », issues des observations (symboles) et issues d'une simulation de 1 000 ans (courbes) pour les postes de Felce (Corse) et des Bonnaux (Var).

Cumulative distribution functions of « test » variables, observed (symbols) and simulated (curves) at Felce station and Bonnaux station.

De plus, la génération d'une multitude de hyétoigrammes, permet de s'affranchir d'une pluie de projet unique. Si l'on reprend l'exemple du poste de Felce, on détermine par la simulation une pluie maximale en 24 h (PM24) décennale de 250 mm. Une simulation de 1 000 années nous a permis de sélectionner 16 événements dont la PM24 est proche de la valeur décennale (entre 245 et 255 mm). On observe alors sur les graphiques *a*, *b* et *c* de la figure 11 différents hyétoigrammes dont la PM24 est décennale.

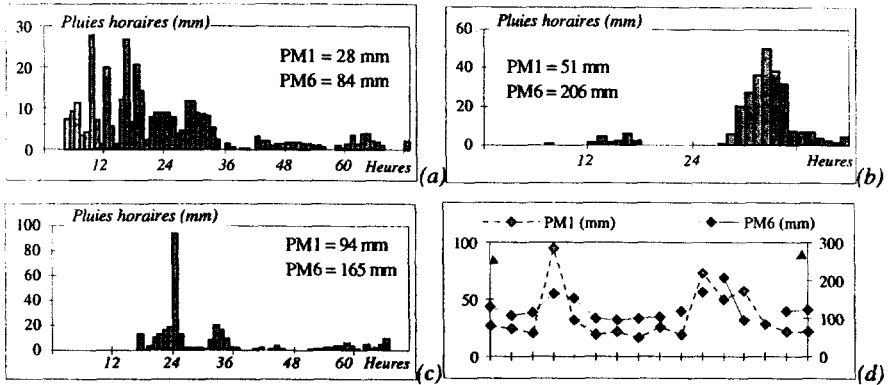


Figure 11 Hyétogrammes simulés dont la pluie maximale en 24 h (PM24) est décennale.

Simulated hyetographs characterized by decennial maximum rainfall of 24 hours.

Le graphique *d* de la figure 11, montre la diversité des PM1 et PM6 pouvant être associées à un hyétogramme dont la PM24 est décennale. Ce graphique montre bien la difficulté d'associer une fréquence à un hyétogramme unique caractérisé par une seule variable. Cette considération sur les caractéristiques temporelles des hyétogrammes prend son importance lors que l'on passe d'une pluie de projet à une crue de projet, la forme de la crue ayant une grande importance par exemple dans la simulation du comportement hydraulique d'un ouvrage.

CONCLUSION

L'étude approfondie d'un modèle stochastique aboutit d'abord à une constatation. La garantie de bons résultats nécessite d'une part le respect des caractéristiques des variables descriptives entre le phénomène observé et celui simulé, ainsi que le respect de leur éventuelle dépendance. Ce dernier point souvent négligé dans les modèles stochastiques s'est pourtant révélé d'une grande importance dans le modèle développé ici. La dépendance entre les durées et les intensités moyennes des averses est d'ailleurs un problème présent aussi dans les modèles basés sur les processus d'agrégation, où ONOF *et al.* (1995) relèvent le problème d'une dépendance entre les durées et les intensités des cellules dont il faudrait tenir compte.

La modélisation de hyétogrammes présente l'avantage de fournir de multiples scénarios respectant la diversité et l'information temporelle des pluies, ainsi que de proposer une nouvelle approche du comportement à l'infini des pluies, qui semble supérieur à une tendance strictement exponentielle.

De plus, l'analyse des pluies est effectuée sur un échantillon très important de variables. On étudie en effet, suivant les variables, des échantillons dont l'effectif varie de 100 à plus de 1 000 valeurs pour chaque poste. L'analyse d'événements observés en moyenne plusieurs fois par an, pour les reproduire statistiquement, semble être une méthode moins sensible aux problèmes d'échantillonnage et plus rapidement fiable avec un minimum d'années d'observations, qu'une approche statistique simple basée généralement sur l'ajustement de valeurs maximales annuelles.

On dispose maintenant d'un modèle de génération de hyétoigrammes horaires, correctement calé sur une cinquantaine de postes, et qui continue d'être testé sur de nouvelles données du pourtour méditerranéen français. Il paraît alors intéressant d'effectuer une cartographie de ses paramètres pour pouvoir générer des hyétoigrammes horaires en tout point de la zone d'étude. Cependant, le faible nombre de pluviographes disponibles et leur mauvaise répartition spatiale rendent impossible une cartographie directe des paramètres. C'est la raison pour laquelle une étude régionale de la paramétrisation du modèle de génération de hyétoigrammes horaires est étudiée à partir de la connaissance exclusive de la pluviométrie journalière. L'étude de la sensibilité du modèle à ses différents paramètres, a permis de fixer certains d'entre eux, peu variables sur la zone d'étude et peu influents sur les performances du modèle. La variabilité de la pluie ne repose alors plus que sur 5 paramètres majeurs à déterminer par saison à partir de l'information contenue dans la pluie journalière (ARNAUD 1997). L'information journalière plus largement disponible sera alors introduite dans un Système d'Information Géographique pour réaliser la cartographie des paramètres du modèle de génération de pluies horaires.

