

Table des matières

Introduction	10
1 Analyse et traitement de données de phénomène pluvieux local. Choix de l'approche de modélisation	14
1.1 Analyse des données d'observation	15
1.2 Caractérisation de la dynamique des données de pluie	17
1.3 Distributions de probabilités	30
1.4 Conclusion du 1 ^{er} chapitre	38
2 Choix et mise en oeuvre de la modélisation de la dynamique du phénomène pluvieux	39
2.1 Première approche du processus pluvieux local	40
2.2 Seconde approche du processus pluvieux local	55
2.3 Conclusion du 2 nd chapitre	60
3 Mise en oeuvre du modèle régi par l'équation FFPE. Application à une simulation du fléchissement de la dynamique des pluies.	61
3.1 Approche stochastique de la mise en oeuvre du modèle	62
3.2 Apparaît-il un aspect d'un changement climatique en 1997?	67
3.3 Résolution numérique du modèle macroscopique	70
3.4 Conclusion du 3 ^{ème} chapitre	72
4 Estimation de l'impact des précipitations pluvieuses sur les fluctuations de la nappe phréatique.	77
4.1 Estimation des fluctuations de nappe corrélées avec les événements de pluie	78
4.2 Caractéristiques du signal de fluctuations de nappe $\psi(t)$	85
4.3 Modélisation du transport de la pluie dans le sol par le logiciel P.A.S.T.I.S.	86
4.4 Analyse des corrélations entre $\lambda(t)$, $\phi(t)$ et $\psi(t)$	89
4.5 Conclusion du 4 ^{ème} chapitre.	96
Conclusion	97

A Dimension fractale et corrélation entre accroissements passés et futurs d'une série	100
B Pluies et sécheresses suivent-elles un processus markovien ?	102
C Méthodes d'estimation de l'exposant de stabilité α	104
D Processus stochastique	109
E Données et mesures expérimentales	113
F Code de calcul PASTIS	115
G Transformation en ondelette	121

Table des figures

1.1	La série temporelle constituée du relevé des hauteurs de pluies sur 40 ans de 1968 à 2008.	17
1.2	Série des précipitations constituées des seules hauteurs de pluie. Sur les 40 années, Il n'apparait pas de changement flagrant des quantités de pluies.	18
1.3	Série des temps d'attente entre deux pluies. Il apparaît, à droite du graphique, une forte augmentation des temps d'attente qui traduit de longues périodes de sécheresse des années 1997 - 2008	18
1.4	Hauteurs de pluie cumulées des précipitations. On remarque évidemment un changement de dynamique vers 1997. Est-ce un indice de changement pérenne de la pluviométrie	19
1.5	Hauteurs de pluies cumulées par saisons.	20
1.6	Auto corrélation et auto information du signal global.	23
1.7	Auto corrélation et autoinformation du signal global en log-log.	23
1.8	Fonctions d'autocorrélation et d'autoinformation des hauteurs de pluie cumulées. .	24
1.9	Auto corrélation des hauteurs de pluie cumulées en coordonnées logarithmiques.	25
1.10	Portrait de phase des données de précipitations.	29
1.11	Portrait de phase du surrogate des données de précipitations.	30
1.12	Distribution des intensités de pluie et la distribution log-normale qui lui est ajustée (fittée).	33
1.13	Comparaison des distributions des intensités de pluie et de la log-normale en coordonnées logarithmiques.	34
1.14	Premier test graphique : évolution de la variance.	35
1.15	Estimation graphique de l'exposant de stabilité α	36
2.1	Distributions log-normales de \tilde{j}_n et j_n simulant les sauts de la marche aléatoire. .	42

2.2	Hauteurs de pluie cumulées simulées par des marches aléatoires non soumises aux caractéristiques de la série temporelle des pluies cumulées	43
2.3	Exemples de trajectoires de pluie cumulées simulées présentant des grands écarts par rapport à la série temporelle des pluies cumulées.	44
2.4	Hauteurs de pluie cumulées simulées par des marches aléatoires non tempérées, non "triées".	45
2.5	Hauteurs de pluie cumulées simulées par des marches aléatoires tempérées mais non "triées".	45
2.6	Hauteurs de pluie cumulées simulées par des marches aléatoires "Tempérées et triées".	46
2.7	1 ^{ière} figure, en haut : hauteurs de pluie cumulées simulées par des marches aléatoires "Tempérées, triées et sélectionnées", elles "voisinent" bien la trajectoire expérimentale. 2 ^{nde} figure, en bas : la trajectoire moyenne des simulations montre bien un certain fléchissement. $V=9$, $\alpha =0,85$ _ $D = 1$ _ $\lambda = 10^{-7}$ _ sélection entre 0,7 et 0,9 soit entre 70 et 90 % de temps d'attente supérieurs à 1 jour entre deux pluies. Temps entre 13000 et 14000 jours.	48
2.8	1 ^{ère} figure en haut : en rouge et vert apparaissent les simulations, "Tempérées, triées et sélectionnées", d'abscisses ; les temps d'attente α -stables d'une part et rapportées à des temps entiers d'autre part. En noir : trajectoire moyenne. En bleu : La courbe expérimentale. $V= 9$, $\alpha =0,85$, $D = 1$, $\lambda = 10^{-7}$, sélection entre 0,7 et 0,9, temps entre 13000 et 14000 jours. 2 ^{nde} figure en bas : vue plus "fine" portant sur 2 simulations.	49
2.9	Hauteurs de pluie cumulées simulées par des marches aléatoires "Tempérées, triées et sélectionnées", parmi les déciles ; trajectoire moyenne. $V=9$, $\alpha =0,85$, $D = 1$, $\lambda = 10^{-7}$, sélection entre 0,7 et 0,9, temps entre 13000 et 14000 jours.	50
2.10	P.D.F obtenu à partir de marches aléatoires "tempérées, triées et sélectionnées".	50
2.11	P.D.F obtenu à partir de marches aléatoires "tempérées, triées et sélectionnées".	51
2.12	P.D.F obtenu à partir de marches aléatoires "tempérées, triées et sélectionnées" extrapolées à l'an 2018.	51
2.13	Processus "tronqué" : Cumuls expérimentaux et cumuls les plus probables obtenus par marche aléatoire.	53

2.14 Processus "tronqué" : Ecarts-types des distances : courbe expérimentale- simulations par marche aléatoire "Tempérées, triées et sélectionnées"	53
2.15 Processus "tronqué" : Ecart-type "relatif" des distances entre la valeur expérimentale et les simulations de marches aléatoires.	54
 3.1 Trajectoires simulées et trajectoire moyenne obtenues à partir d'une discréétisation de l'équation différentielle stochastique d'Itô. Les temps sont "subordonnés, tempérés, sélectionnés". En vert : la trajectoire moyenne. Paramètres : $\alpha = 0.95$, $V = \mu/\tau$, $\eta = 0.5396$, $D = 1$, $\lambda = 10^{-7}$	64
3.2 Exemple de graphe permettant d'apprécier la hauteur de pluie cumulée la plus probable pour un temps donné, ici pour le 300 <i>ième</i> jour.	66
3.3 Comparaison des hauteurs de pluie les plus probables par le "processus tronqué" : (*) par des simulations de marches aléatoires. Paramètres : $V = 3.4$; $D = 1$; $\alpha = 0.9$. (+) par discréétisation de l'E.D.P.Stochastique de Îto. Paramètres : $V = 5$; $D = 1$; $\alpha = 0.9$	68
3.4 Comparaisons des hauteurs de pluies cumulées les plus probables avec les observations. Les paramètres des simulations sont donnés dans le tableau (3.4) . . .	69
3.5 Hauteurs de pluie cumulées les plus probables obtenues avec la résolution numérique de F.F.P.E. (Période 1968-2008)	70
 4.1 Relevés horaires des variations de hauteur de nappe phréatique et de précipitations.	79
4.2 Fonction de cohérence normalisée et déphasage entre λ , η et entre λ et le surrogate de η	82
4.3 Haut : Variation de hauteur de la nappe $\eta(t)$ (mesuré en m). Milieu : Intensité des précipitations $\lambda(t)$ (en mm/h). Bas : Fluctuations de la nappe $\psi(t)$ (in m) .	83
4.4 Fonction de cohérence normalisée et déphasage entre le signal de hauteur de nappe $\eta(t)$ et le signal de fluctuations $\psi(t)$	84
4.5 Fonction de cohérence normalisée et déphasage entre le signal de pluie $\lambda(t)$ et le signal de fluctuations de nappe $\psi(t)$	84
4.6 Fonction densité spectrale de puissance des fluctuations de la nappe phréatique .	86
4.7 Densité de probabilité du signal de fluctuations (symboles "o") et loi stable avec un exposant caractéristique $\alpha = 1.9$	87

4.8 Corrélation croisée entre précipitations et fluctuations de l'état hydrique	91
4.9 Densités spectrales de puissance des précipitations et des fluctuations de l'état hydrique à 1.6m	92
4.10 Corrélation croisée entre fluctuations de l'état hydrique à 1.6m et fluctuations de la nappe	93
4.11 Spectres de puissance des fluctuations de l'état hydrique à 1.6m et celles de la nappe	94
4.12 Corrélation croisée entre fluctuations de la nappe et précipitations	95
4.13 Spectres de puissance des précipitations et des fluctuations de la nappe	95
E.1 Schéma de l'instrumentation du site atelier "Fontanille". Les données issues du site 3C ont été utilisées dans ce mémoire. Elles sont équivalentes à celles obtenues sur 1C et 2L.	114
G.1 Ondelette de Morlet - (a) Forme de sa partie réelle (b) Ondelette de Morlet (en pointillé bleu), résultante d'une cloche Gaussienne (en rouge) et d'une onde plane (en vert)	122

Liste des tableaux

1.1	Valeurs des différents exposants de Hurst.	28
1.2	Estimation des paramètres de la loi stable des temps d'attente.	37
2.1	Le modèle approche d'autant mieux la réalité expérimentale que le nombre d'observations est grand.	52
3.1	Comparaison des hauteurs de pluie les plus probables (en mm) obtenues avec la marche aléatoire et l'équation différentielle stochastique d'Itô. Les écarts par rapport aux valeurs mesurées sont en %.	67
3.2	Comparaison des hauteurs de pluie les plus probables (en mm) obtenues avec la marche aléatoire et l'équation différentielle stochastique d'Itô. Les écarts par rapport aux valeurs mesurées sont en %.	67
3.3	Estimation des exposants de stabilité de la loi stable des temps d'attente avant et après 1997.	68
3.4	Paramètres des simulations par marches aléatoires.	69
3.5	Paramètres des simulations avec l'équation différentielle d'Itô.	69
3.6	Tableau de prévision de la hauteur cumulée la plus probable en 2018	71
4.1	Statistiques des niveaux de la nappe en trois points différents sur le site	78
4.2	Statistiques des événements de pluies.	79
A.1	Les dimensions fractales sont inférieures à celles d'une série normale ce qui induit une certaine persistance, sauf pour les pluies cumulées où la très forte persistance est due à la construction de la série.	101
C.1	Méthode de McCulloch	105
C.2	Méthode de Koutrouvelis.	107

Introduction

Cette thèse porte sur la modélisation du processus pluvieux local et l'étude de son impact sur la dynamique des fluctuations de la nappe phréatique. Elle s'inscrit dans le cadre du projet "Sol Virtuel" du département "Environnement et Agronomie" de l'I.N.R.A. (Institut National de la Recherche Agronomique) qui porte sur le développement d'une plateforme de modélisation numérique des processus physiques, chimiques et biologiques se déroulant dans le sol et de leurs interactions. Une des finalités de ce travail concerne les recherches menées sur la qualité des ressources en eau, l'étude du processus de transfert de masse de la surface du sol vers la nappe phréatique constitue un élément primordial pour la compréhension de la pollution de cette dernière. En effet, les éléments polluants à la surface notamment ceux qui sont solubles (tels que les produits liés aux activités humaines comme les engrains, les pesticides, etc...) peuvent transiter vers la nappe à travers le milieu poreux que représente la couche de sol. Ce scénario de transfert de pollution repose sur deux phénomènes à savoir la pluie qui génère la masse d'eau à la surface et la dispersion de celle-ci à travers le milieu poreux. La dispersion de masse dans un milieu poreux naturel comme le sol forme un sujet de recherche vaste et difficile aussi bien sur le plan expérimental que théorique. Sa modélisation constitue une préoccupation importante du laboratoire EMMAH, en particulier dans le cadre du projet Sol Virtuel dans lequel un modèle de transfert PASTIS (Predicting.Agricultural.Solute.Transport in Soils) a été développé. Ce modèle permet de simuler le transport de masse d'eau à travers la couche du sol avec comme entrées principales la composition du sol, la couverture végétale, les caractéristiques du phénomène pluvieux et le taux d'évapo-transpiration au niveau du sol.

Ce travail de thèse aborde cet objectif en s'appuyant d'une part sur des résultats d'observations expérimentaux et d'autre part sur de la modélisation inspirée par l'analyse des données d'observation. La première partie du travail est consacrée à l'élaboration d'un modèle stochastique de pluie. Le choix et la nature du modèle sont basés sur les caractéristiques obtenus à partir de l'analyse de données de hauteur de pluie recueillies sur 40 ans (1968-2008) sur le

Centre de Recherche de l' INRA d'Avignon. Pour cela, la représentation cumulée des précipitations sera assimilée à une marche aléatoire dans laquelle les sauts et les temps d'attente entre les sauts représentent respectivement les amplitudes et les durées aléatoires entre deux occurrences d'événements de pluie. Ainsi, la loi de probabilité des sauts (loi log-normale) et celle des temps d'attente entre les sauts (loi alpha-stable) sont obtenus en analysant les lois de probabilité des amplitudes et des occurrences des événements de pluie. Nous montrons alors que ce modèle de marche aléatoire tends vers un mouvement brownien géométrique subordonné en temps (quand les pas d'espace et de temps de la marche tendent simultanément vers zéro tout en gardant un rapport constant) dont la loi de densité de probabilité est régie par une équation de Fokker Planck fractionnaire (FFPE). Deux approches sont ensuite utilisées pour la mise en oeuvre du modèle. La première approche est de type stochastique et repose sur le lien existant entre le processus stochastique issu de l'équation différentielle d'Itô et la FFPE. La deuxième approche utilise une résolution numérique directe par discrétisation de la FFPE. Cette première partie est déclinée en trois chapitres. Le premier chapitre est consacré à l'étude et à la caractérisation du relevé chronologique journalier des précipitations pluvieuses. Les démarches d'analyse proposées visent à fournir à partir des données des caractéristiques qui contribueront à définir une modélisation adaptée du phénomène pluvieux local. Les démarches d'analyse proposées sont regroupées en deux parties. D'une part, avec une approche d'analyse de système dynamique, nous avons étudié l'importance des non-linéarités et l'existence des "effets mémoires". D'autre part, en adoptant une approche plus probabiliste, nous avons découpé les données sous la forme d'une série d'intensité de pluies et de temps d'attente entre deux pluies qui représentent mathématiquement deux variables aléatoires indépendantes. En tenant compte des caractéristiques d'effets de mémoires à long terme, les distributions de probabilité des intensités et des temps d'attente ont été modélisées par des lois de probabilités à ailes épaisses. Le second chapitre portera sur le choix et la mise en oeuvre d'un modèle de marche aléatoire simulant la dynamique décrite par la série temporelle de précipitations cumulées. Ce modèle de marche aléatoire est adapté pour "coller " au mieux aux données qui constituent la seule réalisation du processus à déterminer. Plus précisément, les adaptations portent sur la distribution α -stable des temps d'attente qui produit des temps trop exagérément longs en regard aux temps effectivement constatés dans les données. Pour pallier cette surestimation on imposera alors des restrictions à ces distributions. De telles restrictions nous conduisent à

qualifier le processus correspondant de "processus tronqué". Enfin, un passage à la limite de la marche aléatoire est réalisé pour aboutir à un processus stochastique modèle dont la loi de distribution de probabilité suit l'équation FFPE. Le troisième chapitre quand à lui est consacré à des simulations de trajectoires à partir de la mise en oeuvre du modèle de mouvement brownien géométrique subordonné trouvé dans le chapitre précédent. A l'instar du modèle de marche aléatoire, des ajustements liés principalement à la surestimation des temps d'attente produite par la loi α -stable sont conduits pour approcher au mieux la mesure expérimentale.

Conformément à l'objectif principal de la thèse, la deuxième partie du travail est consacrée à la mise en oeuvre du modèle de transfert PASTIS dans le contexte de mesures couplées et synchronisées de précipitations pluvieuses et de variation du niveau de la nappe effectué sur le site atelier du centre de recherche de l'INRA Avignon. Cette partie du travail de thèse est présentée au chapitre quatre du manuscrit et a pour objectif la conduite d'une étude comparative entre la simulation par le modèle avec les données réelles en entrée et les fluctuations mesurées. Pour cela, les données de variations de hauteur de nappe sont d'abord analysées et traitées pour isoler les fluctuations cohérentes avec les événements de pluie. Une étude statistique entre les résultats de la modélisation et ces fluctuations de la nappe est alors menée pour montrer la pertinence de la modélisation proposée à représenter le transfert de masse dans un contexte réaliste.

Par ailleurs, compte tenu de la diversité des domaines mise en oeuvre dans la réalisation de cette thèse, nous n'avons pas choisi de commencer ce manuscrit par un chapitre de bibliographie comme on peut le trouver traditionnellement dans les mémoires de thèse. Au lieu de cela, des rappels bibliographiques sont donnés tout au long du manuscrit quand c'est nécessaire. Ceux-ci sont complétés par des annexes où sont rappelés plus en détails des aspects techniques ou des informations utiles mais qui ne sont pas au premier plan dans le travail de thèse.

Le dernier chapitre du manuscrit résume les grandes lignes des résultats. Nous y développons des discussions sur la portée de nos résultats et les perspectives d'applications, en particulier sur la possibilité de coupler le modèle de transfert PASTIS avec notre modèle stochastique de pluie pour réaliser des simulations adaptées avec en vue d'un objectif de prévision.

Première partie

Modélisation d'un phénomène pluvieux local.

Chapitre 1

Analyse et traitement de données de phénomène pluvieux local. Choix de l'approche de modélisation

L'étude menée ci-après aborde le traitement de données de précipitations pluvieuses relevées au quotidien en un lieu pour lesquelles on veut construire un modèle mathématique représentatif. Afin de réaliser cet objectif, des analyses et des traitements visant à établir les caractéristiques de ces données seront nécessaires pour obtenir des résultats qui seront à la base de notre approche de modélisation. Comme on pourra le voir, les premières observations des données montrent les caractères aléatoires et très erratiques des variations de la quantité de pluie. De telles caractéristiques sont généralement observées et admises depuis longtemps. Elles ont donné lieu dans la littérature à de nombreux travaux d'analyse et de modélisation adoptant l'approche stochastique [19][42][89][90][83]. Au-delà des constatations que l'on retrouve largement dans de nombreuses publications, l'étude plus approfondie de la structure probabiliste associée à ces données et les caractérisations du type de dynamique de variations observées doivent permettre d'améliorer la réflexion sur la modélisation de ces précipitations. En particulier, nous nous sommes intéressés à l'existence de propriétés de mémoire ou de caractère fractal dans les données. En effet, ces caractéristiques peuvent être les signes d'existence de distribution de probabilités associée à des lois dites à ailes épaisses [58]. On se focalisera ici sur les deux aspects principaux de la chronique de pluie à savoir la variation de l'intensité de la pluie et la distribution des occurrences. Avant d'aller plus loin dans cette étude, une première analyse des données avec des outils statistiques simples est effectuée afin de décrire rapidement les premières propriétés de la série temporelle. On s'intéressera ensuite à une analyse plus quantitative des données en étudiant les liens statistiques à différentes échelles de temps à l'intérieur.

On s'intéressera en particulier à la manifestation des effets de mémoire évalués par l'exposant de Hurst [27]. La structure fractale du phénomène décrit par la série temporelle, sera explorée par la dimension fractale et la corrélation entre accroissements passés et futurs. Nos résultats suggèrent que la série temporelle étudiée présente une mémoire à long terme. Néanmoins, cela ne dispense pas d'avoir aussi une mémoire à très court terme, décrit par exemple par un modèle de type chaîne de Markov. Dans la deuxième partie de ce chapitre, l'intensité des précipitations et le temps d'attente entre deux pluies seront considérés comme deux variables aléatoires dont on déterminera les lois de probabilité.

1.1 Analyse des données d'observation

Comme il vient d'être dit, l'analyse des données s'attachera à les représenter sous diverses formes puis à rechercher sous ces formes les caractéristiques sous-jacentes de la dynamique des variations. Si ce signal recèle des "effets mémoire", ils peuvent être inscrits dans la fonction d'auto-corrélation et d'auto-information [17] et mesurés par l'exposant de Hurst [27], lequel peut conduire à un lien avec les phénomènes non-linéaires. La dynamique de la série qui caractérise son évolution dans le temps peut se révéler au travers de "portraits de phases", sa transcription dans l'espace de Fourier et la technique du surrogate [84] permettent de souligner les relations de phase existant dans la série temporelle.

Contexte de l'étude des données

Notre approche de l'analyse de ces données de pluie consiste à le considérer comme la réalisation d'un processus stochastique [4]. Le concept de processus stochastique constitue un cadre général, assez puissant permettant d'une part de mettre en oeuvre des outils plus efficaces que les outils statistiques habituels. D'autre part, tenter de définir un processus stochastique lorsque l'on ne dispose que d'une seule réalisation est à priori une tâche très difficile voire même impossible. En effet, théoriquement nous ne pouvons pas assurer que notre problème ait une solution bien définie. Néanmoins une telle situation est assez classique dès lors qu'on souhaite décrire par un modèle probabiliste un résultat d'expérience. Dans ce cadre, il est nécessaire de prendre certaines précautions portant sur des hypothèses d'ergodicité du signal, de stationnarité (au moins au second ordre) et d'homogénéité.

Hypothèses d'érgodicité et de stationnarité

Historiquement, l'hypothèse ergodique fut formulée initialement par Boltzmann. De manière simple, on peut dire qu'elle suppose [9], que la valeur moyenne d'une grandeur calculée de manière statistique rejoint la moyenne d'un très grand nombre de mesures prises dans le temps. Dans ce sens, l'hypothèse d'ergodicité est fondamentale pour un bon rapprochement entre la théorie et l'expérience. Elle est souvent jugée vraisemblable a posteriori quand elle permet de faire des estimations correctes [4]. Une autre hypothèse toute aussi importante concerne la stationnarité qui suppose que les moments statistiques du processus concerné sont indépendants de l'origine des temps [4]. Pour mieux décrire ces hypothèses, nous donnons ci-dessous la formulation de Yaglom [94] de l'érgodicité sous l'hypothèse de stationnarité au second ordre :

Théorème 1 (Yaglom) *Si $x(t)$ est un processus aléatoire (du temps t) stationnaire du second ordre alors le moment d'ordre 1 et la fonction d'autocorrelation de $x(t)$ obtenus par moyennage stochastique et ceux obtenus par un moyennage en fonction du temps sont identiques.*

$$\lim_{T \rightarrow +\infty} \left\{ \frac{1}{T} \int_0^T x(t) dt \right\} = \langle x(t) \rangle \quad (1.1)$$

$$\lim_{T \rightarrow +\infty} \left\{ \frac{1}{T} \int_0^T x(t)x(t+\tau) dt \right\} = \langle x(t)x(t+\tau) \rangle \quad (1.2)$$

Quant à l'homogénéité des données, on l'estimera en observant que les parties du signal correspondant à chacune des quatre saisons présentent le même comportement statistique que le signal global. En effet on peut noter que des fléchissements comparables à celui de la série annuelle (figure 1.4) apparaissent à peu près au même moment pour toutes les saisons (figure 1.5). Nous adoptons "a priori" ces hypothèses dans le cadre de cette étude.

