

UNIVERSITÉ PARIS VI
et
ECOLE NATIONALE DU GÉNIE RURAL

THÈSE

présentée par

Keltoum Chaouche

pour obtenir le grade de

Docteur de L'ENGREF

en

Hydrologie stochastique

**Approche multifractale de la modélisation
stochastique en hydrologie**

Soutenue publiquement le 4 Janvier 2001 devant le jury suivant:

M. Jean-Noel BACRO
M. Philippe BOIS
M. Pierre HUBERT
M. Gabriel LANG
M. Christian ONOF
M. Daniel SCHERTZER

Le mémoire de thèse de Keltoum Chaouche a été approuvé par:

Professeur Ph. Bois

Date

Professeur C. Onof

Date

Université Paris VI

Janvier 2001

Résumé

La plupart des séries d'observations hydrologiques possèdent des caractéristiques peu communes (grande variabilité sur une large gamme d'échelles spatiales et temporelles, périodicité, corrélation temporelle à faible décroissance), difficiles à mesurer (séries tronquées et intégrées sur des pas de temps qui ne respectent pas la nature du phénomène) et compliquées à intégrer dans un modèle stochastique classique (ARMA, Markov). Le modélisateur doit aussi faire face aux problèmes liés à l'échelle : en hydrologie (et en météorologie) où les données sont issues de pas de temps très divers, il est particulièrement intéressant de disposer de modèles à la fois capable d'intégrer des données à pas de temps différents et de fournir des résultats à un pas de temps différent de celui des entrées (désagrégation ou agrégation de séries temporelles).

Ce travail de thèse explore les possibilités d'application d'un nouveau type de modèle, les modèles multifractals, qui traduisent le plus simplement possible des propriétés d'invariance de certains paramètres (invariance spatiale, temporelle ou spatio-temporelle). En particulier, dans les modèles de cascades multifractales de générateur algébrique, c'est le coefficient de décroissance algébrique qui est un invariant d'échelle. Des résultats issus de la théorie probabiliste des valeurs extrêmes sont dès lors très utiles pour estimer ce paramètre. L'élaboration d'un outil statistique capable de détecter un comportement algébrique et d'estimer le paramètre de décroissance algébrique constitue une étape préalable au développement de cette thèse. Il est aussi montré dans ce travail, que la forme de dépendance (longue ou courte) des modèles en cascade multifractale diffère selon le générateur de la cascade.

L'étude exploite une base de données de 232 séries annuelles de divers sites, de nombreuses séries de pluie à divers pas de temps (mois, jour, heure, minute) ainsi que quelques séries de débits. Elle conduit aux résultats suivants :

- Les lois de type algébrique sont adaptées à la modélisation des grandes périodes de retour des séries de pluie étudiées.

- Sur ces mêmes séries, le coefficient de décroissance algébrique est un paramètre invariant d'échelle (sur des gammes d'échelles supérieures à l'heure).

- L'estimation de ce coefficient en divers sites à travers le monde est très peu variable.

- La propriété de longue dépendance est décelable au sein de certaines séries de débits, notamment des séries de rivières sur craie.

Ces résultats incitent donc à l'emploi de cascades multifractales pour la modélisation des séries de pluie, bien qu'un travail concernant la détection et l'estimation de longue dépendance reste à accomplir pour que le choix du générateur respecte la forme de dépendance de la série.

Table des matières

Liste des Figures	iv
Liste des Tableaux	vii
1 Introduction	2
1.1 Problème général	2
1.2 Objectifs de la thèse	6
1.3 Démarche proposée	6
2 Champs de pluie et modélisation stochastique	8
2.1 Les champs de pluie	9
2.1.1 Mécanismes générateurs des champs de pluie	9
2.1.2 Mesure de la pluie	10
2.1.3 Echelles spatio-temporelles	12
2.1.4 Conclusion	13
2.2 La modélisation stochastique de la pluie	13
2.2.1 Pourquoi une modélisation stochastique ?	14
2.2.2 Méthodes statistiques d'estimation spatiale	14
2.2.3 Modèles temporels	20
2.2.4 Modèles spatio-temporels	22
2.2.5 Conclusion	24
2.3 Les modèles fractals	25
2.3.1 Les fractals	25
2.3.2 Application à quelques séries d'occurrence	32
2.3.3 Modèle en cascade multifractale	36
2.3.4 Revue bibliographique de modèles hydrologiques multifractals	39
3 Modélisation des extrêmes hydrologiques	41
3.1 Introduction	41
3.2 Notations	42
3.3 Rappels sur la théorie des valeurs extrêmes	43
3.3.1 Loi des extrêmes	44