RÉFÉRENCES BIBLIOGRAPHIQUES

- ACREMAN M.C., 1990. A simple stochastic model of hourly rainfall for Farnborough. *Hydrolog. Sciences Journ.*, 35(2), 119-148.
- ANDRE G., 1991. Modèle stochastique de construction de hyétoigrammes horaires. Rapport de 3^e année de l'Institut des Sciences et Technologies de Montpellier, Sciences et Technologies de l'Eau.
- ARNAUD P., 1997. Modèle de prédétermination de crues basé sur la simulation - Extension de sa zone de validité, paramétrisation horaire par l'information journalière et couplage des deux pas de temps. Th. Doct. Univ. Montpellier II, 286 p.
- CERNESSON F., 1993. Modèle simple de prédétermination des crues de fréquences courantes à rares sur petits bassins versants méditerranéens. Th. Doct. Univ. Montpellier II, 240 p.
- CERNESSON F., LAVABRE J., MASSON J.M., 1996. Stochastic model for generating hourly hyetographs. *Atmospheric Research*, vol. 42, n°1-4 (1996), 149-161.
- COWPERTWAIT P.S.P., 1991. Further developments of the Neyman-Scott clustered point process for modelling rainfall. *Water Resour. Res.*, 27(7), 1431-1438.
- DAGNELIE P., 1973. Théorie et méthodes statistiques. Volume 1 et 2. Les Presses Agronomiques de Gembloux-Belgique, 378 p. et 451 p.
- ECONOPOULY T.W., DAVIS D.R. WOOLHISER D.A., 1990. Parameter transferability

- for a daily rainfall disaggregation model. *Journal of Hydrology*, 118 (1990) 209-228.
- GEORGAKAKOS K.P., BRAS R.L., 1984. A hydrologically useful station precipitation model: 1 - Formulation. *Water Resour. Res.*, 20(11), 1585-1596.
- KOUTSOYIANNIS D., 1994. A stochastic disaggregation method for design storm and flood synthesis. *Journal of Hydrology* 156 (1994) 193-225.
- LEBEL T., 1984. Moyenne spatiale de la pluie sur un bassin versant : estimation optimale, génération stochastique et gradex des valeurs extrêmes. Th. Institut National Polytechnique de Grenoble 350 p.
- LE CAM L. 1961. A stochastic description of precipitation. *Preprints : Fourth symp. on Mathematical Statistics and Probability, University of California, Berkeley*, 3, 165-186.
- MASSON J.M., 1991. Un problème parmi d'autres dans l'analyse des distributions des variables hydrologiques : les horsains (outliers). *Statistiques Appliquées, Séminaire 5*, ed. ORSTOM.
- ONOF C., CHANDLER R., KAKOU A., NORTHROP P., 1995. Rainfall modelling using Poisson-cluster process. Recueil de publications à l'International Conference in honour of Jacques BERNIER, Chap. II, PARIS, UNESCO, 11-13 sept. 1995.
- RODRIGUEZ-ITURBE I., COX D.R., ISHAM F.R.S et V., 1987. Some models for rainfall based on stochastic point processes. *Proc. Royal Society London*, A410, 269-288.
- RIBERO E., 1992. Eléments de régionalisation. Modèle stochastique de génération de hyétoigrammes horaires. Mémoire Mastère ENGREF, Cemagref d'Aix-en-Provence, 53 p.
- ROSSI F., FIORENTINO M., VERSACE P., 1984. Two component extreme value distribution for flood frequency analysis, *Water Resour. Res.*, 20(7), 847-856.
- TOURASSE P., 1981. Analyses spatiales et temporelles des précipitations et utilisation opérationnelle dans un système de prévision de crues. Application aux régions cévenoles. Th. IMG, Univ. scientifique et médicale et Institut National Polytechnique de Grenoble, 190 p.
- VUILLET M.F., 1993. Analyse d'un modèle stochastique de simulation de hyétoigrammes horaires. Rapport de 3e année, Institut des Sciences de l'Ingénieur de Montpellier, Sciences et Technologies de l'Eau. 44 p.
- WAYMIRE E., GUPTA V.K., 1981. The mathematical structure of rainfall representations : a review of the stochastic rainfall models. *Water Resour. Res.*, 17(5), 1261-1272.
- WOOLHISER D.A., OSBORN H.B., 1985. A stochastic model of dimensionless thunderstorm rainfall. *Water Resour. Res.*, 21(4), 511-522.