La série temporelle et sa représentation

La représentation graphique des données constituées des relevés de quantités de pluie recueillies quotidiennement pendant 40 ans (chronique de pluie entre 1968 et 2008) est donnée sur la figure (1.1). Les précipitations sont exprimées en millimètres et l'unité de temps est en jours. Les variations observées sont très erratiques et suggèrent naturellement une évolution aléatoire. Après un examen rapide, on peut remarquer une densité de points bien moindre sur le dernier quart de la représentation. Cette constatation nous a amené à nous poser des questions sur la fiabilité des données et donc sur la fiabilité à long terme du système de mesure. En particulier, nous nous sommes demandés si un incident dans le système de mesure (remplacement de capteurs, de système d'acquisition,...) aurait pu être à l'origine de ce phénomène. Notons que nos

recherches sur l'historique de l'activité de mesures n'ont pas permis d'établir des indications dans ce sens. Cet aspect de l'évolution des données sera repris et analysé ultérieurement dans ce chapitre. On peut noter que ces données renferment deux grandeurs fondamentales. D'une part les hauteurs de pluie (non nulles) qui caractérisent l'intensité des précipitations, d'autre part les jours de sécheresse ou "temps d'attente" de la pluie. Compte tenu du caractère aléatoire de la chronique des pluies, l'intensité des précipitations (figure 1.2) ainsi que l'occurrence des événements de pluie (figure 1.3) semblent aussi évoluer de manière aléatoire. Si l'on cumule les précipitations en fonction des temps d'attente entre deux événements de pluie (figure 1.4), on obtient alors un phénomène à deux dimensions, l'une spatiale, la hauteur atteinte par les valeurs cumulées et l'autre temporelle qui mesure la durée pour atteindre une hauteur donnée.

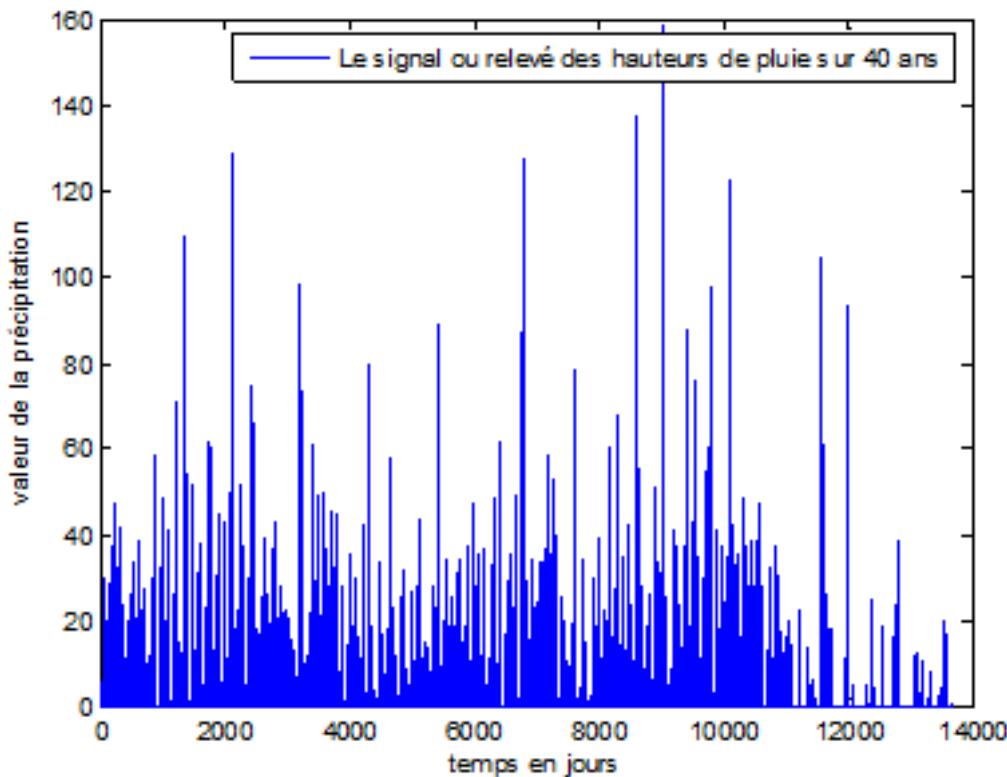


FIGURE 1.1 – La série temporelle constituée du relevé des hauteurs de pluies sur 40 ans de 1968 à 2008.

1.2 Caractérisation de la dynamique des données de pluie

Dans cette section, nous présentons une première caractérisation quantitative des données de pluie. Rappelons que le signal de base considéré ici est le relevé des hauteurs de pluie nulles ou non (figure 1.1). C'est une série temporelle à partir de laquelle on peut espérer décrire

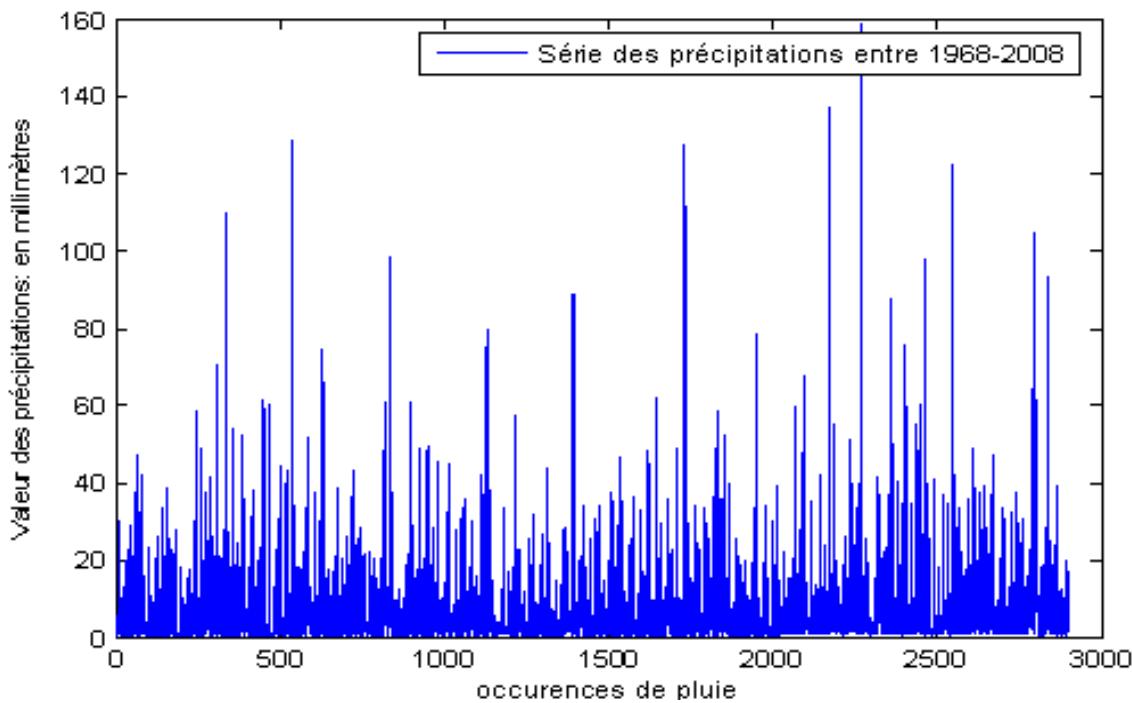


FIGURE 1.2 – Série des précipitations constituées des seules hauteurs de pluie. Sur les 40 années, Il n'apparaît pas de changement flagrant des quantités de pluies.

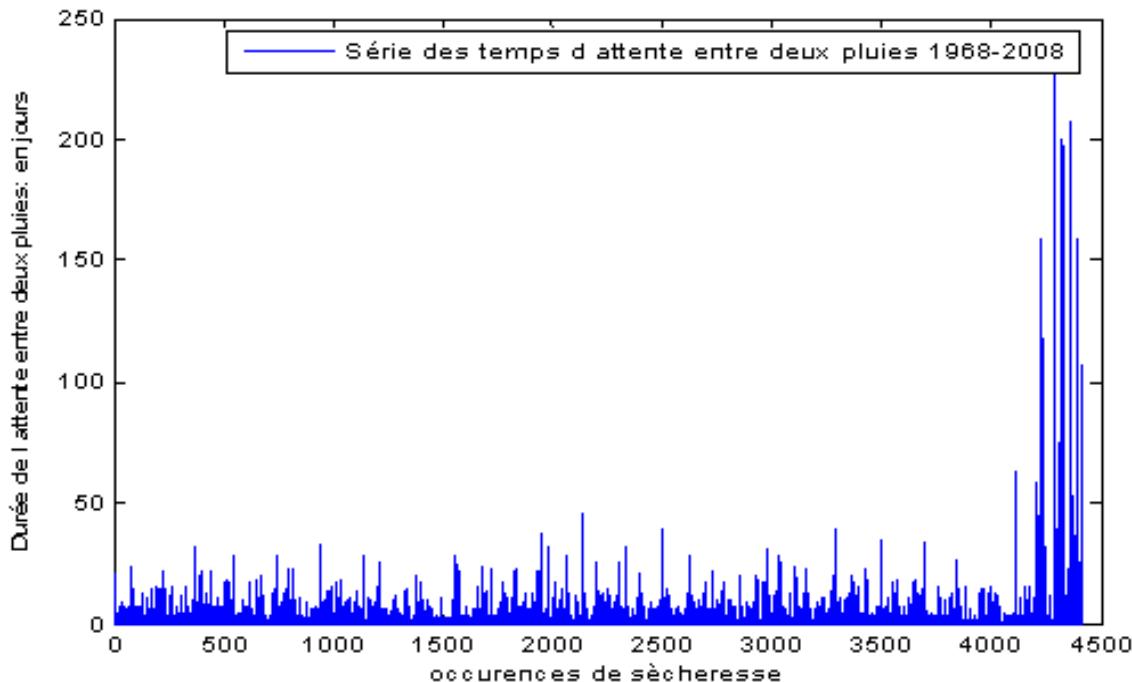


FIGURE 1.3 – Série des temps d'attente entre deux pluies. Il apparaît, à droite du graphique, une forte augmentation des temps d'attente qui traduit de longues périodes de sécheresse des années 1997 - 2008

le processus qui gère le système pluvieux local. Rappelons aussi que cette série n'est alors qu'une réalisation de ce processus inconnu. Comme il a été dit plus haut, afin d'effectuer le rapprochement entre l'observation et la théorie, il est pratique de supposer l'"ergodicité" de ce processus ce qui impose que l'on admet alors que l'évolution d'un seul signal observé au cours

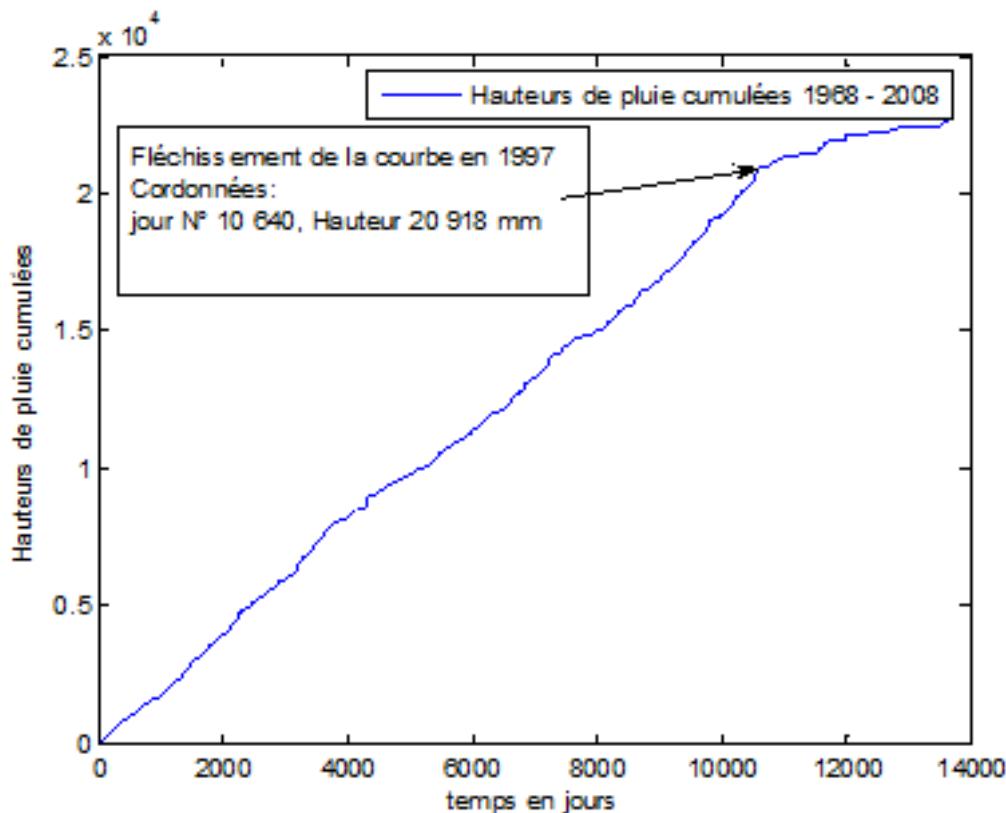


FIGURE 1.4 – Hauteurs de pluie cumulées des précipitations. On remarque évidemment un changement de dynamique vers 1997. Est-ce un indice de changement pérenne de la pluviométrie

du temps apporte la même information statistique qu'un ensemble de réalisations.

Un premier examen quantitatif de la série temporelle (figure 1.1) permet de montrer des périodes de "persistance" i-e un nombre de jours de pluie consécutifs et ceci dans 15% du temps. De même, la persistance de la sécheresse représente 80% du temps. Les 5% restant constituent les alternances pluie-sècheresse sur un jour. Cette persistance dominante laisse supposer que le signal possède un certain effet mémoire à court terme et peut-être même à long terme. Cette propriété peut constituer le caractère fondamental du processus que nous voulons décrire.

Pour corroborer cette assertion, on vérifiera successivement deux points ; le premier consistera à calculer la fonction d'auto corrélation [9][4]. Si cette dernière peut présenter une décroissance lente en loi de puissance, un tel comportement peut représenter un effet de mémoire. Le second point reposera sur l'estimation de l'exposant de Hurst[27]. En effet, si cet exposant est supérieur à 0.5 le phénomène observé présentera une certaine persistance, s'il est égal à 0.5 le phénomène n'aurait pas de mémoire, il correspondrait alors à la réalisation d'un processus stochastique semblable à un mouvement brownien. Si l'exposant de Hurst est inférieur à 0.5 le phénomène serait anti-persistant.

Par ailleurs, les effets de mémoire peuvent être liés à des caractères non-linéaires. Nous allons aussi estimer l'existence de caractère non-linéaire dans les données. L'idée que nous adopterons repose sur le fait que la non-linéarité du signal sera contenue dans la phase de sa transformée de Fourier. Pour cela nous allons faire une étude comparative du portrait de phase du signal original et de son surrogate [84]. Le signal surrogate est construit avec les mêmes amplitudes de Fourier que celles de l'originale mais pour lequel la phase a été remplacée par des valeurs aléatoires uniformément distribuées. Si les deux portraits sont fondamentalement différents, c'est que la signature des non-linéarité présente dans les phases de la transformée de Fourier du signal a été totalement détruite par la procédure de surrogate.

Si les résultats sont cohérents, ces analyses nous apporteront un faisceau de résultats sur l'influence des effets mémoires et de la non-linéarité dans les données. Rappelons ici que des résultats dans ce sens ont été publiés dans la littérature [28][87][71]. En revanche, si le système formé par les deux états : pluie et sécheresse ne présente pas d'effet mémoire, c'est-à-dire que l'état d'un jour ne dépend que de celui de la veille, vu qu'il est apériodique et irréductible, on peut supposer qu'il suit un processus de Markov [24] et, connaissant le pourcentage de temps passé dans les deux états, nous essayerons de déduire les probabilités de passage d'un état à l'autre lorsqu'il atteint un état stationnaire i-e pour un temps infini (Cf. Annexe B).

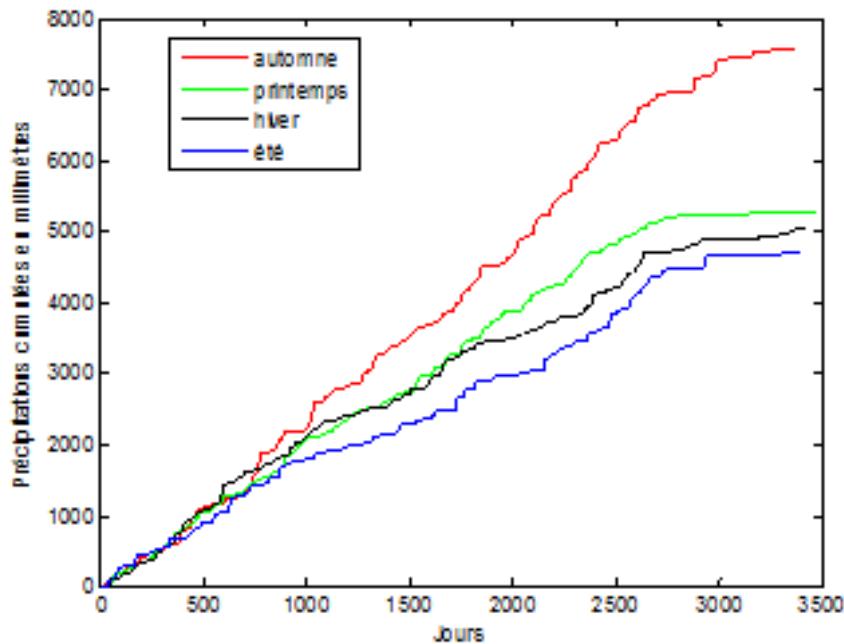


FIGURE 1.5 – Hauteurs de pluies cumulées par saisons.

Recherche d'effets mémoire dans les données

Les pluies et l'importance des précipitations, faible, forte ou très forte sont saisonnières et se retrouvent d'année en année. C'est cette dépendance entre les valeurs mesurées que l'on peut évaluer par la recherche de corrélations. A contrario, la persistance de la pluie ou de la sécheresse d'un jour sur l'autre peut laisser penser que seul l'état d'un jour conditionne l'état du lendemain, celui-ci étant alors indépendant des états antérieurs à la veille. Dans quel cas le processus pourrait être markovien. Les deux cas de dépendance seront envisagés ci-dessous.

Estimation des fonctions d'autocorrélation et d'auto-information

Rappels et définitions

Deux résultats particulièrement intéressants sont à retenir pour la présente étude : le comportement asymptotique de la fonction d'auto corrélation [5] et d'auto-information ainsi que la valeur de l'exposant de Hurst. Le distinguo entre mémoire courte et longue porte sur la décroissance de la fonction d'auto corrélation et d'auto-information au fur et à mesure que le retard augmente. La décroissance est rapide, exponentielle pour une mémoire à court terme et plus lente, hyperbolique pour la mémoire longue. Ce qui est assez intuitif. Plus rigoureusement, localement un processus "sans mémoire" montre une distribution qui se comporte comme : $p(x) = \lambda e^{-\lambda x}$ de moyenne λ et un processus à mémoire longue une distribution $p(x) = k x^{-(1+\alpha)}$ où α représente un exposant fractal. Il quantifie l'importance du regroupement local des valeurs prises par la variable aléatoire x . Si α est non nul, le processus est corrélé et peut présenter une certaine mémoire.

Rappelons la définition suivante relative à la **fonction d'autocorrélation** d'un processus aléatoire $x(t)$ à mémoire longue (pour plus de détails voir par exemple [?]).

Un processus stationnaire $X(t)$ est dit "à mémoire longue", s'il existe un réel α , $0 < \alpha < 1$ et une constante positive C telle que ρ étant la fonction d'auto corrélation (voir la définition à l'équation 1.2) et τ le retard, nous avons

$$\lim_{\tau \rightarrow \infty} \frac{\rho(\tau)}{C \cdot \tau^{-\alpha}} = 1$$

ou encore asymptotiquement : $\rho(\tau) \approx \tau^{-\alpha}$ ou encore la fonction d'auto corrélation présente une décroissance de type puissance et sa représentation en log-log est à tendance linéaire. A contrario la mémoire est courte si la décroissance est exponentielle comme écrit ci-dessus.

Dans l'espace de Fourier, l'exposant α est le coefficient d'échelle de la loi de puissance décroissante que suit la densité spectrale du processus : $S(f) \propto f^\alpha$. Cette dernière exprime combien le processus est concentré sur diverses bandes de fréquences. Notons enfin que la fonction d'autocorrélation mesure et donne une estimation de la dépendance linéaire entre deux variables aléatoires qu'on pourrait prendre comme $x(t)$ et $x(t+\tau)$ dans notre cas.

La **fonction d'autoinformation** (ou de mutuelle information) [17] mesure les dépendances statistiques linéaire mais aussi non-linéaires entre deux variables aléatoires. Sa définition est liée aux probabilités marginales de $x(t)$ et $x(t+\tau)$ dans notre cas. Pour un signal discret, la mutuelle information est donnée [17][62][85] par :

$$I(s, q) = \sum_{i,j=1}^m P_{sq}(s_i, q_j) \log \left\{ \frac{P_{sq}(s_i, q_j)}{P_s(s_i)P_q(q_j)} \right\}$$

où s_t et $q_{t+\tau}$ sont respectivement les états à l'instant t et à l'instant $(t + \tau)$. P_s et P_q sont les densités de probabilité de s et de q . P_{sq} , est la densité de probabilité jointe du couple (s,q) .

Résultats et commentaires

Les figures ci-après (figure 1.6 et figure 1.7) montrent les évolutions des fonctions d'autocorrélation et d'autoinformation de la chronique des précipitations (le signal original) respectivement en représentation linéaire et logarithmique. A première vue, on remarque que les deux fonctions évoluent dans une même gamme de valeurs situées autour de 0.2 (pour une représentation normalisée entre -1 et 1). Les premiers minimum des deux fonctions s'établissent autour de 6 jours pour la fonction d'autoinformation et de 7 jours pour la fonction d'autocorrelation. La proximité de ces deux valeurs semble suggérer que les événements de pluie espacés d'une semaine sont statistiquement peu correlés. On remarque aussi qu'à court terme (moins de 4 jours), l'évolution des deux fonctions présentent une assez grande similarité. Ces constatations semblent indiquer que l'autocorrelation et l'autoinformation mesurent les mêmes phénomènes à court terme.