3.3.2	Le domaine d'attraction de la loi de Fréchet	46
3.4	Estimation du comportement algébrique	51
3.4.1	Revue d'outils statistiques d'exploration	52
3.4.2	Revue d'outils d'estimation	59
3.4.3	Nouvel outil statistique	65
3.4.4	Conclusion	76
3.5	Application à des séries hydrologiques	76
3.5.1	Hauteurs journalières de pluie	77
3.5.2	Données horaires	80
3.5.3	Données de l'Ile de la Réunion	86
3.5.4	Application à 232 longues séries annuelles	86
3.5.5	Application à des séries de débits	88
3.5.6	Conclusion	91
3.6	Données à basculement d'auget	92
3.6.1	Distribution des fréquences de basculement	92
3.6.2	Application aux données	95
3.7	Conclusions	97
4	Les cascades multifractales et la longue dépendance	99
4.1	Notations	100
4.2	Définitions de la longue dépendance	100
4.3	La longue dépendance des séries hydrologiques	102
4.4	Le mouvement brownien fractionnaire	102
4.5	Les cascades sont-elles à longue dépendance ?	104
4.5.1	Définition de la cascade	104
4.5.2	Type de dépendance des cumuls	105
4.5.3	Cascades particulières	107
4.6	Application aux données	109
4.6.1	Outils statistiques	109
4.6.2	Séries simulées	111
4.6.3	Séries de débits	112
4.7	Conclusions	112
5	Conclusion	114
	Bibliographie	116
A	Comptage de boîtes sur l'Ensemble de Cantor	122
B	Caractérisation des lois de type algébrique	124
C	Calculs du biais et de la variance de \hat{h}	127
C.1	Biais	127
C.2	Variance	128

D	Auto-similarité et Multifractalité	129
E	Liste des stations	131

Liste des Figures

1.1	Effet de la troncature d'une série.	4
1.2	Sur-estimation des périodes de retour estimées par la loi normale.	5
2.1	Asymétrie des cumuls en fonction du pas d'intégration (série de Athènes). . .	15
2.2	Asymétrie de la distribution des cumuls pluviométriques.	15
2.3	Pépite, portée et palier d'un variogramme.	19
2.4	Processus de Neyman-Scott.	23
2.5	Construction de l'Ensemble de Cantor.	27
2.6	Graphes de la fonction de Weirstrass pour s variable.	28
2.7	Mesure et dimension de Hausdorff.	30
2.8	Comptage de boîtes sur un Cantor simulé à 3^9 données, avec boîtes se dé- mplant.	31
2.9	Comptage de boites sur les données Badinage.	33
2.10	Comptage de boites sur la station 1.	34
2.11	Comptage de boites sur la série de la Réunion.	35
2.12	Cascade multifractale.	37
3.1	Densité de probabilité des lois de Gumbel, Fréchet et Weibull (Embrechts et al. [38]).	45
3.2	Quelques résultats sur le domaine d'attraction de Fréchet.	50
3.3	Comparaison des estimateurs de Weibull et standard de la fonction de survie.	54
3.4	Graphe de quantiles d'un 1000-échantillon de loi exponentielle comparé à une loi exponentielle (a), normale (b), pareto (c) et log-gamma (d).	55
3.5	57
3.6	Graphes des FMD de quelques lois (Embrechts et al. [38]).	57
3.7	FMD empirique sur des 5000-échantillons de loi (a) exponentielle, (b) Pareto et (c) log-gamma.	58
3.8	Estimateur de Hill pour N simulations de lois de fonction de survie $G(s) =$ $s^{-3.33}$ (ligne 'x') et $G(s) = (s + 1)^{-3.33}$ (ligne '+') pour : (a) $N = 500$, (b) $N = 1000$ et (c) $N = 5000$	59
3.9	Histogrammes des estimations de ξ, ψ et μ par les méthodes des moments pondérés et du maximum de vraisemblance dans un modèle GVE ($\xi = 0, \mu = 2.07, \psi = 0.45$)	