Un examen plus attentif de la forme de la fonction d'autocorrelation sur la figure (1.6) montre une loi de décroissance qui oscillent autour d'une droite. On note qu'à de plus grandes échelles de temps, l'évolution de la fonction d'autoinformation décroît moins lentement et se décale de la tendance de décroissance linéaire que présente la fonction d'autocorrelation. Ce fait exprime l'idée [17][62][85] que la fonction d'autoinformation mesure au-delà de la dépendance linéaire estimée habituellement par la fonction d'autocorrelation, des liens statistiques à plus

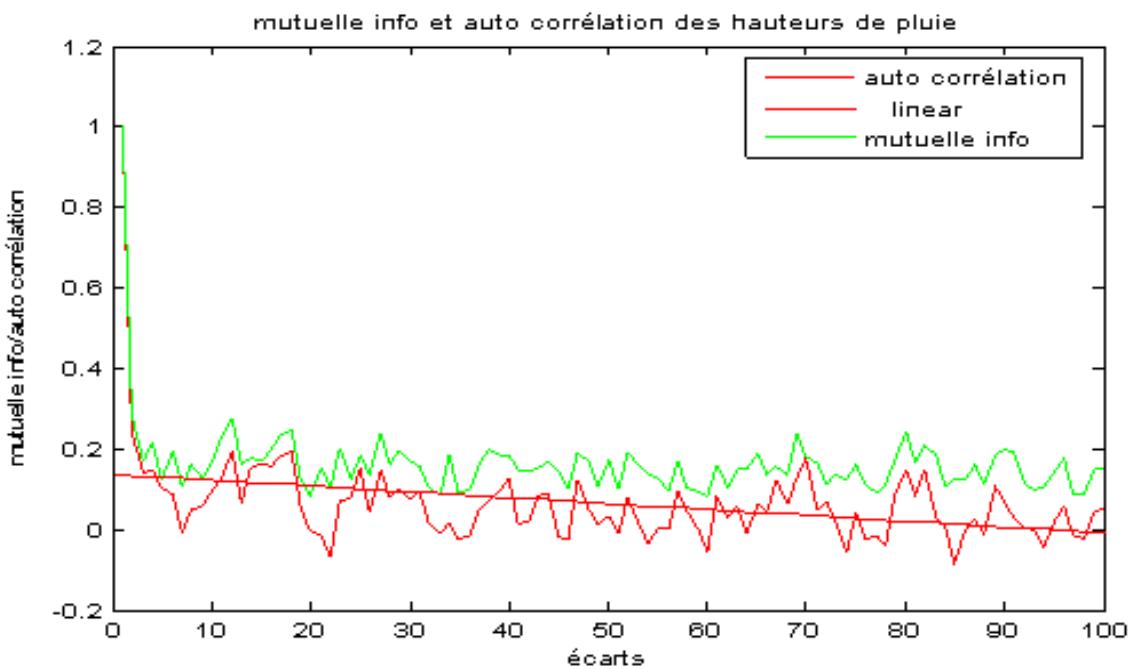


FIGURE 1.6 – Auto corrélation et auto information du signal global.

grande échelle pouvant être associés à des caractères non-linéaires de la dynamique.

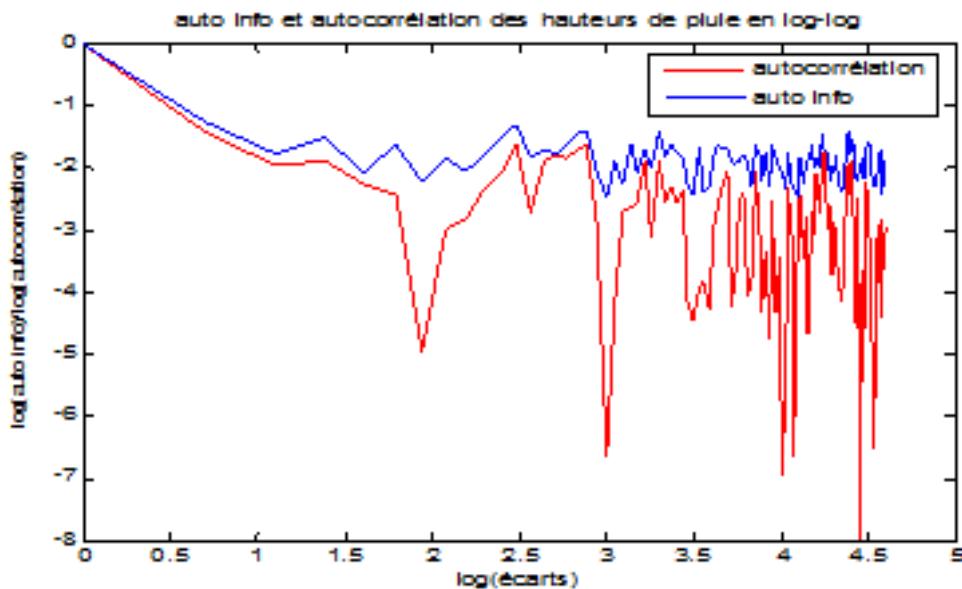


FIGURE 1.7 – Auto corrélation et autoinformation du signal global en log-log.

Par ailleurs sous une représentation logarithmique (figure 1.7), l'absence de "tendance linéaire" écarte la possibilité d'une décroissance de type puissance de la fonction d'autocorrélation. Ce qui suggère que le signal global c'est à dire la série temporelle des chroniques de pluie n'est pas associé à des distributions de probabilité de type "ailes épaisse".

Les résultats pour les précipitations cumulées (figure 1.8) semblent indiquer l'existence de

mémoire à long terme. En effet, la figure (1.9) montre une décroissance quasi linéaire (sous une représentation log-log) ce qui induit, pour les courbes elles-mêmes une décroissance de type puissance. On peut interpréter ceci par le fait que la persistance d'un temps pluvieux ou d'une période de sécheresse n'influe pas sur le comportement conjugué des deux phénomènes. Autrement dit les successions pluie-sécheresse sont relativement indépendantes ce qui décrit des alternances de périodes assez longues de pluie et de sécheresse et non pas un émettement. Ceci est un comportement pluvieux bien remarquable dans notre région.

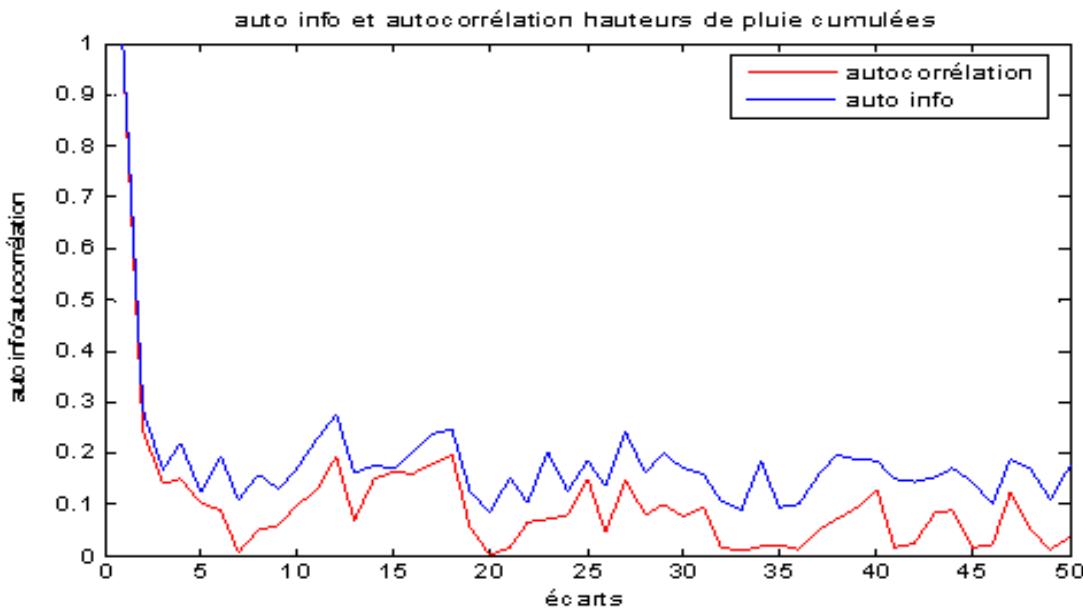


FIGURE 1.8 – Fonctions d'autocorrélation et d'autoinformation des hauteurs de pluie cumulées.

Evaluation de l'exposant de Hurst

Hurst [27] à partir de 1951, Mandelbrot à partir de 1968 [46][48], ont apporté une large contribution à l'étude des effets de mémoire dans les séries temporelles. Nature fractale d'un signal, autosimilarité, comportement asymptotique qui évolue en fonction du temps sont autant de concepts qui trouvent une mesure par l'exposant de Hurst. C'est à partir de cet exposant que l'on peut classer un signal parmi les phénomènes diffusifs, normaux ou anormaux. Techniquement, l'exposant sera calculé par l'analyse R/S et la précision du résultat évaluée sur des séries test.

Rappels et définitions

Il existe différentes façons d'aborder l'exposant de Hurst, (noté H ici). Sous un angle statistique

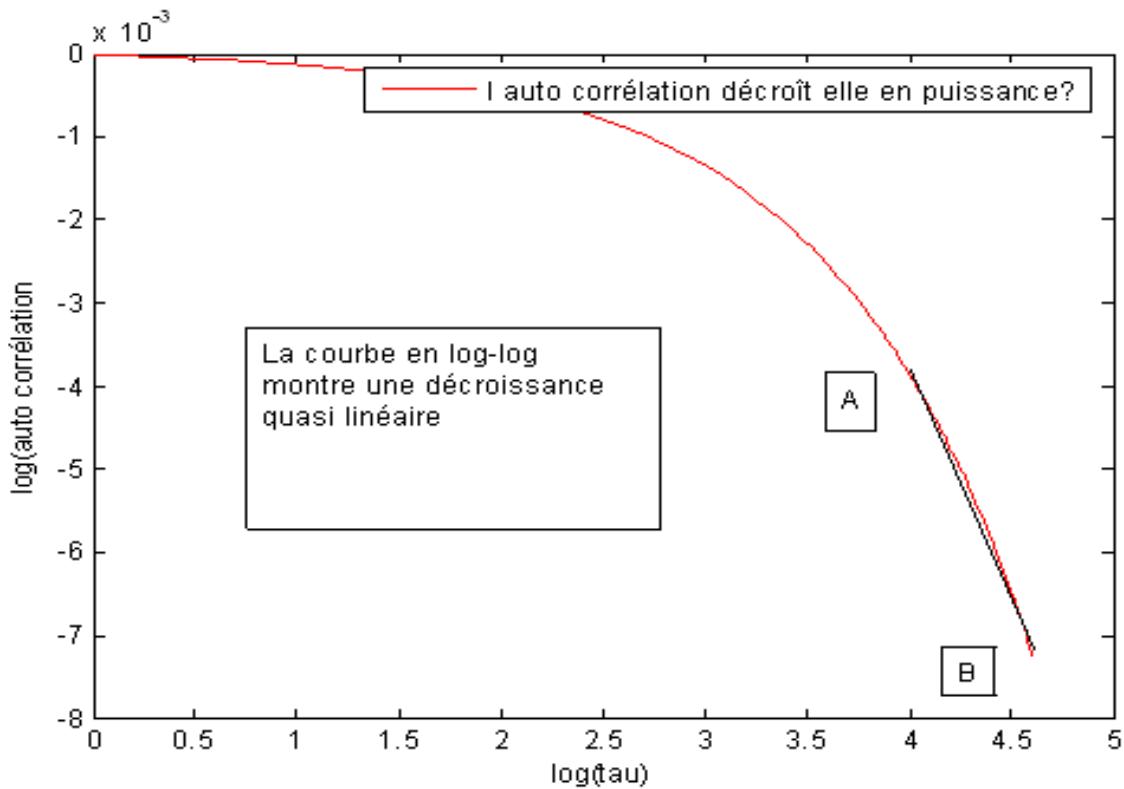


FIGURE 1.9 – Auto corrélation des hauteurs de pluie cumulées en coordonnées logarithmiques.

et probabiliste, l'exposant de Hurst est lié aux processus stochastiques qui définissent les phénomènes de diffusion. En particulier, il entre dans la caractérisation du "mouvement brownien fractionnaire" (noté ici $B_H(t)$) qui est une généralisation du mouvement brownien standard [47]. Le processus $B_H(t)$ est caractérisé par une variance (équation (1.3)) de comportement asymptotique proportionnel à la puissance $2H$ du temps.

$$E (B_H(s) B_H(t)) = 0,5 \operatorname{var}(B_H(t)) [|s|^{2H} + |t|^{2H} - |s-t|^{2H}] \quad (1.3)$$

Rappelons ici que pour un mouvement brownien standard qui n'a aucun effet mémoire, on montre que la variance possède un comportement asymptotique linéaire en temps (propriété classique des phénomènes de diffusion [30]). Il est donc facile à partir de l'équation (1.3) d'en déduire que l'exposant H vaut 0.5 pour un processus aléatoire dépourvu de mémoire. Sous un angle plus géométrique l'exposant de Hurst permet de caractériser l'apparence fractale et auto similaire d'un objet (Cf. Annexe A). Ce phénomène d'autosimilarité est statistique et signifie alors pour un signal donné qu'en moyenne, les fluctuations sur une échelle de temps sont statistiquement similaires aux fluctuations sur d'autres échelles de temps. Un signal chaotique peut ainsi se caractériser par l'exposant de Hurst qui quantifie la probabilité pour qu'un événement soit suivi par un événement similaire. Cette définition implique pour une fonction aléatoire $\Psi(t)$

possédant un exposant H, que $\Psi(ct)$ et $c^H \Psi(t)$ ont des distributions de probabilité équivalentes.

$$\Psi(ct) \triangleq c^H \Psi(t) \quad \forall t \text{ et } \forall c > 0 \quad (1.4)$$

Ce qui se traduit en dimension finie par le fait que pour toute séquence de Δ valeurs de t et toute constante c positive, les deux processus ont la même distribution. Cette relation décrit ce que l'on appelle une "loi d'échelle". La loi d'échelle s'interprète aussi comme la signature d'effet de mémoire dans un signal. En effet, lorsque l'exposant H admet un écart avec la valeur 0,5, le signal peut présenter des corrélations (linéaire ou nonlinéaire) entre événements actuels et futurs. Les liens entre événements actuels et futurs peuvent être persistants et anti-persistants suivant que H soit supérieur ou inférieur à 0,5. Un exposant supérieur à 0,5 montre donc un phénomène persistant et la persistance augmente lorsque H se rapproche de 1. En d'autres termes, l'exposant de Hurst se comporte donc comme un index quantifiant la mémoire à long terme d'un phénomène.

L'exposant de Hurst permet aussi de classer les phénomènes diffusifs anormaux. La sous-diffusion correspond à des valeurs de $H < 1/2$ donc à de l'anti-persistante. Ce qui s'interprète par un mécanisme qui devient de plus en plus difficile à mesurer que le temps s'écoule alors que pour $H > 1/2$, c'est le contraire, on est en présence d'une diffusion plus facile, c'est la super diffusion. Les images souvent citées sont celles de la fourmi dans un labyrinthe pour la sous diffusion et le tourbillon ou le vol de Lévy pour la super diffusion. Dans ce dernier cas, la trajectoire suivie par le phénomène peut évoluer longtemps dans la même direction avant d'en changer [58].

Méthode d'estimation de l'exposant de Hurst

Le calcul de l'exposant relève de l'analyse R/S (rescaled adjusted range analysis) [27] dont l'intérêt vient de ce qu'elle n'impose ni hypothèse ni exigence sur la forme de la distribution supposée. Pour une série X_t avec $t=1, \dots, n$, la quantité ci-dessous se comporte asymptotiquement comme n^H

$$\frac{R}{S_t} = \text{var}(X_t)^{-0.5} \left[\max_{1 \leq k \leq t} \sum_{i=1}^k (X_i - \bar{X}_n) - \min_{1 \leq k \leq t} \sum_{i=1}^k (X_i - \bar{X}_n) \right] \quad (1.5)$$

$$\text{Et } H \sim \frac{\log?(\frac{R}{S_t})}{\log?(n)}.$$

Pour tester la méthode et calibrer les résultats, nous l'avons appliquée à 500 simulations de mouvements browniens de types distribution normale et mouvement brownien, donc d'exposant de Hurst vaut théoriquement 0,5. La moyenne des valeurs issues de l'analyse R/S étant

relativement conforme aux résultats attendus, la méthode a été considérée comme valide. Pour une éventuelle confirmation, nous avons appliqué l'analyse statistique R/S modifiée proposée par Lo [41]. Celle-ci tient compte de la dualité mémoire courte - mémoire longue en ce sens que l'analyse R/S étant biaisée à la hausse, Lo [41] fait apparaître au dénominateur de R/S une auto covariance pondérée en fonction des décalages (q) pour apprécier le comportement à long terme. La variance est donc remplacée par la racine carrée de l'expression ci-après :

$$s_n^2(q) = \text{var}(X_t) + \frac{2}{t} \sum_{j=1}^q \left(1 - \frac{j}{q+1}\right) \left[\sum_{i=j+1}^n (X_i - \bar{X}_n)(X_{i-j} - \bar{X}_n) \right] \quad (1.6)$$

q est ici un décalage qui peut être choisi arbitrairement ou comme le préconise Lo [41] en fonction de la dimension de la série et de l'auto corrélation de 1er ordre. q prenant la forme :

$$q = \text{partie entière de } \frac{3n^{1/3}}{2} \left| \frac{2 \text{ autocorrelation}}{1 - \text{autocorrelation}^2} \right|^{2/3} \quad \text{et } q < n. \quad (1.7)$$

Le choix de q pouvant induire des différences notables entre les résultats, le calcul de l'exposant de Hurst a été réalisé en fonction de q jusqu'à ce qu'il se stabilise. 200 valeurs de q en général ont suffi excepté pour la série des temps d'attente qui a décru sans se stabiliser.

Résumé des valeurs de H

Les divers résultats sont résumés ci-dessous dans tableau (1.1). Les séries tests excèdent de 5% environ dans les deux analyses l'exposant théorique 0.5. L'analyse modifiée étant légèrement meilleure on retiendra les valeurs de cette colonne. Toutes les valeurs de H excèdent nettement 0.5 même si l'on tient compte d'une dérive de 5%, les séries montrent donc une certaine persistance, peu importante pour les intensités des précipitations et forte pour les temps d'attente ce qui pourrait bien traduire la persistance du beau temps en bordure de la Méditerranée. En tenant compte de la marge d'erreur citée plus haut, on peut conclure que toutes les séries issues du relevé des hauteurs de pluie ont un exposant de Hurst entre 0,5 et 0,7 exceptée la série des pluies cumulées pour laquelle H peut être situé entre 0,7 et 0,9. Les séries considérées sont alors réputés avoir une mémoire à long terme, ce qui semblent compléter et corroborer les précédents résultats obtenus sur l'auto corrélation et l'auto information.

Test de Non-linéarités

L'influence des caractères non-linéaires dans la dynamique du signal peut transparaître, d'une part, dans un "portrait de phases" [28][80] et d'autre part par la technique du "surrogate"

Séries Temporelles	Méthode R/S modifiée (H_m)	Méthode R/S originale (H)
Hauteurs de pluies (1968-2008)	$H_m = 0,6545$	$H = 0,6254$
Hauteurs de pluie par décades 1968-1977	$H_m = 0,5173$	$H = 0,5$
1978-1987	$H_m = 0,5643$	$H = 0,5416$
1988-1997	$H_m = 0,5863$	$H = 0,5596$
1998-2008	$H_m = 0,6113$	$H = 0,5552$
Hauteurs de pluie "Avant 1997" : 1968-1997	$H_m = 0,5756$	$H = 0,5553$
Hauteurs de pluie "Après 1997" : 1997-2008	$H_m = 0,6112$	$H = 0,5580$
Intensités des précipitations	$H_m = 0,5608$	$H = 0,5456$
Temps d'attente	$H_m = 0,6417$	$H = 0,67$
Pluies cumulées	$H_m = 0,9126$	$H = 0,7477$

TABLE 1.1 – Valeurs des différents exposants de Hurst.

[28][84]. Il est possible de statuer sur l'origine du caractère désordonné de la série temporelle en observant le portrait de phase. En effet, si l'état chaotique du signal est dominé par des effets non-linéaires, il est possible que le portrait de phase montre une topologie centrée autour d'un point attracteur [80]. Plus généralement, l'existence d'effets nonlinéaires peut être mis en avant par la comparaison du portrait de phase du signal et de son signal surrogate.

D'une manière exacte, la reconstruction de trajectoire de phase d'un signal temporel nécessite la connaissance des dérivées successives du signal temporel $x(t)$. Pour une série temporelle expérimentale, la méthode utilisée repose sur le principe des retards : il s'agit plus précisément d'une reconstruction d'une trajectoire dans un espace à m dimension (\mathbf{R}^m) en utilisant pour chaque instant t la valeur $x(t)$ et les valeurs aux $(n-1)$ -instants décalés d'un retard τ . Chaque élément de la trajectoire dans (\mathbf{R}^m) sera alors de la forme :

$$X(t) = (x(t), x(t - \tau), x(t - 2\tau), \dots, x(t - (m - 1)\tau)) \quad (1.8)$$

L'importance de cette représentation est par sa nature même directement fonction de la valeur qu'on attribue au temps de retard τ et à la dimension m de l'espace dans quel le signal temporel est plongé. Deux considérations importantes sont prises en compte dans le choix de ces deux paramètres de construction à savoir la taille des données obtenues et la bonne représentativité du phénomène. Différentes méthodes ont été élaborés pour optimiser ce choix, notamment la

méthode basée sur la recherche d'un temps minimum permettant de s'assurer de l'indépendance des informations contenues dans chaque point de la trajectoire [17].

Un signal surrogate peut être construit à partir de la transformée de Fourier de la série temporelle $x(t)$ à laquelle la phase est rendu aléatoire. La transformée inverse de Fourier du signal $x(t)$ avec la phase rendue aléatoire donne un nouveau signal $x'(t)$ dont les propriétés habituelles telles que le spectre de Fourier ou la loi de probabilité restent inchangées.

L'étude comparative des données $x(t)$ et $x'(t)$ constitue un moyen efficace pour déterminer la nature du chaos inclus dans $x(t)$. En effet la libération des phases implique la disparition des structures cohérentes typiques des effets non-linéaires dans l'espace de phase. Par conséquent si la dynamique de série temporelle $x(t)$ est influencée par des effets non-linéaires, ses propriétés topologiques dans l'espace de phase et celles de $x'(t)$ seront très très différentes.

Résultats et commentaires

La figure (1.10) représente un portrait de phase à trois dimensions en utilisant la représentation vectorielle donnée à l'équation (1.8) avec comme paramètres $m = 3$ et $\tau = 7$ correspondant au premier minimum de la fonction d'autoinformation. Sur cette figure l'ensemble des points du portrait sont dessinés en bleu et en rouge les 1ers points tracés pour montrer comment ils se répartissent dans le temps. On remarque que l'évolution se distribue suivant les trois axes de coordonnées qui semblent jouer un rôle d'attracteur. Ce phénomène s'explique globalement par la répartition assez dense à l'intérieur de la série temporelle des valeurs nulles qui représentent des jours sans pluie.

Portrait de phase "quotidien" dans l'espace: $A(t), A(t+7), A(t+14)$

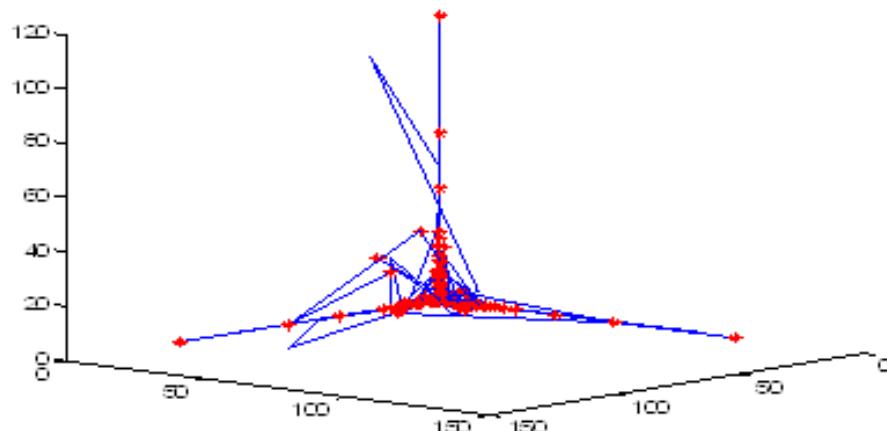


FIGURE 1.10 – Portrait de phase des données de précipitations.

La figure (1.11) représente le portrait de phase du signal surrogate de la série temporelle des précipitations. On note que le portrait de phase du surrogate présente un comportement "sphérique" qui contraste considérablement avec celui du signal original. Puisque la série des précipitations et son surrogate manifestent des différences très nettes au niveau de leurs portraits de phase, ceci semble suggérer que des relations de phase qui existaient dans la série originale ont été détruites dans le signal surrogate. En d'autres termes, ces observations tendent à confirmer l'existence de caractères nonlinéaires qui pourraient corroborer les résultats précédents obtenus avec l'analyse des variations des fonctions d'autocorrelation et d'autoinformation, et l'estimation de l'exposant de Hurst. Notons aussi que nous avons observés des résultats comparables avec ceux des figures (1.10) et (1.11) sur les séries des intensités des pluies et des temps d'attente.

Portrait de phase du surrogate dans l'espace: $A(t)$, $A(t+6)$, $A(t+12)$

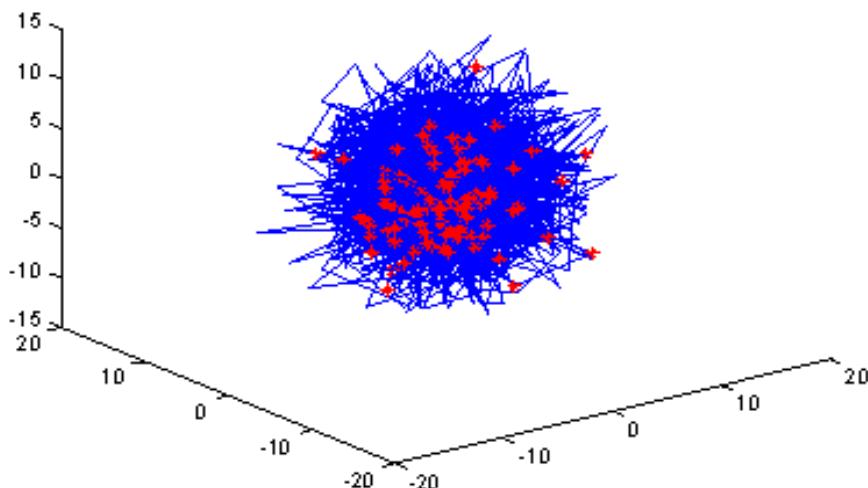


FIGURE 1.11 – Portrait de phase du surrogate des données de précipitations.

1.3 Distributions de probabilités

L'hypothèse qui est faite dans cette étude est que le phénomène pluvieux local suit un processus stochastique qui reste à déterminer et pour cela, l'intensité des précipitations comme les temps d'attente entre deux pluies sont des variables aléatoires dont on s'attachera à déterminer les caractéristiques. Les variables aléatoires étant définies, il importe de les caractériser afin de les introduire dans le processus. Les résultats de nos analyses montrent que les intensités des précipitations suivent une loi de distribution de probabilité loi log-normale alors que la série

des temps d'attente entre deux pluies, semble avoir une distribution de type α -stable.

Rappels et définitions

La notion de variable aléatoire se situe à la base de la caractérisation des phénomènes aléatoires. De façon simple, on peut définir une variable aléatoire η comme une grandeur dont la valeur est associée au résultat d'un événement (ϕ) qui dépend du hasard. De ce fait, la valeur en question n'est pas unique mais appartient à un ensemble de valeurs possibles. Sur le plan mathématique, on ne peut définir que la chance (la probabilité) pour que la variable aléatoire prenne, à la suite de l'événement, l'une de ces valeurs possibles.

Notons Γ l'ensemble de tous les résultats possibles ϕ de l'épreuve considérée. Soit A une σ -algèbre de sous-ensemble de Γ . Soit P une mesure définie sur A à valeur dans $[0, 1]$. Le triplet (Γ, A, P) constitue un espace de probabilisé sur lequel on peut définir η comme une application vers l'ensemble des réels \mathbf{R} :

$$(\Gamma, A, P) \rightarrow \mathbf{R} \quad (1.9)$$

$$\phi \rightarrow \eta(\phi)$$

Le caractère aléatoire de η est complètement défini par sa fonction de répartition donnée[91]par :

$$F(x) = Pr(\eta(\phi) < x) \quad (1.10)$$

Dans le cas où F admet une dérivée alors on définit la densité de probabilité de η par

$$f(x) = \frac{dF(x)}{dx} \quad (1.11)$$

Dans ce cas la probabilité pour que $\eta(\phi) = x$ vaut $f(x)dx$. Nous complétons cette propriété par l'énoncé du théorème suivant[91] :

Théorème 2 (Ventsel) *Pour qu'une fonction réelle F à valeur dans $[0, 1]$ soit une fonction de répartition liée à une loi de probabilité P , il faut et il suffit que $F(x_0)$ soit croissante, continue à gauche et tende vers +1 (resp. 0) si x_0 tend vers $+\infty$ (resp. $-\infty$).*

Pour la suite, nous supposons que la densité de probabilité $f(x)$ existe.

Dans cette section, notre recherche des lois de probabilités s'oriente plus particulièrement vers les lois à "ailes épaisse". En effet, différents travaux théoriques [72, 73] et expérimentales

[70, 77] dans la littérature montrent qu'il peut exister des liens étroits entre l'existence de longue mémoire et un comportement statistique à ailes épaisses.

Une loi de probabilité est dite à 'ailes épaisses' si elle possède des ailes qui ne peuvent pas être bornées par des exponentielles [2][79]. Une variable aléatoire η admet une telle loi si sa probabilité évolue asymptotiquement comme une puissance

$$P[\eta > x] \sim x^{-\alpha}, \text{ quand } x \rightarrow \infty, 0 < \alpha < 2 \quad (1.12)$$

où le paramètre α définit l'exposant caractéristique de la loi de probabilité à ailes épaisses. De telles lois ont été observées dans différents phénomènes naturels [29, 60, 64, 70]. Parmi les lois de probabilités à ailes épaisses, la famille des lois alpha stables de Lévy constitue une classe à part à cause de sa propriété de stabilité par l'addition. Cette stabilité par l'addition est à la base de la possibilité de généraliser le théorème central limite (voir par exemple [15][72]) et qui confère à la famille des lois stables une si grande importance.

De manière simple on peut dire qu'une loi de probabilité est dite stable si pour une suite X_1, X_2, \dots, X_n de variables aléatoires distribuées selon cette loi, la variable aléatoire $\frac{X_1+X_2+\dots+X_n}{c_n}$ possède la même loi que les X_n et c_n est une puissance de n [15][20]. Une autre propriété fondamentale des lois stables est aussi d'être les seules lois en probabilité à posséder le caractère d'attracteur pour presque toutes les autres lois de probabilités. Ces propriétés sont intimement liées comme le prouvait le théorème établi par Kolmogorov qui dit que seules les lois stables possèdent un domaine d'attraction non-vide [20]. Une des conséquences de ces propriétés est qu'un grand nombre de phénomènes aléatoires obtenus par addition de variables aléatoires indépendantes possédant des propriétés communes mais pas forcément distribuées selon une loi stable tendraient vers des lois stables [?]. Rappelons que pour les lois stables, l'exposant caractéristique (qui est aussi l'ordre de stabilité) $0 < \alpha \leq 2$. Pour $\alpha = 2$, on retrouve la loi gaussienne. Pour ces lois stables, seuls les moments statistiques d'ordres inférieurs ou égaux à l'exposant caractéristique sont finis. En particulier, on note que seule la loi gaussienne possède une variance finie et que de plus seules les lois ayant un exposant caractéristique $\alpha > 1$ admettent une espérance finie. Nous avons donc porté tout naturellement un grand intérêt à cette catégorie de lois de probabilité, partant de l'idée que les observations pourraient être la somme d'un grand nombre de fluctuations statistiquement comparables. Rappelons néanmoins que d'autres lois à ailes épaisses existent comme par exemple les lois log-normales que nous allons utiliser pour caractériser la distribution de probabilité des intensités des pluies.

Loi de probabilité des intensités de pluie

L'histogramme des intensités des précipitations est représenté sur la figure (1.12) en représentation linéaire et sur la figure (1.13) en représentation log-log. Une modélisation de l'histogramme par une loi log-normale est aussi représentée. Les paramètres de la loi log-normale sont, moyenne = 1.3944, écart-type = 1.149. On remarque des écarts entre l'histogramme et la loi lognormale pour les petites valeurs de la variable, ceux-ci proviennent, en partie, du défaut de mesures des quantités de pluies reçues inférieures au calibre des pluviomètres à augets. On note un assez bon accord entre les deux courbes au niveau des ailes de la loi de probabilité. Il est admis, dans la littérature, que les intensités des précipitations pluvieuses jusqu'à un seuil de 50 mm/h, dans nos régions, suivent une distribution log-normale et gamma au-delà [10].

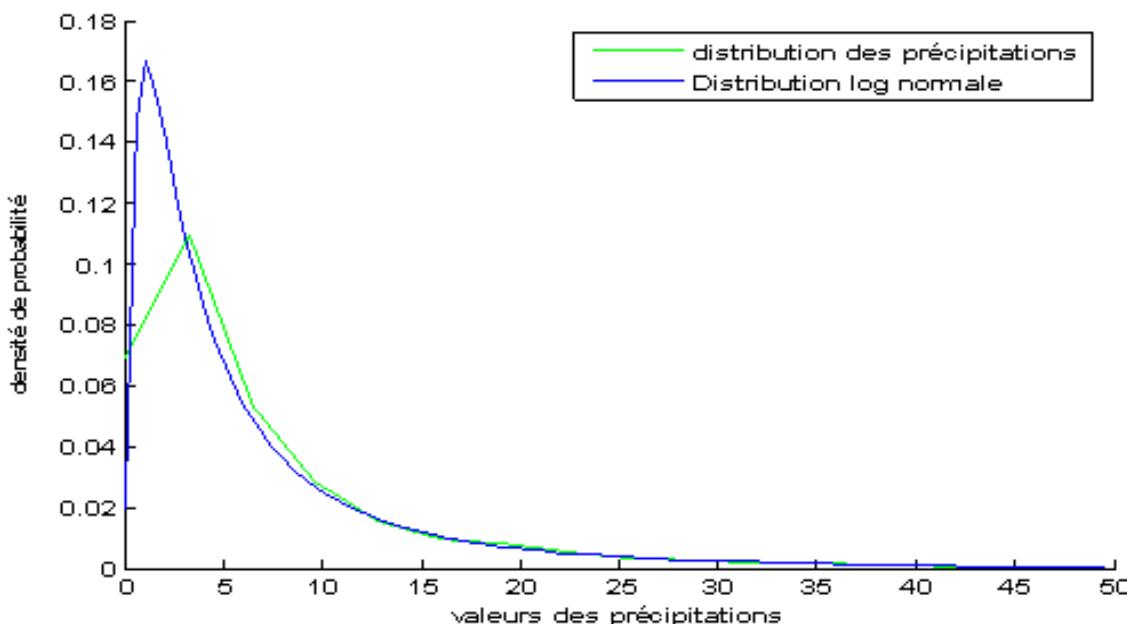


FIGURE 1.12 – Distribution des intensités de pluie et la distribution log-normale qui lui est ajustée (fittée).

Loi de probabilité des temps d'attente

Nous avons vu dans la première section de ce chapitre que les temps d'attente par définition ne prennent que des valeurs positives. D'autre part, des évènements extrêmes tels que les sécheresses des années 2005-2007 apparaissent sur le graphique des temps d'attente (figure 1.4). De plus, compte tenu des résultats obtenus dans les sections précédentes (existence de longue mémoire, d'effet nonlinéaire), nous allons chercher un modèle de loi de distribution de probabilité à ailes épaisses pour la série des temps d'attente. Une première étape pour vérifier

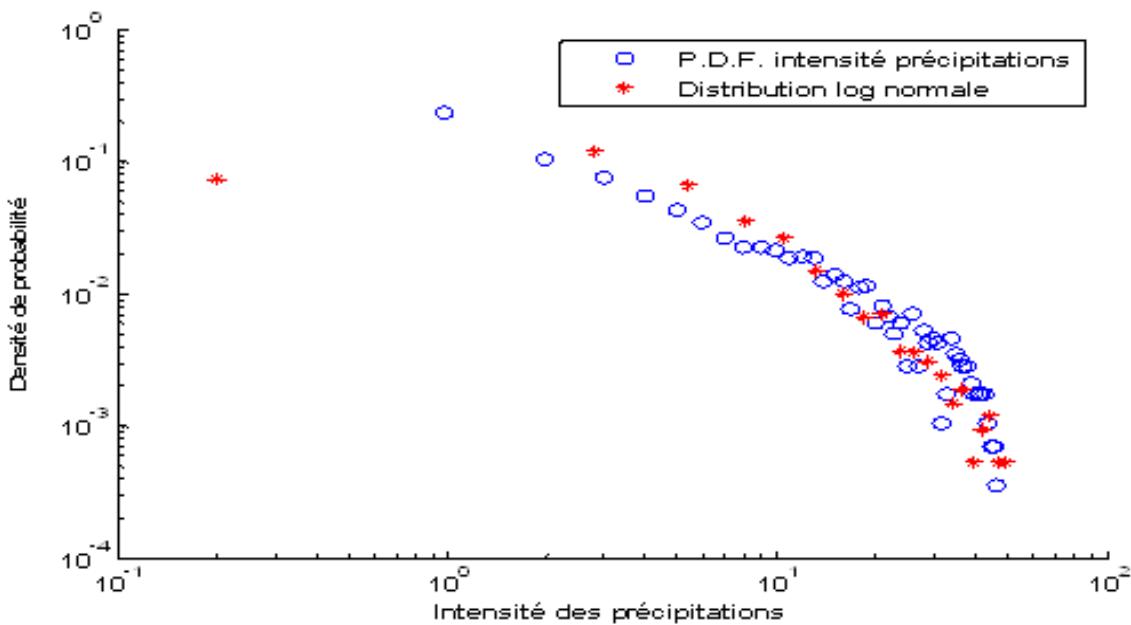


FIGURE 1.13 – Comparaison des distributions des intensités de pluie et de la log-normale en coordonnées logarithmiques.

le bien fondé du choix d'un modèle à aile épaisse consiste à évaluer si les moments statistiques de la série des temps d'attente divergent ou non. Bien évidemment, comme la série est d'une longueur finie, un calcul direct ne pourrait pas donner lieu à une réponse satisfaisante.

Les temps d'attente ont-ils une variance finie et une espérance finie ?

Pour vérifier cette éventualité sur la variance, nous allons appliquer deux tests graphiques. Le premier pour apprécier la variance de la série qui devrait diverger dans le cas d'une distribution à aile épaisse et le second pour estimer l'exposant de stabilité α .

a) Test graphique 1

Le test consiste à calculer, pour diverses valeurs de n : $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ et $\text{var} = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$ puis tracer la représentation de (n, var) . Intuitivement, lorsque n augmente et lorsque la variance est finie, le tracé doit converger. Au contraire, si on est en présence d'une loi à variance infinie, le tracé diverge.

b) Test graphique 2

Ce deuxième test est basé sur le fait que le comportement asymptotique de la fonction de survie ($1-F(X)$) de la variable aléatoire X est de type puissance de $(-\alpha)$ soit :

$$\lim_{t \rightarrow \infty} t^\alpha \Pr(|X| > t) = \gamma C(\alpha)$$

Donc à l'infini $\frac{d\log(P(|X|>t))}{d\log(t)}$ est équivalent à α .

D'où : Fixer t et calculer $f(t) = \log\left(\frac{1}{n} \sum_{i=1}^n 1_{|x_i|>t}\right)$. Tracer le graphique de $(\log(t), f(t))$ et voir si la pente devient finie pour de grandes valeurs de t . Cette pente est alors en valeur absolue une évaluation de l'exposant α .

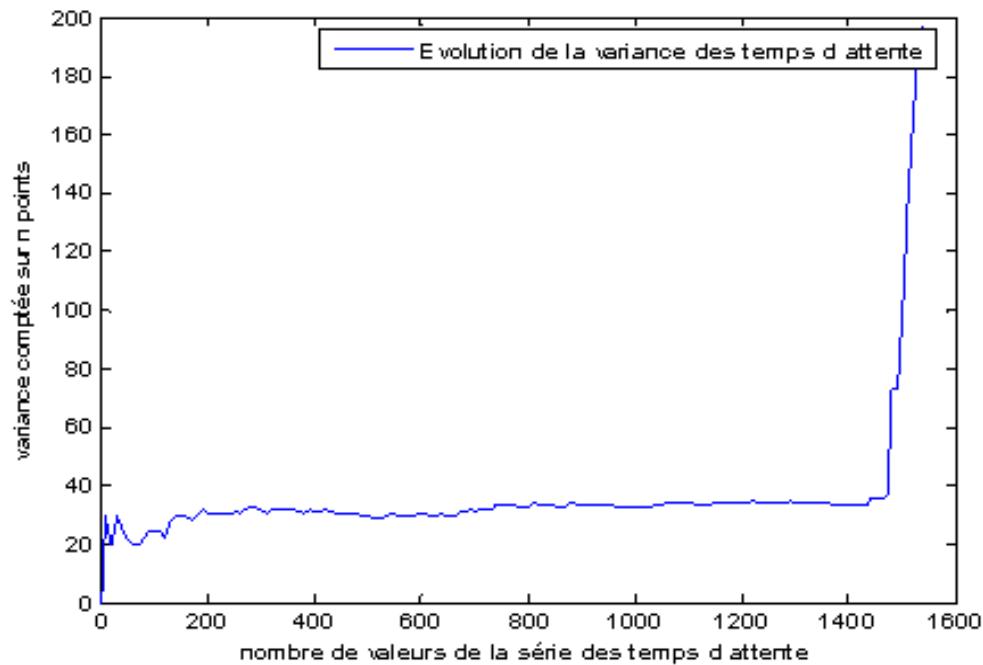


FIGURE 1.14 – Premier test graphique : évolution de la variance.

La figure (1.14) donne le résultat du premier test. La variance semble bien diverger avec le nombre d'observations. On note que l'augmentation brutale de la variance est liée à l'accentuation de l'occurrence de longues périodes de sécheresse (figure 1.3) qui est aussi à l'origine du fléchissement de la dynamique de la série des précipitations cumulées (figure 1.4). Le résultat du deuxième test est présenté sur la figure (1.15). On remarque un comportement asymptotique globalement linéaire. Le graphique dans son comportement asymptotique est ajusté par un segment de droite entre A et B dont la pente est de 0.96 ce qui donne une estimation de l'exposant $\alpha = 0.96$ avec une erreur de l'ordre de 5%. Ces deux résultats semblent suggérer d'une part que la distribution de probabilité de la série des temps d'attente peut être représentée par une loi à ailes épaisses dont la variance n'est pas finie. D'autre part, l'exposant caractéristique estimé par le deuxième test de cette loi indique une valeur plutôt inférieure à 1. Par ailleurs, comme nous l'avons déjà dit, les temps d'attente doivent être tous positifs. L'ensemble de ces caractéristiques laisse supposer qu'une variable aléatoire de loi α -stable constitue un candidat potentiel pour modéliser la distribution aléatoire des temps d'attente.

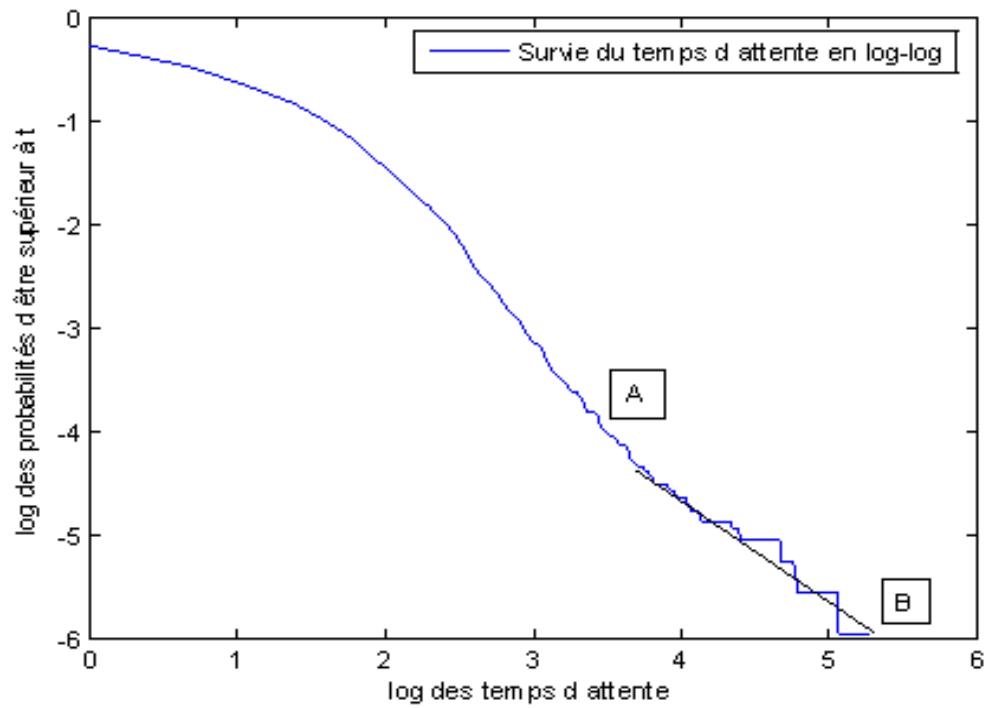


FIGURE 1.15 – Estimation graphique de l'exposant de stabilité α .