3.10	Histogrammes des estimations de β et ξ par les méthodes des moments et du maximum de vraisemblance dans un modèle $PG(\beta, \xi)$ sur 50 échantillons simulés de $PG(2, 0.31)$ de taille 500.	66
3.11	Histogrammes des estimations de β et ξ par les méthodes des moments et du maximum de vraisemblance dans un modèle $PG(\beta, \xi)$ sur 50 échantillons simulés de $PG(2, 0.31)$ de taille 1000.	67
3.12	Histogrammes des estimations de β et ξ par les méthodes des moments et du maximum de vraisemblance dans un modèle $PG(\beta, \xi)$ sur 50 échantillons simulés de $PG(2, 0.90)$ de taille 1000.	68
3.13	Durée et cumul de dépassement.	70
3.14	Variations du biais de la variance avec le seuil.	71
3.15	Estimation de h sur un 5000-échantillon de loi $Exp(0.45)$	73
3.16	Répartition des estimations de la $FMDR$ h au seuil maximal sur 50 8000-échantillons simulés de loi $Exp(0.45)$	74
3.17	Estimation de la $FMDR$ h et de q sur un 5000-échantillon simulé de loi $Al(3.2)$	74
3.18	Répartition des estimations de ξ par la $FMDR$ sur 50 échantillons simulés de loi $Al(3.2)$ ($\xi = 0.31$) de taille 1000.	75
3.19	Série de Dédougou (5% des plus hautes observations) : Estimation de q par la $FMDR$ et comparaison des estimations du MV et de la $FMDR$ par la fonction de survie ('+' : estimateur standard de la fonction de survie).	79
3.20	Estimation de q par la $FMDR$ sur les séries de Athènes et Bra.	81
3.21	Athènes et Bra sur 5% des plus hautes observations : Comparaison des estimations du MV et de la $FMDR$ de la fonction de survie comparées à la fonction de survie empirique.	82
3.22	Série de Larnaca (5% des plus hautes observations) : Estimation de q par la $FMDR$ et comparaison des estimations du MV et de la $FMDR$ par la fonction de survie ('+' : estimateur standard de la fonction de survie).	83
3.23	Série de Alabama (5% des plus hautes observations) : Estimation de q par la $FMDR$ et comparaison des estimations du MV et de la $FMDR$ par la fonction de survie ('+' : estimateur standard de la fonction de survie).	84
3.24	Série de Bordeaux (5% des plus hautes observations) : Estimation de q par la $FMDR$ et comparaison des estimations du MV et de la $FMDR$ par la fonction de survie ('+' : estimateur standard de la fonction de survie).	85
3.25	Série de la Réunion (5% des plus hautes observations) : Estimation de q par la $FMDR$ et comparaison des estimations du MV et de la $FMDR$ par la fonction de survie ('+' : estimateur standard de la fonction de survie).	87
3.26	Les 230 stations étudiées ainsi que la répartition de la durée des séries.	88
3.27	Boîte à moustaches et histogramme des estimations de q sur les 232 séries annuelles.	89
3.28	Effet de la longueur de la série sur l'estimation de q	89
3.29	Histogramme des estimations de q	90
3.30	Estimation de la fonction de survie de la série des grandes fréquences de basculement de la Réunion (troncature à 5%) : Méthodes du maximum de vraisemblance (MV) et de la $FMDR$	96

3.31	Estimation de la fonction de survie de la série des grandes fréquences de basculement EPSAT (troncature à 5%) : Méthodes du maximum de vraisemblance (MV) et de la FMDR.	98
E.1	Liste des stations des séries annuelles.	132