Paramètres de la distribution de la loi de probabilité des temps d'attente

L'estimation graphique ci-dessus nous a donné l'ordre de grandeur de l'exposant caractéristique mais elle n'est pas suffisante pour obtenir un résultat fiable et précis. Par ailleurs, en dehors de l'exposant caractéristique (appelé aussi indice de stabilité) α , la caractérisation d'une loi stable implique la recherche de trois autres paramètres à savoir un paramètre d'asymétrie β , un paramètre d'échelle σ et un paramètre de localisation μ [15][82] (voir en annexe un détail sur une définition plus précise de la loi de probabilité α -stable). Il existe de multiples méthodes pour déterminer ces quatre paramètres. En ce qui nous concerne, on s'attachera avant tout à déterminer l'exposant caractéristique α puisque la série des temps d'attente étant toujours positive, ceci impose que le paramètre β doit être égal à 1. Cette asymétrie maximale est la condition nécessaire pour qu'une loi α -stable (avec $\alpha < 1$) fournisse des échantillons uniquement positifs. Les paramètres σ et μ seront utilisés comme des variables d'ajustement. Le tableau tableau (A.1) regroupe les résultats obtenus avec la mise en oeuvre de cinq méthodes d'estimation de l'exposant α .

Les détails sur la mise en oeuvre des différentes méthodes énumérées ci-dessous sont donnés dans l'annexe C. Afin de tester la fiabilité de l'implémentation des méthodes, des tests avec des échantillons pseudo-aléatoires construits numériquement suivant l'algorithme Chambers-

Méthodes utilisées	Echantillons test $\alpha=0.4$	Echantillons test $\alpha=0.9$	Série des temps d'attente
Méthode de McCULLOCH (quintiles)	ne fonctionne que pour $\alpha > 0.6$	$\alpha = 0.94$	$\alpha = 0.92$ $0.87 < \alpha < 0.97$
Méthode des moments (fonction caractéristique)	$\alpha = 0.42$ $\beta=1.36$ $\sigma=1.59$ μ non significatif	0.89 $\beta=1$ $\sigma=1$ μ =non significatif	Valeurs très différentes en fonction du choix des valeurs de t
Méthode de Koutrouvelis	$\alpha = 0.35$ $0.31 < \alpha < 0.39$	$\alpha = 0.89$ $0.87 < \alpha < 0.91$	$\alpha = 0.97$ $0.95 < \alpha < 0.99$
Méthode de Hill	$\alpha = 0.44$	$1.4 < \alpha < 1.5$	$1.7 < \alpha < 1.8$
Méthode de régression des queues	$\alpha = 0.4$	$\alpha = 0.9$ $0.85 < \alpha < 0.94$	$\alpha = 0.94$ $0.89 < \alpha < 0.98$

TABLE 1.2 – Estimation des paramètres de la loi stable des temps d'attente.

Mallows-Stuck [8] ont été faits. Des échantillons avec $\alpha = 0.4$ et $\alpha = 0.9$ ont été simulés et testés. Les résultats de l'estimation de α sur les échantillons tests sont donnés sur les deux premières colonnes du tableau. Pour les échantillons $\alpha = 0.4$, en dehors de la méthode de McCulloch [54, 55] qui n'est pas adaptée à des exposants faibles (< 0.6), les méthodes ont donné une bonne estimation de la valeur théorique de 0.4. La méthode de Koutrouvelis [32] a abouti à une mauvaise estimation avec une erreur relative inférieure à 10%. A l'inverse, pour une valeur de α proche de l'unité (valeur limite pour la positivité des échantillons), la méthode de Hill [25] aboutit vraisemblablement à des estimations érronées tandis que les autres méthodes donnent des estimations avec des marges d'erreurs entourant la valeur exacte qui est de 0.9. En ce qui concerne la série des temps d'attente, les méthodes de McCulloch, de Koutrouvelis et la méthode de régression des queues [1] donnent respectivement les valeurs 0.92, 0.97 et 0.94. La proximité de ces valeurs (en tenant compte des marges d'erreurs) suggère que l'exposant caractéristique α de la série des temps d'attente se situe entre 0.92 et 0.97.

1.4 Conclusion du 1^{er} chapitre

Dans ce chapitre, nous avons étudié sous différentes approches les caractéristiques des données de relevés quotidiens de précipitation recueillies sur quarante ans. L'étude de la série temporelle a permis de montrer qu'au delà de l'aspect désordonné et aléatoire des valeurs observées, des caractéristiques de mémoires à long termes liées à des manifestations de non-linéarité apparaissent dans les données. Les fonctions d'auto corrélation et d'auto information ont permis de révéler des effets de persistance à court terme et de mémoire à long terme du signal. Les portraits de phase des données comparés à ceux de son surrogate ont montré que des relations de phases du signal étaient détruites, ce qui sous entend qu'une certaine non linéarité apparaît dans le signal. Ces caractéristiques sont connus pour être étroitement liées à des propriétés statistiques de lois de probabilité à ailes épaisses. En accord avec cette idée qui est largement reprise dans la littérature, nous avons ensuite cherché à modéliser les distributions de probabilité des intensités des precipitations et des temps d'attente avec des lois à ailes épaisses. Une telle distribution devait d'une part montrer une variance qui diverge avec le temps, ce que nous avons effectivement vérifié sur un test graphique et d'autre part estimé, graphiquement aussi l'index de stabilité. Ces résultats constituent des informations importantes qui seront utilisées pour décrire et construire une modélisation mathématique de la dynamique de cette série temporelle des précipitations. La démarche adoptée dans le chapitre suivant consiste à considérer la dynamique de la precipitation sous la forme d'une marche aléatoire dont les sauts représentent la variable aléatoire intensité de la pluie et les temps d'attente entre deux sauts modélisent la durée aléatoire entre deux pluies successives. Cette démarche et l'exploitation du modèle qui en sortira feront l'objet du prochain chapitre de ce manuscrit.

Chapitre 2

Choix et mise en oeuvre de la modélisation de la dynamique du phénomène pluvieux

Le phénomène pluvieux local, concrétisé par la série temporelle des hauteurs de pluie dont les caractéristiques statistiques ont été établies au 1^{er} chapitre, peut être modélisé par un processus stochastique. Nous abordons cette approche ici en deux parties. Dans la première partie, la série temporelle des hauteurs de pluie cumulées sera modélisée par une marche aléatoire dont les sauts et les temps d'attente entre deux sauts représenteront respectivement l'intensité des précipitations et les durées entre deux événements de pluies. Les trajectoires simulées avec ce modèle de marche aléatoire seront étudiées comparativement avec les hauteurs de pluies cumulées. Des conditions physiques comme le respect de la proportion de sécheresse et de pluie seront prise en compte dans la modélisation. Ainsi la démarche développée consiste à créer une marche aléatoire dite "tronquée" en ce sens que les temps d'attentes sont distribués suivant une loi α -stable à l'origine mais restreinte et calibrée par la suite pour n'en garder que les simulations offrant un certain voisinage avec la réalisation expérimentale dont on dispose. Ces restrictions consistent à "tempérer" la distribution en la bornant par une distribution exponentielle pour éviter les temps d'attente exagérément longs, calibrage, à "trier" les durées pour qu'elles soient conformes à la durée expérimentale et enfin à "sélectionner" les simulations de manière à présenter les mêmes proportions de périodes humides et de sécheresse que la série expérimentale. La 2^{nde} partie s'attache à rechercher la limite spatio-temporelle de la marche aléatoire ci-dessus. Cette limite sera une F.F.P.E. (Equation Fractionnaire de Fokker-Planck), exploitée avec des temps d'attente "subordonnés" et toujours avec les mêmes restrictions et calibrages. Les simulations sont réalisées à partir de l'E.D.P.S. (Equation aux Dérivées Partielles

Stochastiques) d'Îto relativement simples d'utilisation [44]. Pour un temps donné, la P.D.F. sera établie par la méthode de Monte-Carlo. L'estimation de la proximité des simulations à la courbe expérimentale sera évaluée, d'une part par les écarts de celle-ci à la trajectoire moyenne et d'autre part par la situation de la courbe expérimentale parmi les déciles des trajectoires et enfin par diverses mesures de leur dispersion.

2.1 Première approche du processus pluvieux local

Conception d'une marche aléatoire simulant le processus pluvieux

La marche aléatoire en temps continu (CTRW) a été introduite par Montroll et Weiss [61]. Dans ce modèle, le marcheur, ici dont la position représente la valeur de la hauteur de pluie, effectue une marche constituée d'un saut distribué suivant une loi log-normale puis s'immobilise durant un certain temps, dit temps d'attente, distribué selon une loi α -stable de type $\tau^{-1-\alpha}$ dans son comportement asymptotique avec $0 < \alpha < 1$ avant d'effectuer le saut suivant. Ces marches aléatoires modélisent souvent des phénomènes soumis à une diffusion anormale car pour $\alpha < 1$ le temps moyen d'attente est infini ce qui traduit un comportement sous diffusif. Comme on l'a déjà dit auparavant, le processus pluvieux qui nous intéresse va être modélisé à partir d'une seule réalisation que constitue le relevé des hauteurs de pluie. Celle-ci a été partagée en deux composantes à savoir les séries des intensités des précipitations et des temps d'attente dont on a analysées les distributions de probabilité dans le chapitre précédent. La marche aléatoire proposée modélisera ainsi la série temporelle des précipitations cumulées. Rappelons que selon les résultats que nous avons établis précédemment, la série temporelle des précipitations cumulées suit un régime sous diffusif exprimé par son exposant de Hurst [43].

D'autre part, si on jette un regard sommaire sur le phénomène physique que constitue la pluie dans une région, on peut admettre qu'il est le résultat du scénario suivant : Les gouttes d'eau sont extraites par évaporation au dessus de zones humides puis sont transportées jusqu'au pluviomètre où elles sont réceptionnées et comptées pour donner une valeur qui est la hauteur de pluie un jour donné. Ce scénario décrit un processus aléatoire unidimensionnel, Un tel processus peut être approché par une marche aléatoire en espace et en temps, les précipitations étant assimilées à des sauts distribués suivant la loi log-normale estimée précédemment et le temps distribué suivant une loi 0,95-stable estimée, elle aussi, précédemment. Chaque saut est

alors la contribution à la marche aléatoire cumulée au temps t.

Construction du modèle de marches aléatoires des hauteurs de pluie cumulées

Cette marche aléatoire comporte donc deux dimensions, l'une spatiale et l'autre temporelle. Pour la première, la hauteur de pluie cumulée x_n du nième jour, est égale à la hauteur du jour précédent augmentée de la précipitation ou encore du saut j_n du jour.

$$x_{n+1} = x_n + j_n \quad (2.1)$$

où j_n est une variable aléatoire de loi log-normale dont la loi de densité de probabilité s'écrit

$$\varphi(x) = \frac{1}{\sigma \sqrt{2\pi}} \frac{1}{x} \exp\left(-\frac{(\ln(x)-\mu)^2}{2\sigma^2}\right) \quad \text{avec} \quad x > 0 \quad (2.2)$$

Comme il a été précisé précédemment, les caractéristiques μ, σ ; moyenne et écart-type de j_n , sont calculées à partir de la moyenne et variance (m, var) de la série expérimentale selon la formule

$$\mu = \ln(m) - 0.5 * \ln\left(1 + \frac{\text{var}}{m^2}\right) \quad \text{et} \quad \sigma^2 = \ln\left(1 + \frac{\text{var}}{m^2}\right) \quad (2.3)$$

L'équation (2.3) donne $\mu = 1,394$ et $\sigma = 1,149$ pour les valeurs $m = 7,8$ et $\text{var} = (12,9)^2$ de la série temporelle. Il est à noter ici que ces valeurs de μ et σ ne constituent qu'une estimation "à priori". Elles seront réévaluées par ajustement des simulations (trajectoires) aux données expérimentales. On peut aussi utiliser une forme centrée de la variable de saut j_n pour re-écrire l'équation (2.1) sous une forme plus standard de marche aléatoire

$$x_{n+1} = x_n + V\tau + \tilde{j}_n \quad (2.4)$$

où τ est une variable temporelle qu'on associera à un pas de temps. Le paramètre $V = \mu/\tau$ possède la dimension d'une vitesse. Ce paramètre V représente une vitesse moyenne d'advection dans la marche aléatoire, dans le sens où elle traduit un déplacement moyen systématique (de valeur μ) à chaque pas de temps. La variable aléatoire \tilde{j}_n , obtenue à partir de j_n par translation, suit donc une loi lognormale centrée de pdf φ définie par :

$$\varphi(x) = \frac{1}{\sigma \sqrt{2\pi}} \frac{1}{(x + \mu)} \exp\left(-\frac{(\ln(x + \mu) - \mu)^2}{2\sigma^2}\right) \quad \text{avec} \quad x > -\mu$$

la marche aléatoire, dans sa dimension spatiale, peut encore s'écrire : $x_{n+1} = x_n + \mu + \tilde{j}_n$ en fonction de la variable aléatoire $\tilde{j}_n = \frac{1}{\sigma} j_n$ qui, elle, suit une loi log-normale centrée réduite (notée $L(0, 1)$) . La figure (2.1) ci-dessous donne une représentation de la fonction de densité de probabilité des sauts.

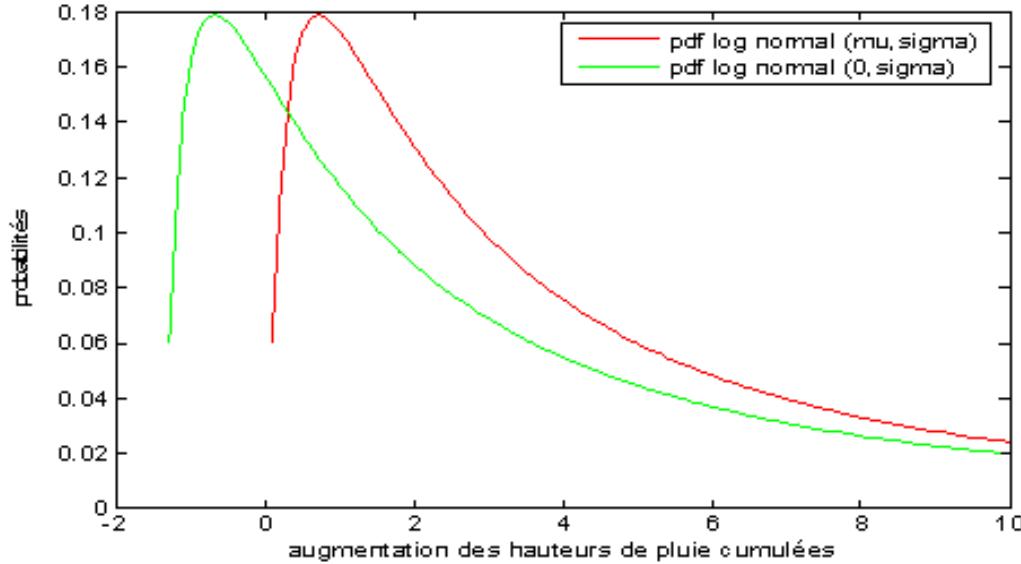


FIGURE 2.1 – Distributions log-normales de \tilde{j}_n et j_n simulant les sauts de la marche aléatoire.

En dimension temporelle, on rappelle qu'un "temps d'attente" est un temps d'attente entre deux relevés de précipitations non nulles, donc que ce temps est strictement positif. Si τ est le pas de temps de la marche aléatoire et W_n est une variable aléatoire α -stable alors la durée entre deux sauts est donnée par $\tau^{1/\alpha} W_n$ [58]. Avec comme condition aux limites $x_0 = 0$ et $t_0 = 0$, la marche aléatoire est alors résumée par l'équation suivante :

$$\begin{cases} x_{n+1} = x_n + V + \ell \tilde{j}_n \\ t_{n+1} = t_n + \tau^{\frac{1}{\alpha}} W_n \end{cases} \quad (2.5)$$

Dans l'équation (2.5) nous avons introduit un pas d'espace ℓ . Le rapport $D = \frac{\ell^2}{\tau}$ constitue le coefficient de dispersion de la marche aléatoire. Comme on le verra dans la section suivante, ce modèle admet comme limite hydrodynamique (quand τ et ℓ tendent vers zéros mais D reste constant) un processus stochastique dont la loi de probabilité marginale est gouvernée par l'équation fractionnaire de Fokker Planck.

Une première mise en oeuvre du modèle (2.5) est illustrée par les trajectoires sur la figure (2.2). Les paramètres de la marche ont été ajustés avec la série temporelle $\ell = 0,0548$, $\alpha = 0,95$,

$D = 0,003$, $V = 2,788$. Cette figure montre un comportement similaire entre la série des pluies cumulées et plusieurs simulations de trajectoires. Du moins sur les 11000 premières mesures de 1968 à 1997. Sur le reste de sa trajectoire la série temporelle des pluies cumulées présente un fléchissement qui n'est pas reproduit par les simulations. Ce fléchissement est dû à des temps d'attente inhabituellement longs liés à des périodes de sécheresse des dernières années entre 1997 et 2008. D'autre part, en multipliant le nombre de simulations, on peut voir apparaître des trajectoires simulées qui présentent des écarts importants avec la série expérimentale. Des tels exemples de simulations sont montrés sur la figure (2.3). Ce fait est directement lié à la loi stable des temps d'attente qui augmentent la probabilité d'occurrence des durées de valeurs extrêmes. Trois méthodes vont être utilisées pour réduire ces écarts de manière à mieux cerner la réalisation dont on dispose et dont on veut établir le processus qui la régit. Ces trois méthodes sont :

1. "Tempérer" les temps d'attente en les bornant par une distribution exponentielle de moyenne lambda [57].
2. "Trier" les temps d'attente en leur imposant une durée voisine de 13660 jours, longueur du relevé des hauteurs de pluies.
3. "Sélectionner" les temps d'attente de manière à présenter les mêmes proportions de temps d'attente entre deux pluies non consécutives que la série expérimentable.

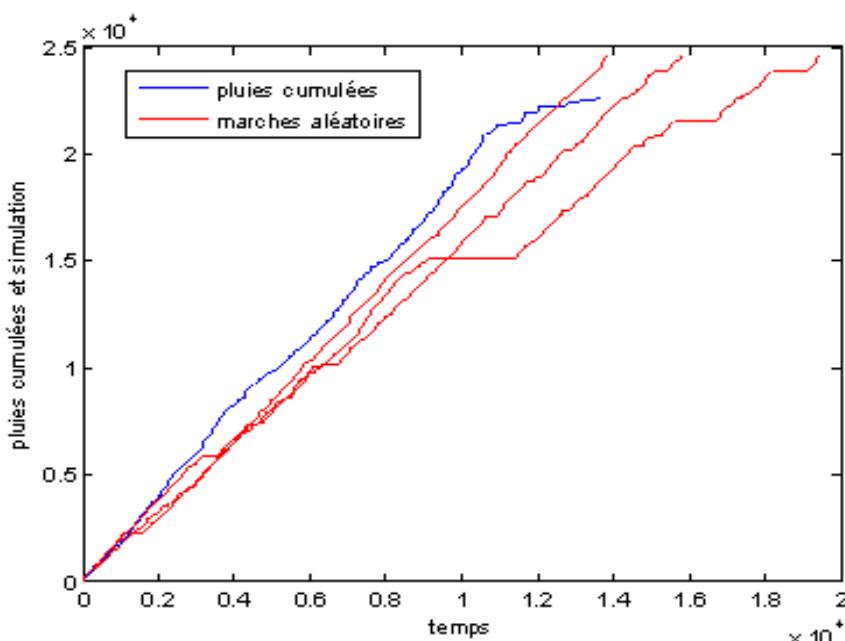


FIGURE 2.2 – Hauteurs de pluie cumulées simulées par des marches aléatoires non soumises aux caractéristiques de la série temporelle des pluies cumulées

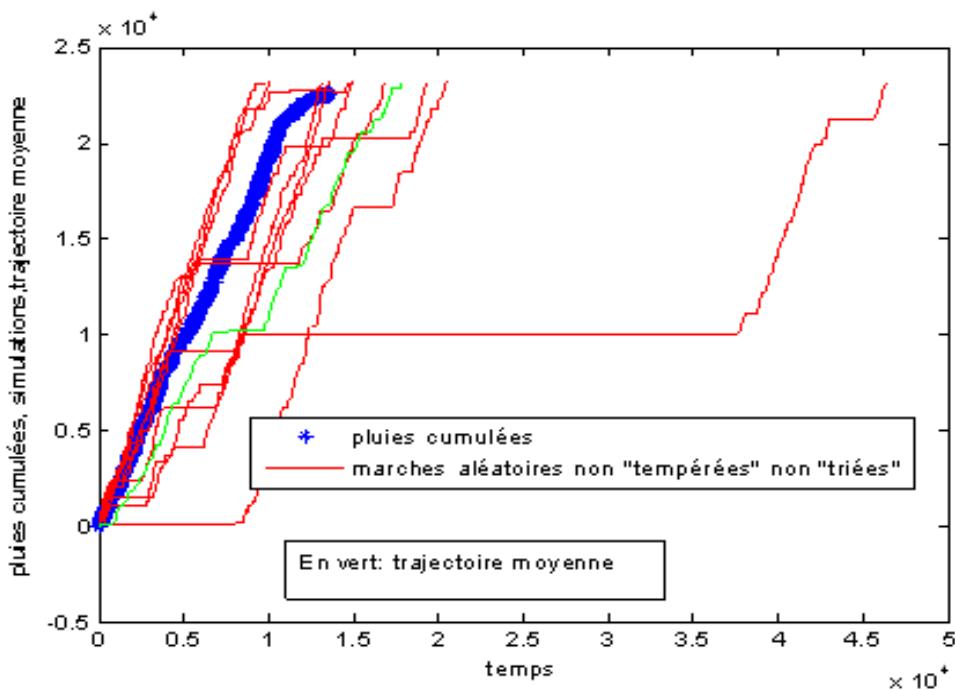


FIGURE 2.3 – Exemples de trajectoires de pluie cumulées simulées présentant des grands écarts par rapport à la série temporelle des pluies cumulées.

Ajustement de la distribution des temps d'attente

Temps d'attente α - stables "tempérés".

Parmi les temps d'attente α -stables utilisés pour simuler les trajectoires ci-dessus certains ont une durée trop importante par rapport au processus que l'on souhaite décrire. Pour limiter la probabilité des événements extrêmes, une méthode consiste à borner la distribution des temps $\tau^{1/\alpha} W_n$ par une distribution "sans mémoire", exponentielle de moyenne λ . [18] [7].

En effet, un modèle subdiffusif présentant une distribution α -stable des temps d'attente donne des simulations qui surestiment les très longues durées. De telles situations ont été observées et corrigées par l'introduction d'une loi de distribution " α -stable tempérées" [18]. Une telle loi est définie par la transformée de Laplace de sa fonction densité de probabilité dont les moments de tous ordres sont finis, soit $f(u) = \exp(-\lambda^\alpha - (u+\lambda)^\alpha)$.

Techniquement, il s'avère que la valeur de λ devra être relativement élevée pour acquérir suffisamment de valeurs ayant les caractéristiques souhaitées dans le cadre de notre phénomène pluvieux local. La figure (2.4) montre bien un encadrement très resserré des hauteurs de pluie cumulées par des simulations de marches aléatoires non tempérées, quasiment linéaires. Par ailleurs, il subsiste des temps exagérément longs qui vont au-delà de la longueur de la série expérimentale.

En revanche, du fait des temps "tempérés" la figure (2.5) affiche pour certaines trajectoires, une tendance à marquer un fléchissement à l'instar de celui qui apparaît dans la partie supérieure de la courbe expérimentale bien que ne l'épousant pas parfaitement. Il reste cependant des distributions présentant des temps d'attente exagérément longs qui doivent être éliminés pour mieux approcher la distribution expérimentale.

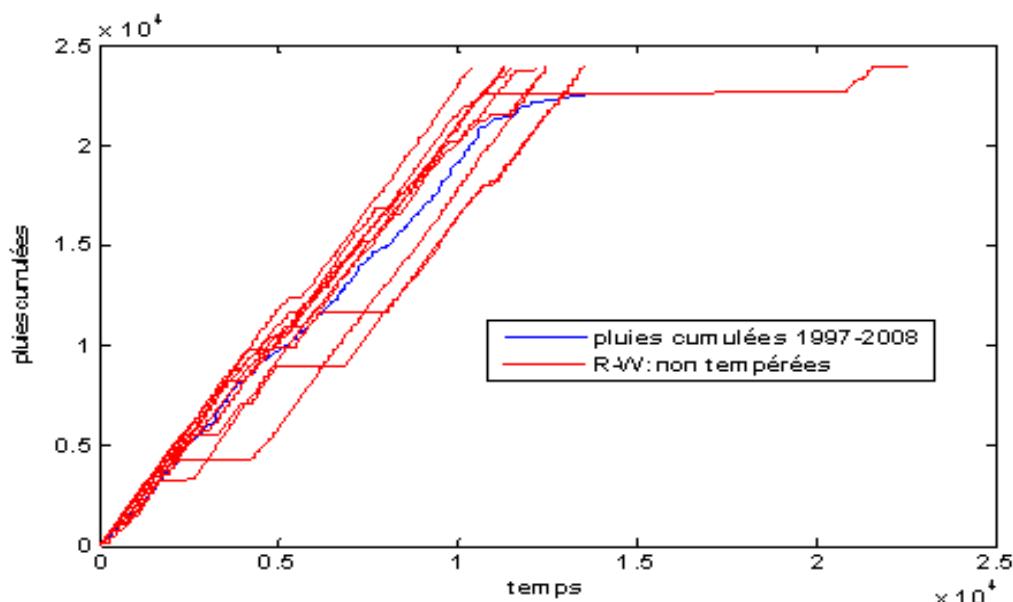


FIGURE 2.4 – Hauteurs de pluie cumulées simulées par des marches aléatoires non tempérées, non "triées".

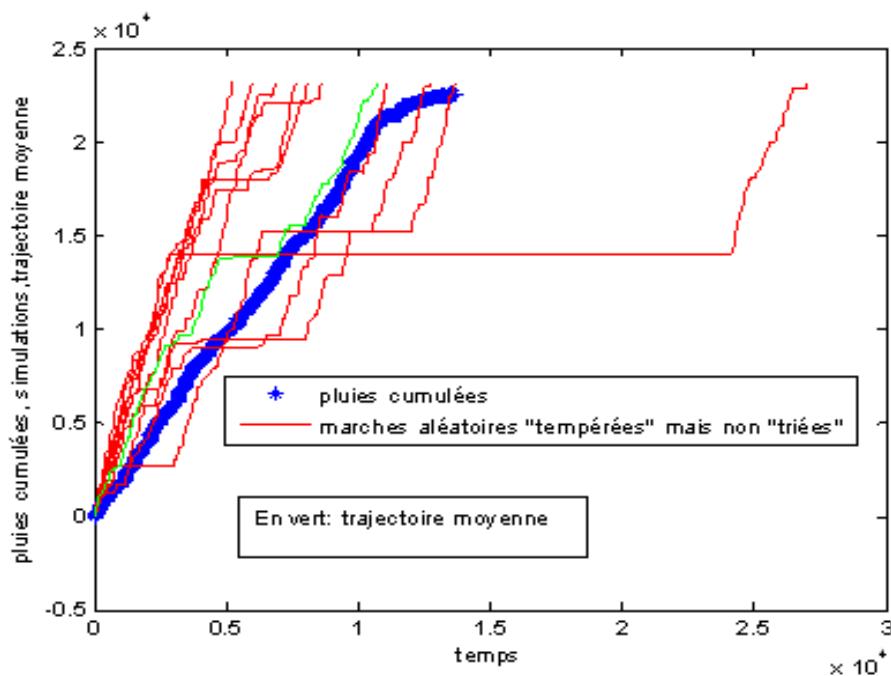


FIGURE 2.5 – Hauteurs de pluie cumulées simulées par des marches aléatoires tempérées mais non "triées".

Temps d'attente α -stables "triés"

Le bornage exponentiel ci-dessus produit des temps d'attente dépassant largement la dimension temporelle de la série expérimentale dont on dispose. Pour éliminer les distributions démesurées, on impose au temps de la marche aléatoire de ne pas dépasser de plus de 10 à 20% le temps du relevé expérimental des hauteurs de pluie considérées. La figure (2.6) ci-dessous révèle bien l'élimination de simulations n'avoisinant pas la longueur temporelle de la série expérimentale et un resserrement des trajectoires simulées autour de la trajectoire expérimentale. On note un certain fléchissement et l'élimination des trajectoires non comparables à la trajectoire expérimentale. Ces résultats ont été obtenus avec les paramètres suivants $V=9$, $\alpha = 0,85$, $D = 1$, $\lambda = 10^{-6}$.

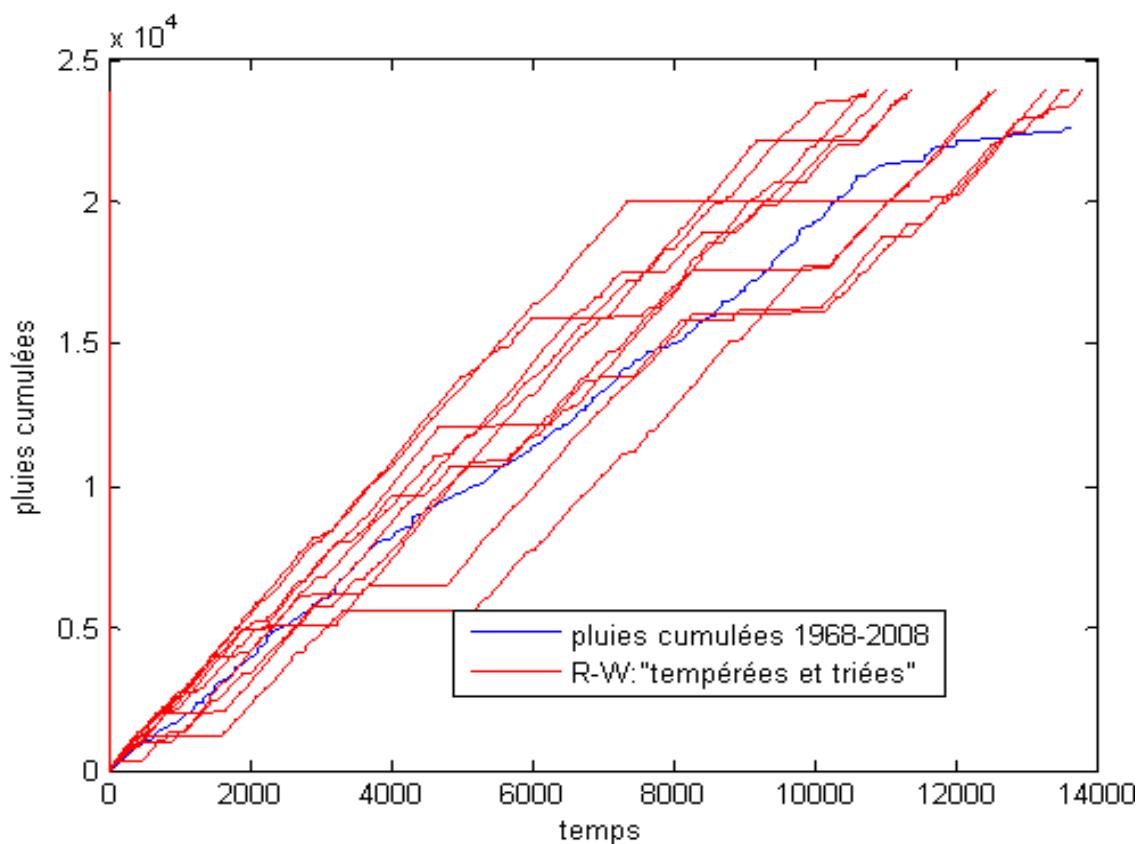


FIGURE 2.6 – Hauteurs de pluie cumulées simulées par des marches aléatoires "Tempérées et triées".

Temps d'attente α -stables "sélectionné"

La seule réalisation du processus dont on dispose, les hauteurs de pluie cumulées, présente 21% "de sauts" qui se succèdent à un jour d'intervalle et par conséquent 79% de temps d'attente supérieurs à un jour entre deux sauts. Graphiquement, ceci se traduit par des "plateaux" au cours

desquels lesquels la hauteur de pluie cumulée n'évolue pas. L'idée retenue ici est de sélectionner, parmi les simulations de trajectoires tempérées et triées, celles qui présentent un taux de temps d'attente comparable; entre 70 et 90%. Il apparaît alors un "resserrement" des trajectoires autour des hauteurs de pluie cumulées. La figure (2.7) ci-après montrent bien un resserrement total autour de la courbe expérimentale lorsque les temps d'attente cumulent successivement les ajustements dits "tempérés", "triés" puis enfin "sélectionnés". Pour une lecture plus aisée des graphiques qui permettra de mieux apprécier la corrélation entre la série expérimentale et les trajectoires simulées, on peut réduire l'ensemble des trajectoires à la trajectoire moyenne d'une part et situer la série expérimentale par rapport aux déciles des trajectoires.

Situer la courbe expérimentale par rapport aux déciles des simulations

Pour chaque trajectoire de la marche aléatoire, apparaît en abscisse son temps opérationnel propre, i-e défini aléatoirement. Afin de représenter la trajectoire moyenne et les "déciles" il est nécessaire de prendre les points des trajectoires qui ont une même abscisse. Un pas de temps de un jour étant choisi et les trajectoires étant "en escalier", il suffit de prendre en ordonnée, celle du point d'abscisse la plus voisine "à gauche" du jour considéré. Les figure (2.8) et figure (2.9) ci-dessous illustrent des simulations ramenées à des abscisses comptées en jour, leur trajectoire moyenne qui voisine fortement la série expérimentale et enfin ; la position de celle-ci parmi les quantiles des trajectoires. Il faut noter le "fléchissement" de la trajectoire moyenne qui marque tout l'intérêt des distributions tempérées des temps d'attente. Ayant cerné le processus qui approche notre phénomène pluvieux local, on va, par la méthode de Monte- Carlo, évaluer les hauteurs de pluie les plus probable pour un temps donné et les comparer à la série initiale.

Détermination graphique de la P.D.F. de la marche aléatoire et évaluation des hauteurs de pluie les plus probables à une date donnée.

Un processus étant bien déterminé par sa fonction de densité de probabilité, nous utiliserons ici la méthode de Monte-Carlo : après avoir simulé 5000 trajectoires répondant au critère explicité ci-dessus, pour un temps donné, on établit la liste des hauteurs de pluie cumulées correspondantes et l'on construit les histogrammes de ces valeurs en faisant varier le nombre de "boites" (ou "bins"). Il doit alors se dégager une tendance où apparaît la valeur de la variable qui a la plus forte probabilité d'être réalisée. Le nombre optimal de boites se situant aux environs de la

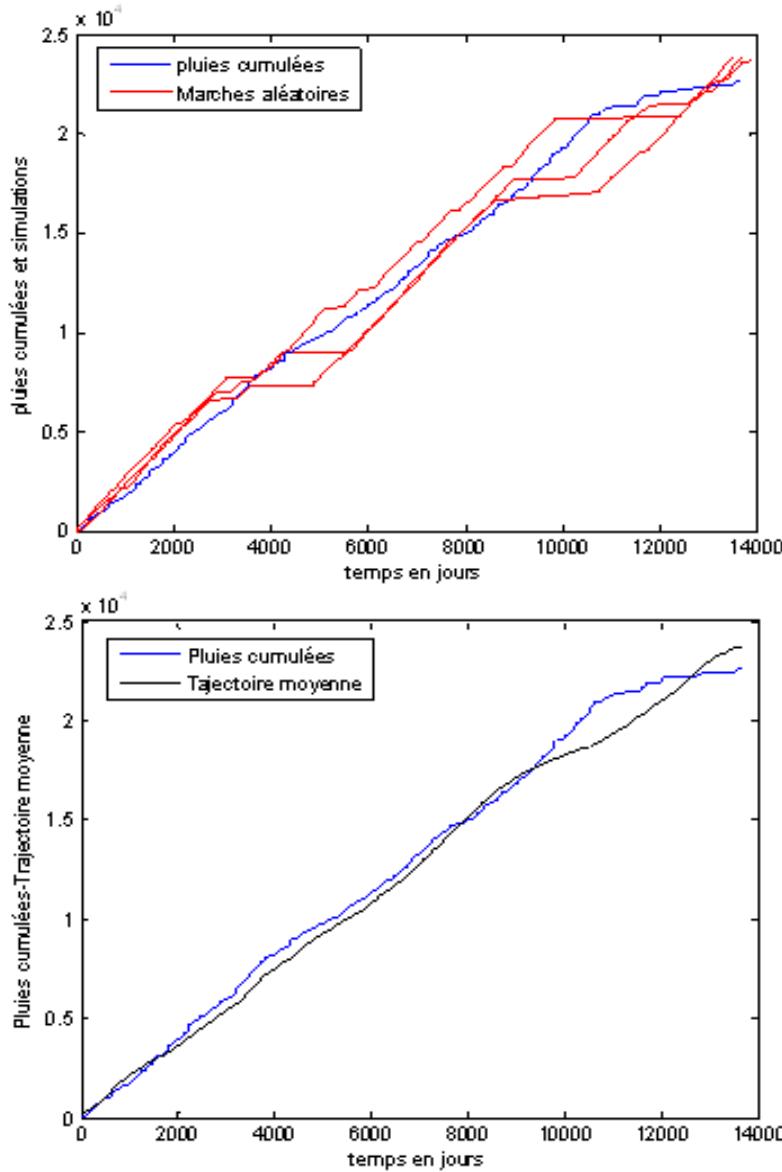


FIGURE 2.7 – 1^{ière} figure, en haut : hauteurs de pluie cumulées simulées par des marches aléatoires "Tempérées, triées et sélectionnées", elles "voisinent" bien la trajectoire expérimentale. 2^{nde} figure, en bas : la trajectoire moyenne des simulations montre bien un certain fléchissement. $V=9$, $\alpha = 0,85$ – $D = 1$ – $\lambda = 10^{-7}$ – sélection entre 0,7 et 0,9 soit entre 70 et 90 % de temps d'attente supérieurs à 1 jour entre deux pluies. Temps entre 13000 et 14000 jours.

racine carrée de l'effectif de la série. C'est cette multiplication des histogrammes qui apparaît à la figure (2.10). La figure (2.11) illustre une application de ce principe pour trois valeurs de t où ne figure que l'histogramme retenu. Enfin la Figure (2.12) montre l'utilisation qui peut en être faite pour prédire la hauteur de pluies cumulées la plus probable en 2018. Le tableau (2.1) ci-dessous met en parallèle, à diverses dates, la réalité expérimentale et les hauteurs de pluies les plus probables, résultats de simulations de type marche aléatoire. Les écarts entre les deux résultats figurent aussi dans le tableau et si ces écarts sont importants en début de la

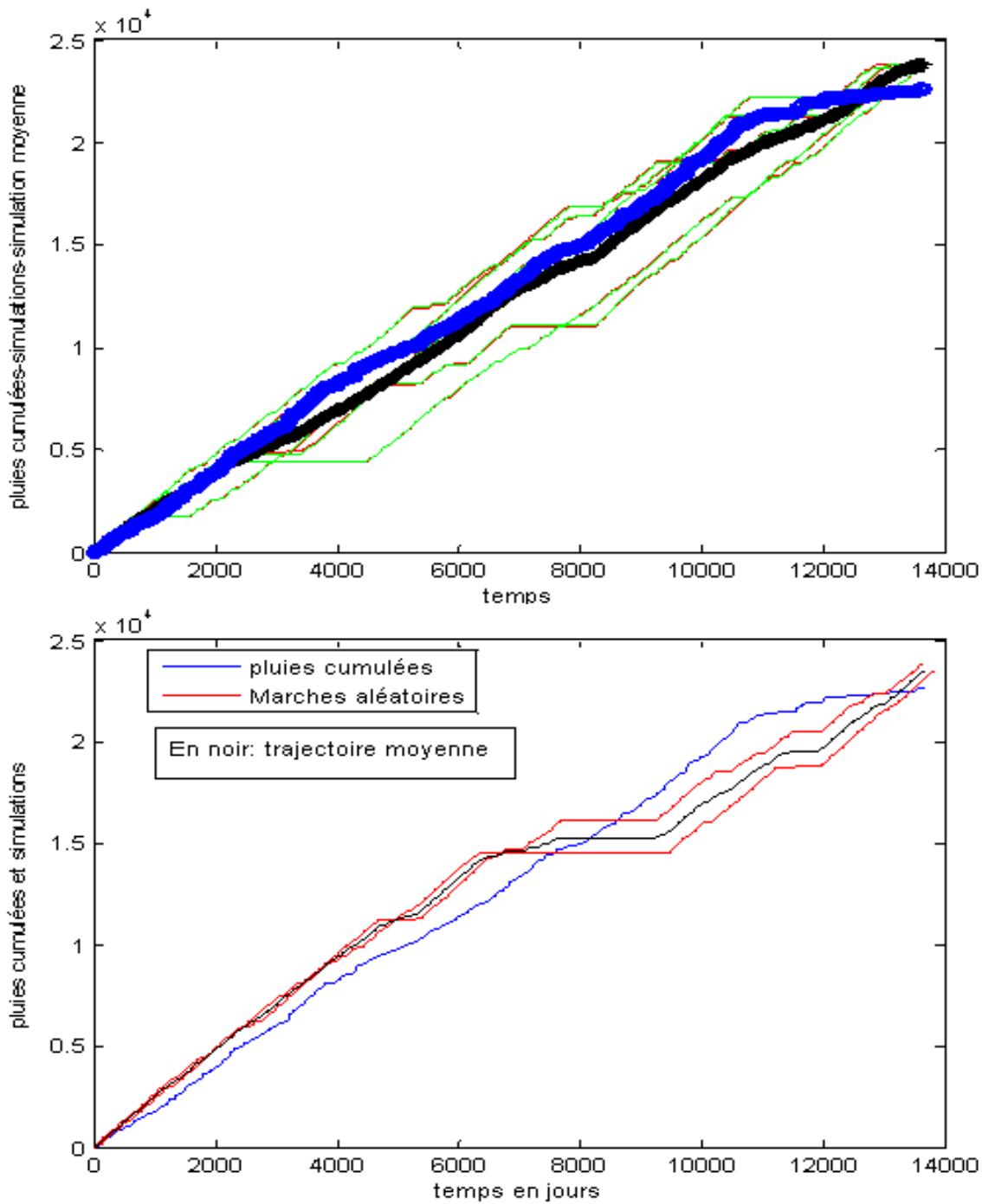


FIGURE 2.8 – 1^{ère} figure en haut : en rouge et vert apparaissent les simulations, "Tempérées, triées et sélectionnées", d'abscisses ; les temps d'attente α -stables d'une part et rapportées à des temps entiers d'autre part. En noir : trajectoire moyenne. En bleu : La courbe expérimentale. $V=9$, $\alpha=0,85$, $D=1$, $\lambda=10^{-7}$, sélection entre 0,7 et 0,9, temps entre 13000 et 14000 jours. 2^{nde} figure en bas : vue plus "fine" portant sur 2 simulations.

trajectoire ils sont quasiment insignifiants dans son dernier quart.

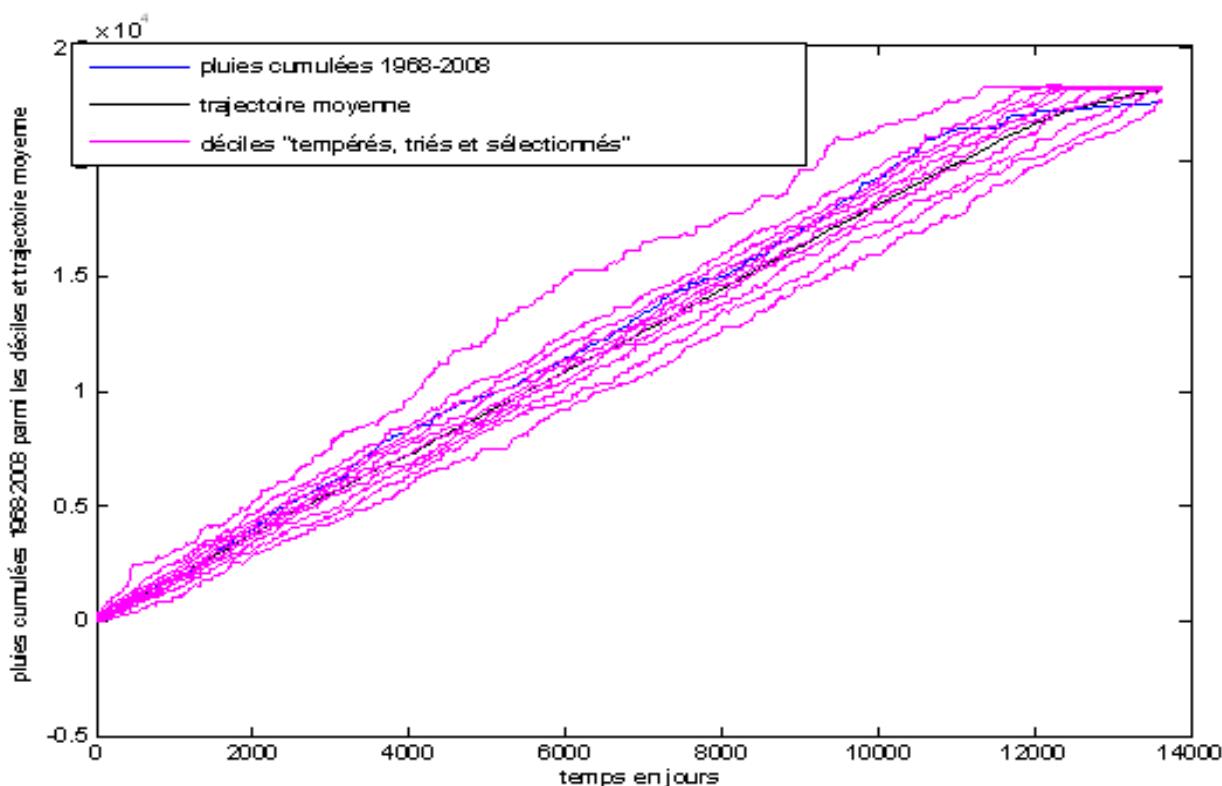


FIGURE 2.9 – Hauteurs de pluie cumulées simulées par des marches aléatoires "Tempérées, triées et sélectionnées", parmi les déciles; trajectoire moyenne. $V=9$, $\alpha = 0.85$, $D = 1$, $\lambda = 10^{-7}$, sélection entre 0,7 et 0,9, temps entre 13000 et 14000 jours.