Liste des Tableaux

2.1	Organisation des champs de pluie	10
2.2	Résolution temporelle et spatiale de quelques satellites.	12
2.3	Ordre de grandeur des tailles et de la résolution des estimations de lames d'eau de chaque type de bassin versant.	13
2.4	Gammes de variation des dimensions fractales.	28
2.5	Comptage de boîtes sur l'Ensemble de Cantor pour plusieurs facteurs d'échelle a (N = nombre moyen de boîtes non vides).	31
2.6	Comptage de boîtes sur les données Badinage (P_1 = pente sur $[5min; 1h40min]$ et P_2 = pente sur $[1h40min; 1j10h]$).	33
2.7	Comptage de boîtes sur la station 1 (P = pente sur $[1h40, 1j10h]$ et m facteur d'échelle).	34
2.8	Méthode du comptage de boîtes sur la série de la Réunion (N = Nombre de basculement, D = Durée d'observation P = Pente sur $[2.5min; 6h]$).	36
2.9	Revue bibliographique des modèles multifractals	40
3.1	<i>FMD</i> pour quelques lois.	48
3.2	Moyennes et variances des estimations dans un modèle <i>GVE</i> ($\xi = 0, \mu = 2.07, \psi = 0.45$) (MV : Maximum de vraisemblance et MP : moments pondérés).	64
3.3	Estimations moyennes de β et ξ par les méthodes des moments et du maximum de vraisemblance dans un modèle $PG(\beta, \xi)$ sur 50 échantillons simulés de tailles 500 et 1000 de $PG(2, 0.31)$	65
3.4	Estimations moyennes de β et ξ sur 50 1000-échantillons simulés de $PG(2, 0.90)$ (par les méthodes des moments et du maximum de vraisemblance dans un modèle de Pareto généralisée).	65
3.5	Moyennes et variances des estimations de ξ par la <i>FMDR</i> sur 50 échantillons simulés de loi $Al(3.2)$ ($\xi = 0.31$) de tailles 500 et 1000.	75
3.6	Caractéristiques des stations.	91
3.7	Estimations du paramètre de décroissance algébrique \hat{q} sur les séries de débits (* : volume moyen écoulé par an (10^6 m ³)).	92
3.8	Récapitulatif des estimations obtenues.	93

4.1	Estimation par le log-périodogramme global sur des cascades simulées (* : nb=nombre d'échantillons simulés).	111
4.2	Estimations de la longue dépendance par le log-périodogramme global (* : volume moyen écoulé par an (10^6 m ³).	113

Remerciements

J'adresse mes remerciements à tous ceux et à toutes celles qui m'ont aidée, par leur encadrement et par leurs conseils scientifiques, ou par leur appui matériel, à réaliser ce mémoire de thèse.

J'exprime ma gratitude à mes deux directeurs de thèse, M. Pierre Hubert et M. Gabriel Lang. P. Hubert, directeur de la formation doctorale Hydrologie et Hydrogéologie Quantitatives de l'Ecole des Mines de Paris, a donné l'impulsion de départ à ce travail et, par ses conseils, m'a permis de prospecter de nouvelles voies. G. Lang, chercheur au GRESE (laboratoire de Gestion du Risque en Sciences de l'Eau) a encadré cette thèse et n'a pas ménagé son temps pour me relire ou pour discuter mes approches.

Je remercie M. Bacro pour l'aide qu'il m'a apportée pour la modélisation des extrêmes.

Je remercie M. Bois, professeur d'Hydrologie à l'ENSHMG et chercheur au LTHE (Laboratoire d Etudes des Transferts en Hydrologie et Environnement) ainsi que M. Onof, professeur de l'Imperial College de Londres qui me font l'honneur d'être les rapporteurs de ma thèse. Cette thèse n'aurait pas vu le jour sans Eric Parent, directeur du GRESE qui m'a accueillie et hébergée dans son laboratoire, sans M. de Marsily et MM. Degoutte et Miller, qui m'ont admise dans l'Ecole doctorale ; à tous ceux-là, j'exprime ma reconnaissance.

Je remercie enfin tous ceux qui ont fourni les données hydrologiques : MM. T. Lebel, A. Barcelo et A. Castéanu.