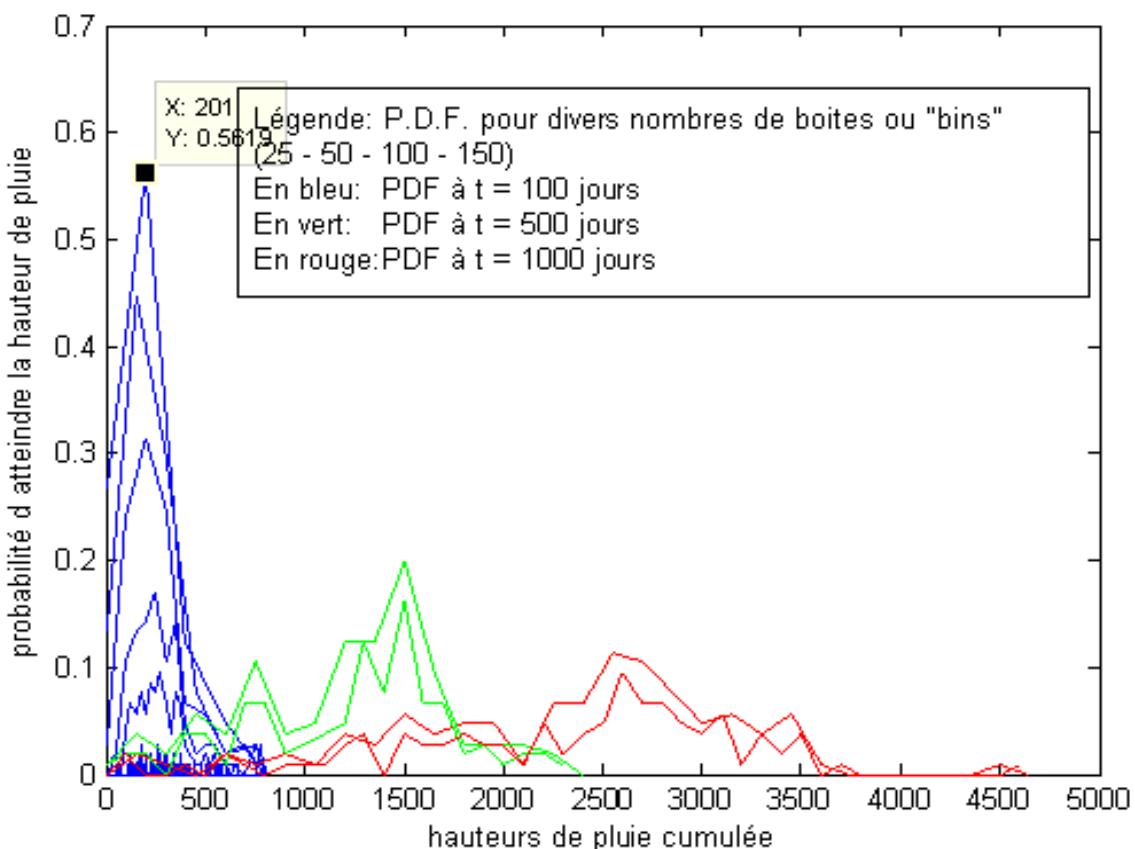


FIGURE 2.10 – P.D.F obtenu à partir de marches aléatoires "tempérées, triées et sélectionnées".

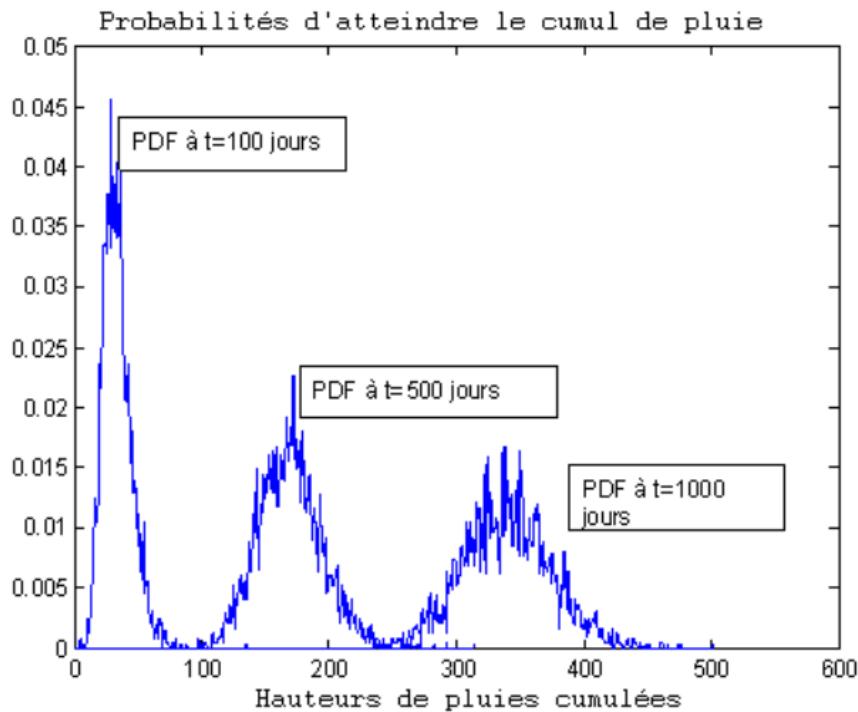


FIGURE 2.11 – P.D.F obtenu à partir de marches aléatoires "tempérées, triées et sélectionnées".

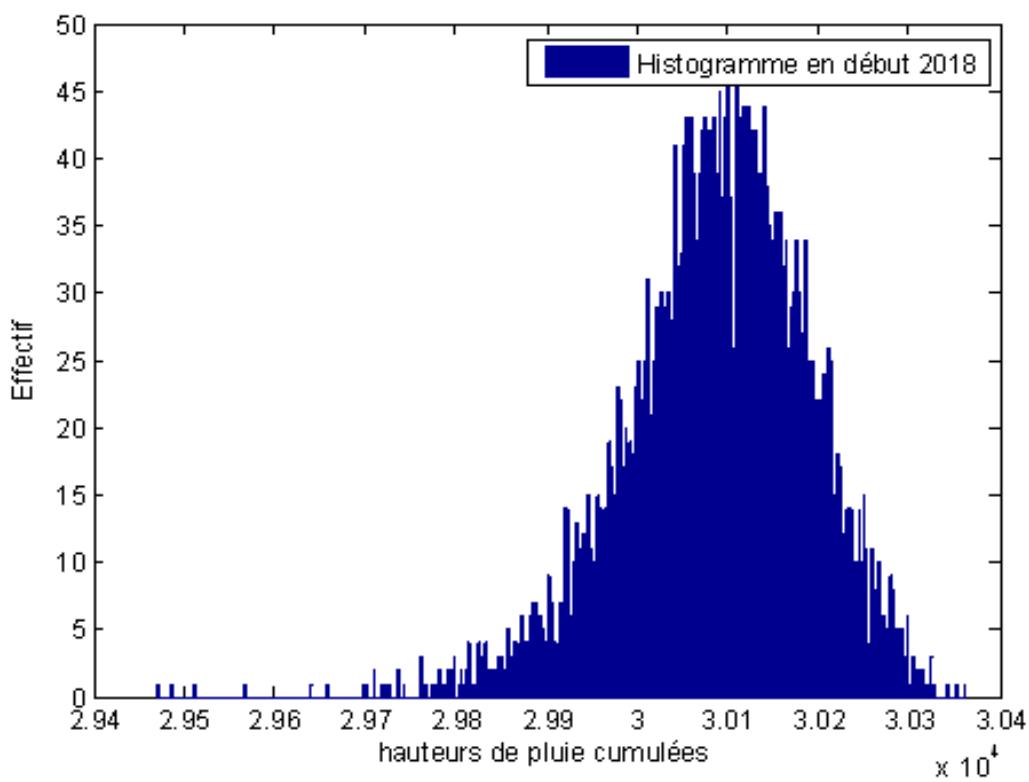


FIGURE 2.12 – P.D.F obtenu à partir de marches aléatoires "tempérées, triées et sélectionnées" extrapolées à l'an 2018.

Evaluation de la pertinence du modèle avancé.

Hauteurs de pluies les plus probables à dates données, issues des simulations de trajectoires construites à partir de temps "tempérés", "triés" et "sélectionnés".

Date/ Num. jour	Hauteur de pluie cumulée mesurée (en mm)	Hauteur de pluie cumulée modélisée (en mm)	Ecart relatif
10 Avril 1968 100 ^{ième} jour	121	140	+15%
15 Mai 1969 500 ^{ième} jour	950	960	+1%
30 Sept. 1970 1000 ^{ième} jour	1707	1800	+5%
20 Sept. 1981 5000 ^{ième} jour	9 742	9020	-7%
10 Nov. 1988 8000 ^{ième} jour	15000	15000	0%
05 Mai 1994 10000 ^{ième} jour	19246	18000	-6%
16 Février 1997 Fléchissement 10640 ^{ième} jour	20918	18700	-10%
06 Oct. 1999 11600 ^{ième} jour	21678	20440	-5%
06 Juillet 2002 12600 ^{ième} jour	22 283	21100	-5%
24 Mai 2008 13659 ^{ième} jour	22593	22 650	+ 0.3%
24 Mai 2018 17 309 ^{ième} jour		25000	

TABLE 2.1 – Le modèle approche d'autant mieux la réalité expérimentale que le nombre d'observations est grand.

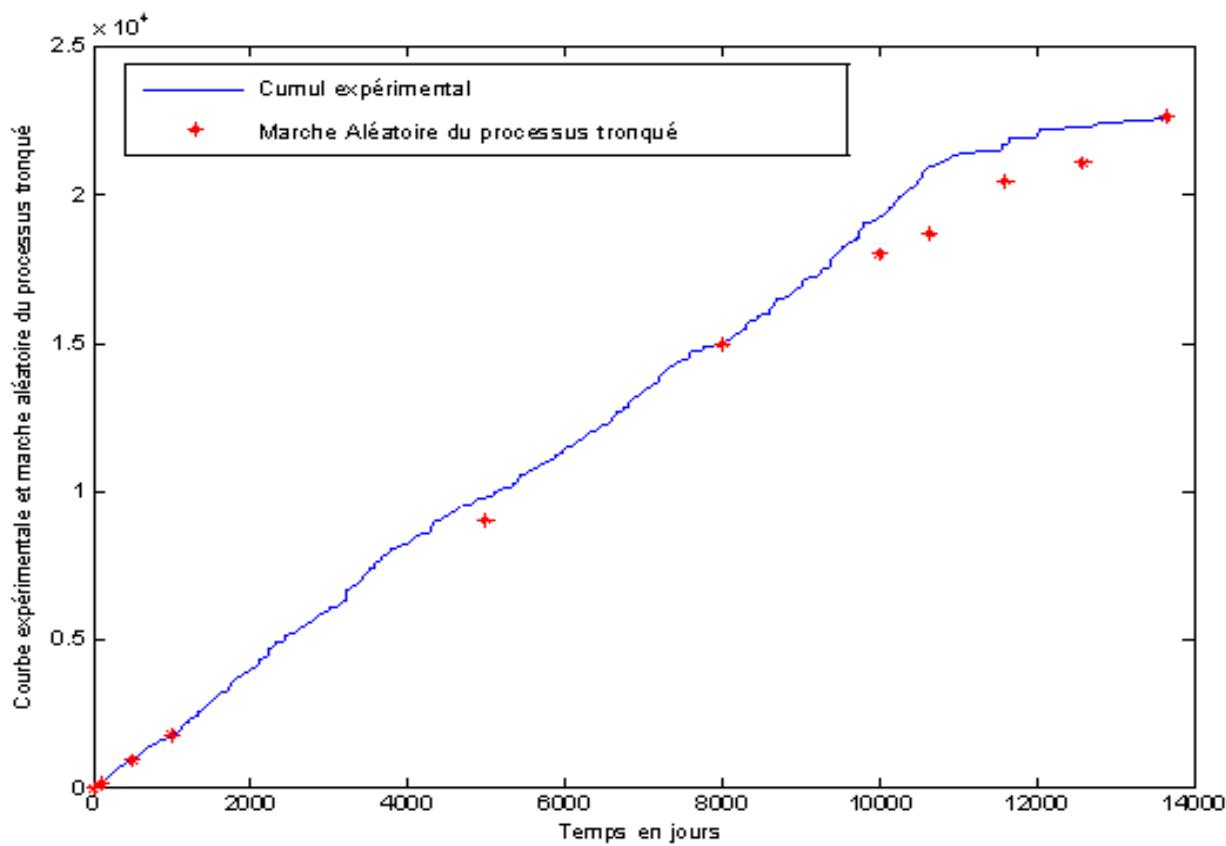


FIGURE 2.13 – Processus "tronqué" : Cumuls expérimentaux et cumuls les plus probables obtenus par marche aléatoire.

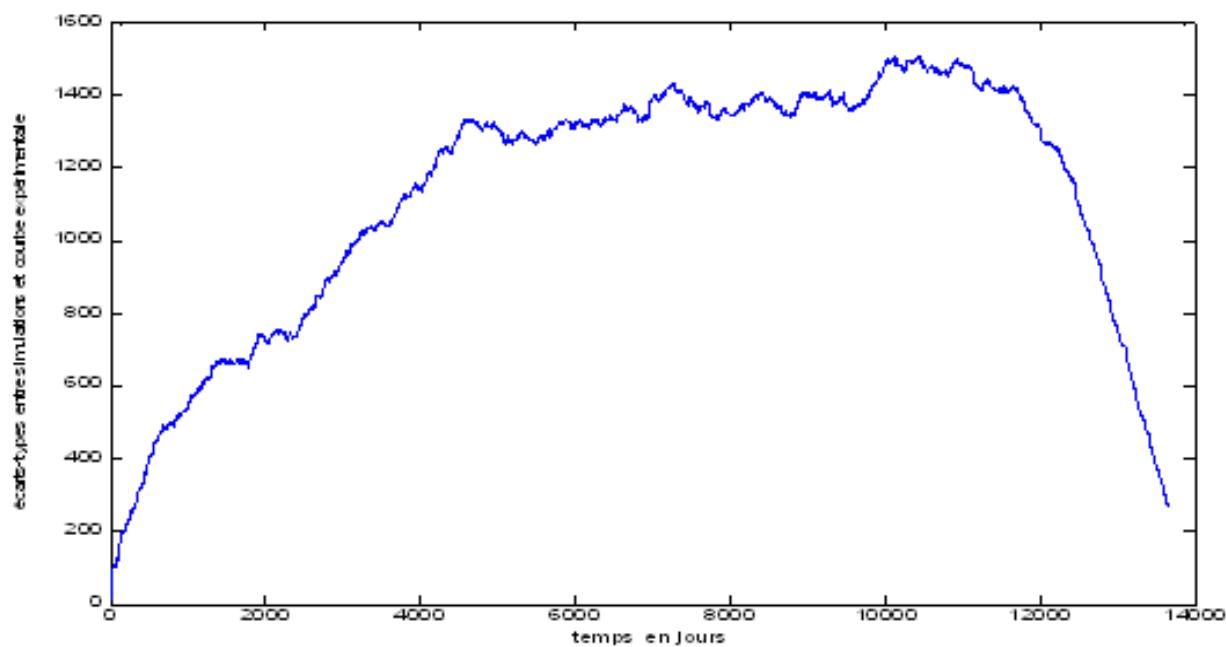


FIGURE 2.14 – Processus "tronqué" : Ecarts-types des distances : courbe expérimentale- simulations par marche aléatoire "Tempérées, triées et sélectionnées".

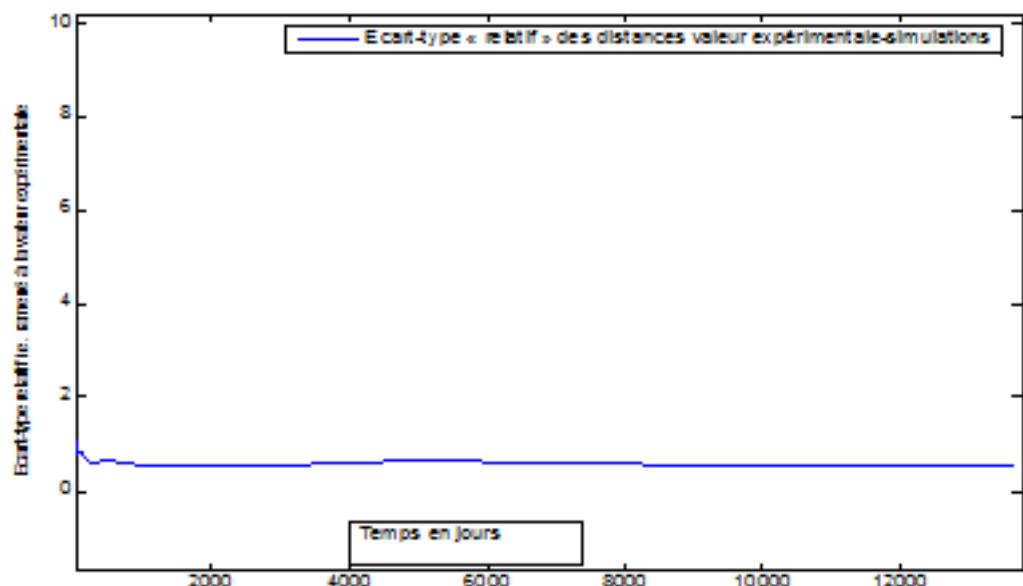


FIGURE 2.15 – Processus “tronqué” : Ecart-type “relatif” des distances entre la valeur expérimentale et les simulations de marches aléatoires.

2.2 Seconde approche du processus pluvieux local

Limites des marches aléatoires

L'objet de ce passage à la limite pour la marche aléatoire ci-dessus est de tenter de gérer le phénomène de manière continue en temps et en espace pour déterminer la probabilité pour qu'une trajectoire suivie par une hauteur de pluie cumulée prenne une valeur x à un instant t . i.e. déterminer la distribution ou encore $P(x,t)$, la densité de trajectoires passant en x à t donné, la probabilité pour que la variable aléatoire que constitue la hauteur de pluie cumulée, prenne la valeur x au temps t . Barkai et al [3] ont montré que dans le cas de distributions α -stables avec $\alpha < 1$ donc de moyenne infinie, en limite d'une C.T.R.W. on obtenait une F.F.P.E. et semblablement, Marie-Christine Néel et al [63] montraient la même chose dans le cas $\alpha > 1$. Ici, par l'estimation du flux qui traduit la tendance à l'étalement des marcheurs on parvient à un model F.F.P.E.

En effet, cette densité dépend donc du flux de trajectoires (noté F) qui traversent un intervalle $[x, x+dx]$ entre t et $t + dt$. Or ce flux est constitué des trajectoires qui, partant d'une position x_n de la marche aléatoire réalisent une augmentation moyenne μ suivie d'un autre saut aléatoire distribué suivant une loi de moyenne nulle et d'écart-type σ . Ces derniers saut étant de moyenne nulle, il en est qui peuvent se faire en deçà de x comme d'au delà de x . C'est la différence d'entre ces deux qui constitue le flux. Comme pour le principe de conservation de la masse lorsqu'il s'agit de particules on considérera ici que toutes les trajectoires participent au mouvement. La variation de densité de trajectoires en x par unité de temps est aussi la variation spatiale du flux. (Au signe près).

Soit :

$$\partial_t P(x, t) dx = F\left(x - \frac{dx}{2}, t\right) - F\left(x + \frac{dx}{2}, t\right) \quad (2.6)$$

Ou encore :

$$\partial_t P(x, t) = - \partial_x F(x, t) \quad (2.7)$$

Déterminons $P(x, t)$ à partir de la marche aléatoire précédente où le déplacement spatial est de la forme :

$$X_{n+1} = X_n + V\tau + \tilde{j}_n$$

Pour passer d'une marche aléatoire discrète à une marche aléatoire continue en temps, il est nécessaire de faire tendre simultanément le temps et l'amplitude des sauts vers 0. La hauteur de pluie cumulée est alors supposée évoluer sur une échelle de temps τ et dans un premier temps, réalise tous les $n\tau$ un saut d'amplitude $\ell' \widetilde{\mathbf{j}_n}$.

En substituant à $\widetilde{\mathbf{j}_n}$ qui suit une loi $L(0, \sigma^2)$ la variable aléatoire $\widetilde{\mathbf{j}_n} = \frac{1}{\sigma} \widetilde{\mathbf{j}_n}$ qui elle, suit une loi $L(0, 1)$ la marche aléatoire s'écrit alors :

$$X_{n+1} = X_n + V\tau + \ell' \widetilde{\mathbf{j}_n}$$

Où . $\widetilde{\mathbf{j}_n}$ suit une loi $L(0, 1)$ définie précédemment.

Ultérieurement on envisagera le cas, plus proche de notre étude où un temps d'attente s'interpose entre deux précipitations pluvieuses.

Les sauts ont lieu à intervalle régulier τ

Les sauts sont supposés indépendants et leur amplitude admet pour pdf :

$$\varphi_\ell(y) = \frac{1}{\ell} \varphi\left(\frac{y}{\ell}\right) \text{ où}$$

$$\varphi(x) = \frac{1}{\sigma \sqrt{2\pi}} \frac{1}{(x+\mu)} \exp\left(-\frac{(\ln(x+\mu)-\mu)^2}{2\sigma^2}\right) \text{ Avec } x > -\mu$$

Pour qu'une trajectoire prenne la valeur x , il y a trois cas :

1. Soit la trajectoire est en x par le saut moyen $V\tau$ ce qui se produit avec la probabilité $V P(x, t)$.
2. Soit, après un saut moyen elle part d'une valeur $x - y$ ($y > 0$), en deçà de x ou de $x + Y$ au-delà de x et fait, dans les deux cas, un saut aléatoire suffisant pour dépasser x .
3. Soit Partir d'une valeur $x - y$, en deçà de x et faire un saut aléatoire suffisant pour dépasser x . Ceci se produit avec une probabilité

$$\int_y^\infty \frac{1}{\ell} \varphi\left(\frac{z}{\ell}\right) dz = \int_{\frac{y}{\ell}}^\infty \varphi(z) dz = \Phi\left(\frac{y}{\ell}\right)$$

Si on pose : $\Phi_+(y) = \int_y^\infty \varphi(z) dz$ donc $\Phi'_+(y) = -\varphi(y)$

La hauteur de pluie cumulée évolue une fois dans chaque intervalle de temps de durée τ , donc elle évolue entre t et $t + dt$ avec la probabilité dt/τ . La probabilité pour qu'une trajectoire évolue de manière à dépasser x en venant de $x-y$ entre t et $t + dt$ est : $\frac{dt}{\tau} \int_0^\infty P(x-y) \Phi_+(\frac{y}{\ell}) dy$.

De même, la probabilité pour qu'une trajectoire évolue de manière à dépasser x en venant de $x+y$ entre t et $t + dt$ est : $\frac{dt}{\tau} \int_0^\infty P(x+y) \Phi_-(\frac{y}{\ell}) dy$

Avec $\Phi_-(y) = \int_{-\infty}^{-y} \varphi(z) dz$ donc $\Phi'_-(y) = \varphi(-y)$

La probabilité de traverser x est alors $\frac{dt}{\tau} \int_0^\infty [P(x-y) \Phi_+(\frac{y}{\ell}) - P(x+y) \Phi_-(\frac{y}{\ell})] dy$

Soit :

$$\begin{aligned} & \frac{dt}{\tau} \int_0^\infty \{[P(x-y) - P(x)] \Phi_+(\frac{y}{\ell}) - [P(x+y) - P(x)] \Phi_-(\frac{y}{\ell})\} dy \\ & + \frac{dt}{\tau} \left[\int_0^\infty [P(x) \Phi_+(\frac{y}{\ell}) - P(x) \Phi_-(\frac{y}{\ell})] dy \right] \end{aligned}$$

Soit, en faisant le changement de variable $\frac{y}{\ell} = Y$

$$dt \left[\int_0^\infty -\frac{P(x-\ell Y) - P(x)}{-\ell Y} \frac{2\ell^2}{2\tau} Y \Phi_+(Y) - \frac{P(x+\ell Y) - P(x)}{\ell Y} \frac{2\ell^2}{2\tau} Y \Phi_-(Y) \right] dY$$

$$+ dt \left[\int_0^\infty [P(x) \frac{\ell}{\tau} \Phi_+(Y) - P(x) \frac{\ell}{\tau} \Phi_-(Y)] \right] dY$$

$$\text{Lorsque } \ell \text{ tend vers } 0 : \lim_{\ell \rightarrow 0} \frac{P(x-\ell Y) - P(x)}{-\ell Y} = \partial_x P(x)$$

et en faisant apparaître $D = \frac{\ell^2}{2\tau}$, il vient alors la somme de 2 termes A + B avec :

$$A = -2D \partial_x P(x) dt \left[\int_0^\infty Y [\Phi_+(Y) + \Phi_-(Y)] dY \right]$$

et

$$B = 2DP(x) dt \left[\int_0^\infty \frac{\Phi_+(Y) - \Phi_-(Y)}{\ell} dY \right]$$

Le crochet de A s'intègre par parties :

$$\int_0^\infty Y [\Phi_+(Y) + \Phi_-(Y)] dY = \left[\frac{Y^2}{2} (\Phi_+(Y) + \Phi_-(Y)) \right]_0^\infty - \int_0^\infty \frac{Y^2}{2} [-\varphi(Y) + \varphi(-Y)] dY$$

Vu que $-\varphi(Y)$ est la dérivée de $\Phi_+(Y)$ et $\varphi(-Y)$ celle de $\Phi_-(Y)$.

Le crochet de cette dernière formule est nul soit :

$$\left[\frac{Y^2}{2} (\Phi_+(Y) + \Phi_-(Y)) \right]_0^\infty = 0.$$

En effet :

$$\Phi_+(Y) = \int_Y^\infty \varphi(z) dz = \int_Y^\infty \frac{1}{\sigma \sqrt{2\pi}} \frac{1}{(z+\mu)} \exp\left(-\frac{(\ln(z)+\mu)-\mu}{2\sigma^2}\right) dz$$

Laquelle se comporte à l'infini comme $\int_Y^\infty \frac{1}{z} \exp(-(\ln(z))^2) dz$

Or dès que z est supérieur à e^3 , $(\ln(z))^2 > \ln(z^3)$ donc $-(\ln(z))^2 < -\ln(z^3)$ et de même, $\exp(-(\ln(z))^2) < \exp(-\ln(z^3))$ soit $\exp(-(\ln(z))^2) < 1/z^3$

encore $\int_Y^\infty \frac{1}{z} \exp(-(\ln(z))^2) dz < \int_Y^\infty 1/z^4 dz$

Or $\int_Y^\infty 1/z^4 dz = [-1/3z^3]_Y^\infty = 1/3Y^3$

Enfin $\frac{Y^2}{2} \cdot \frac{1}{3Y^3}$

Se comporte comme $1/y$ et tend vers zéro si y tend vers l'infini.

La seconde partie de la formule s'écrit :

$$\int_0^\infty \frac{Y^2}{2} [\varphi(Y) - \varphi(-Y)] dY = \int_0^\infty \frac{Y^2}{2} \varphi(Y) dY - \int_0^\infty \frac{Y^2}{2} \varphi(-Y) dY$$

Changeons de variable dans la 2nde intégrale :

$$\begin{aligned} \int_0^\infty \frac{Y^2}{2} \varphi(-Y) dY &= \int_{-\infty}^0 \frac{Y^2}{2} \varphi(Y) dY \\ \int_0^\infty \frac{Y^2}{2} [\varphi(Y) - \varphi(-Y)] dY &= \int_0^\infty \frac{Y^2}{2} \varphi(Y) dY - \int_{-\infty}^0 \frac{Y^2}{2} \varphi(Y) dY \end{aligned}$$

Or $\int_{-\infty}^\infty \frac{Y^2}{2} \varphi(Y) dY = \frac{1}{2} \sigma^2$ d'où :

$$\begin{aligned} \int_0^\infty \frac{Y^2}{2} [\varphi(Y) - \varphi(-Y)] dY &= \frac{1}{2} \sigma^2 - \frac{2}{2} \int_{-\infty}^0 Y^2 \varphi(Y) dY \\ \text{Si on pose : } 2K &= 2 \int_{-\infty}^0 Y^2 \varphi(Y) dY \end{aligned}$$

Il vient : $A = -D(\sigma^2 - 2K) \partial_x P(x) dt$

Détermination de B :

En intégrant par parties :

$$\begin{aligned} \int_0^\infty 1 \frac{\Phi_+(Y) - \Phi_-(Y)}{\ell} dY &= \int_0^\infty \frac{\Phi_+(Y) - \Phi_-(Y)}{\ell} dY \\ &= \left[\frac{Y}{\ell} (\Phi_+(Y) - \Phi_-(Y)) \right]_0^\infty - \int_0^\infty \frac{Y}{\ell} [-\varphi(Y) - \varphi(-Y)] dY \end{aligned}$$

Comme précédemment le crochet est nul et si on procède au changement de variable $\frac{Y}{\ell} = z$ dans l'intégrale restante, il vient :

$$\begin{aligned} \int_0^\infty \frac{\Phi_+(Y) - \Phi_-(Y)}{\ell} dY &= \int_0^\infty \ell z [\varphi(\ell z) + \varphi(-\ell z)] dz \\ \int_0^\infty \frac{\Phi_+(Y) - \Phi_-(Y)}{\ell} dY &= \int_{-\infty}^\infty \ell z \varphi(\ell z) dz - 2 \int_{-\infty}^0 \ell z \varphi(\ell z) dz \end{aligned}$$

Ce qui représente une fraction de l'espérance de la variable ℓz , laquelle tend vers zéro avec ℓ .

Et enfin, $P(x, t)$ qui la probabilité pour qu'une trajectoire prenne la valeur x à l'instant t est une solution de l'équation différentielle partielle :

$$\partial_t P(x, t) = -\partial_x (V(P(x, t))) + D (\sigma^2 - 2K) \partial_{x^2}^2 (P(x, t)).$$

$$\text{Avec } 2K = 2 \int_{-\mu}^0 Y^2 \varphi(Y) dY$$

Valeur approchée de : $2K = 0.0445$

On remarque que cette équation est celle de Fokker-Planck adaptée à une variable aléatoire distribuée suivant une loi log-normale c'est-à-dire que c'est le logarithme de la variable qui suit une loi normale et non pas la variable elle-même. Un tel mouvement est dit "brownien géométrique". La différence entre les deux porte sur le terme multiplicatif ($\sigma^2 - 2K$).

Les temps d'attente sont distribués suivant une loi α - Stable

L'étude de la série des temps d'attente entre deux pluies faite précédemment a montré que celui-ci pouvait être approché par une distribution qui suivrait une loi de type α -stable avec ici $\alpha = 0.95$. α étant inférieur à 1, non seulement la variance est infinie mais la moyenne aussi. Ce cas a fait l'objet d'études spécifiques [3] retour qui conduisent à l'équation fractionnaire de Fokker-Planck notée F.F.P.E. (Fractionnaire Fokker-Planck Equation). Celle-ci, en tenant compte du paragraphe précédent où le mouvement brownien géométrique se substitue au mouvement brownien prend la forme :

$$\partial_t P(x, t) = {}_0D_t^{1-\alpha} \left[-\partial_x (V(P(x, t))) + D (\sigma^2 - 2K) \partial_{x^2}^2 (P(x, t)) \right].$$

Où ${}_0D_t^{1-\alpha}$ est la dérivée fractionnaire de Riemann-Liouville pour α compris entre 0 et 1, ce qui est le cas ici. Cette équation régit la dynamique d'un système, ici l'évolution des hauteurs de pluie cumulées, de type anormal, sous diffusif, qui suit une décroissance assez lente, en "puissance" et non "exponentielle".

Dans le cas qui nous occupe, $\alpha = 0.95$, l'équation ci-dessus qui régit le processus stochastique suivi par les hauteurs de pluie cumulées, n'a pas aujourd'hui de solution analytique. Un moyen d'approcher une solution est de simuler des trajectoires répondant au processus. Une telle approche sera faite au moyen de deux processus qui substituent un "temps opérationnel τ " au "temps physique t ". C'est le principe de subordination qui sera introduit dans le prochain chapitre.

2.3 Conclusion du 2nd chapitre

Les analyses menées au chapitre précédent ont permis de concevoir le cumul des hauteurs de pluie comme étant une marche aléatoire unidimensionnel non-gaussienne avec des sauts log-normaux et des temps d'attente α -stable ($\alpha = 0.95$). Dans le présent chapitre, les simulations de réalisation (trajectoire de marcheur) effectuées ont montré qu'on pouvait approximer de façon satisfaisante l'évolution des cumuls de pluie mesurés. Néanmoins, les résultats ont nécessité des adaptations de la marche aléatoire d'origine. En particulier, il nous a fallu soumettre les distributions des temps d'attente à trois conditions qui ont conduit à dénommer le processus correspondant : "processus tronqué". Ces conditions imposent aux tirages aléatoires des temps d'attente, premièrement de rester bornés par une distribution exponentielle pour éliminer les tirages excessivement grands sans rapport avec les observations réelles, deuxièmement de ne pas dépasser la durée des observations et enfin de présenter les mêmes proportions de temps humides et de sécheresse. Dans la deuxième partie du chapitre, nous présentons une démarche qui consiste à déduire du modèle de marche aléatoire un modèle macroscopique de la dynamique des cumuls de pluie. Ce dernier a été obtenu par un passage à la limite de la marche aléatoire. En passant à la limite lorsque le pas des sauts et le pas de temps tendent simultanément vers zéro en gardant un rapport constant, on montre que la probabilité d'atteindre une hauteur cumulée de pluie fixée pour un temps donné est régie par les solutions de l'Equation Fractionnaire de Fokker-Planck (F.F.P.E.). L'obtention de cette équation, qui n'admet pas de solution analytique connue permet cependant d'avoir directement par calcul numérique cette probabilité et par conséquent de définir la hauteur de pluie la plus probable en localisant au mieux le maximum de la loi de probabilité par exemple pour un temps donné. Cet aspect du modèle macroscopique dont la loi de probabilité suit la FFPE suggère un potentiel d'utilisation dans le cadre de modélisation opérationnel comme celui du projet "sol virtuel". Le prochain chapitre sera consacré à la résolution de l'équation FFPE sous deux approches différentes ainsi qu'à la mise en oeuvre de sa solution pour réaliser des simulations de la dynamique des cumuls de pluie. Des études comparatives entre les données observées et des simulations issues du modèle macroscopique régi par FFPE seront effectuées. En particulier, nous examinerons le phénomène du fléchissement de la courbe représentative des cumuls de pluie apparu aux environs de 1997 en scindant les données en deux périodes : "avant 1997" et "après 1997".

Chapitre 3

Mise en oeuvre du modèle régi par l'équation FFPE. Application à une simulation du fléchissement de la dynamique des pluies.

Le présent chapitre traite de la résolution du modèle régi par l'équation FFPE établi dans le chapitre précédent du manuscrit. Ce modèle est représenté par un processus stochastique de type diffusion dont nous avons montré auparavant que la loi de densité de probabilité est régi par l'équation fractionnaire de Fokker-Planck (FFPE). La donnée de cette loi de densité de probabilité permet d'accéder à toutes les caractéristiques statistiques du processus stochastique étudié. Un point particulièrement important dans cette résolution concerne la possibilité de simuler des réalisations du processus stochastique modèle et ainsi d'utiliser ces réalisations à l'instar des mesures expérimentales pour construire des statistiques et faire des prévisions. Pour cela, nous avons mis en oeuvre deux types de résolution de l'équation FFPE régissant la densité de probabilité du modèle. La première approche repose sur une démarche stochastique qui consiste à simuler les réalisations d'un processus d'Itô. Dans cette démarche nous exploitons la simulation directe du processus d'Itô pour aboutir à une solution de l'équation FFPE en s'appuyant sur les propriétés mathématiques de ce dernier. Compte tenu du caractère proche des observations expérimentales des réalisations que nous avons effectuées, une application pour simuler le fléchissement de la dynamique de la pluie sur la période en aval de l'année 1997 a été réalisé. En effet ce fléchissement pose des questions sur l'éventualité d'un changement définitif ou pas d'une dynamique. Plus précisément, le fléchissement du cumul des pluies qui apparaît vers 1997 est-il un élément annonciateur d'un éventuel changement de dynamique pérenne ? Pour cela, une utilisation plus fine de la résolution stochastique proposée ci-dessus sera mis en

œuvre comme première application du modèle. Une étude par morceau sera menée pour les périodes d'avant et puis d'après 1997. Une interprétation des résultats est ensuite menée pour caractériser ce fléchissement.

La deuxième méthode de résolution repose sur une discréétisation numérique de l'équation modèle. C'est une approche directe qui permet d'obtenir une forme de la loi de densité de probabilité des hauteurs de pluies. Nous montrons ensuite une comparaison des résultats obtenus par ces deux méthodes.

3.1 Approche stochastique de la mise en oeuvre du modèle

Comme on vient de le dire, le processus recherché, présumé régir le phénomène pluvieux étudié, possède une loi de densité de probabilité (pdf) relevant de l'équation fractionnaire de Fokker-Planck (FFPE). Notre démarche consiste à exploiter d'une part le fait qu'il y a un lien direct entre l'équation classique de Fokker-Planck et le processus stochastique décrit par l'équation différentielle d'Itô [44]. D'autre part, il est prouvé [93] que l'extension de l'équation de Fokker-Planck à la version fractionnaire FFPE correspond à la subordination du processus décrit auparavant par l'équation différentielle d'Itô. En simulant ce processus stochastique subordonné par une méthode de Monte-Carlo appliquée à l'équation d'Itô, nous obtenons des réalisations qui représentent les trajectoires. Nous pouvons ainsi accéder par reconstruction à partir de ces trajectoires à des histogrammes approchant la solution de l'équation FFPE.

Définition

Soit $X(t)_{t>0}$ un processus stochastique à valeur réelle (Cf. Annexe D). Soit $S(t)_{t>0}$ un processus stochastique croissant à valeur positive. Alors le processus $Y(t) = X(S(t))$ est appelé processus subordonné de X et $S(t)$ est appelé processus directeur. Le processus $S(t)$ est mathématiquement défini par

$$S(t) = \inf\{\tau, U(\tau) > t\} \quad (3.1)$$

où $U(\tau)$ représente un processus de Lévy α -stable strictement croissant à valeurs positives ($\alpha \in (0, 1)$).

Le principe de subordination d'un processus stochastique

La subordination entre processus s'écrit $Y(t) = X(S_t)$ où $X(\tau)$ est un processus solution de

l'équation différentielle stochastique de Îto ci-dessous [44] :

$$dX(\tau) = V X(\tau) + \sqrt{2(\sigma^2 - 2K)D} dB(\tau) \quad (3.2)$$

avec $B(\tau)$ un mouvement brownien standard.

Simulations de trajectoires du processus subordonné

L'algorithme utilisé a pour objectif de simuler le processus subordonné $Y(t) = X(S_t)$ sur un maillage discret $\{t_i = i \delta t : i = 1, 2, \dots, N\}$ avec $\delta t = \frac{T}{N}$ où T représente la durée de la simulation.

La première étape de la simulation consiste à trouver la valeur du processus directeur $S(t)$. Pour cela nous simulons le processus $U(\tau)$ avec un maillage plus $\{\tau_j = j \delta \tau : j = 1, 2, \dots, M\}$ avec $\delta \tau < \delta t$ sous la forme

$$U(\tau_j) = U(\tau_{j-1}) + \delta \tau^{\frac{1}{\alpha}} \zeta_j \quad (3.3)$$

avec ζ_j une variable aléatoire de loi α -stable. L'estimation de $S(t_i)$ repose ensuite sur la définition (3.1) qui revient à sélectionner pour chaque temps t_i , l'instant τ_j tel que $U(\tau_{j-1}) < t_i < U(\tau_j)$. Ce qui permet de trouver la valeur $S(t_i) = \tau_j$. Nous calculons ensuite les solutions de l'équation d'Îto (3.2) approximée par un schéma d'Euler. Le maillage sur lequel portera cette discrétisation est $\delta \bar{\tau}$ lui aussi inférieur à δt et avec un nombre de points L , suffisant pour couvrir l'ensemble des S_t . Pratiquement, nous prenons L comme le premier entier dépassant la quantité $\frac{S(t_N)}{\delta \bar{\tau}}$. La discrétisation, partant de 0, prend alors la forme :

$$X(\bar{\tau}_k) = X(\bar{\tau}_{k-1}) - V \Delta \bar{\tau} + \sqrt{2(\sigma^2 - 2K)D} \Delta \bar{\tau}^{0.5} \bar{\xi}_k \quad (3.4)$$

k variant de 1 à L et $\bar{\xi}_k \sim N(0, 1)$. La dernière étape de la simulation consiste à interpoler les valeurs du processus subordonné $X(S(t_i))$ avec la formule

$$X(S(t_i)) = \frac{X(\bar{\tau}_{k+1}) - X(\bar{\tau}_k)}{\bar{\tau}_{k+1} - \bar{\tau}_k} (S(t_i) - \bar{\tau}_k) + X(\bar{\tau}_k) \quad (3.5)$$

Simulations de trajectoires solutions de F.F.P.E. "Subordonnées et tempérées"

La figure (3.1) illustre la comparaison des trajectoires avec la série expérimentale. Comme pour la marche aléatoire, le fléchissement qui s'opère à partir de 1997 soit le 10 000e jour environ, peut être provoqué, par l'introduction aux côtés de la subordination, de temps d'attente tempérés et le resserrement des trajectoires autour de la série des observations obtenu par la "sélection". Sans épouser parfaitement la courbe empirique, la trajectoire moyenne représentée

la suit assez bien. A l'instar de la marche aléatoire, la série expérimentale voisine bien avec la trajectoire moyenne des simulations et se situe sur en majeur partie entre les 5ième et 8ième déciles alors que la trajectoire moyenne se situerait, elle, entre les 3ième et 5ième déciles. Ceci traduit un fléchissement insuffisamment marqué, par rapport aux observations. On remarque, en outre, que lorsque l'on sélectionne les trajectoires en fonction de la proportion de temps d'attente comprise entre 70 et 90%, les déciles qui " encadrent " la courbe des hauteurs de pluie cumulée sont trois fois moins distants les uns des autres que les déciles dans le cas des trajectoires seulement subordonnées et tempérées.

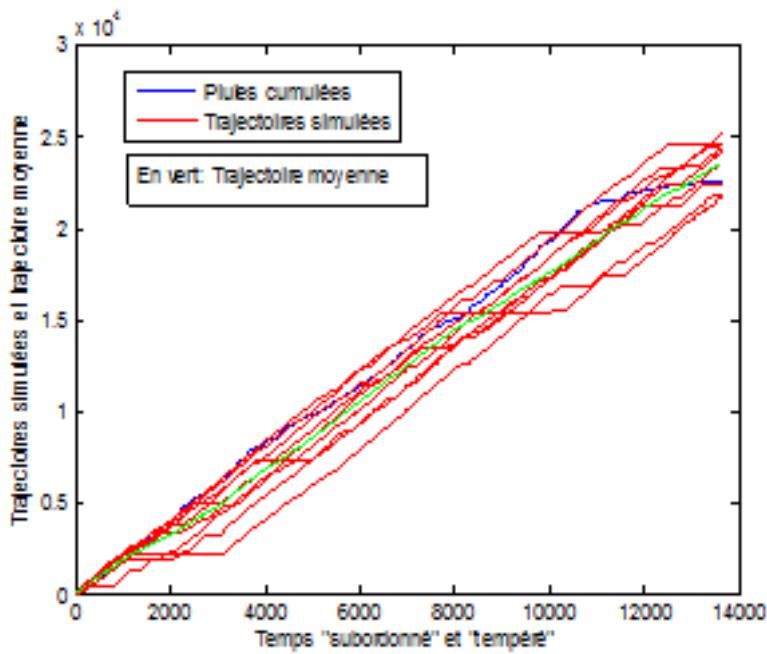


FIGURE 3.1 – Trajectoires simulées et trajectoire moyenne obtenues à partir d'une discréétisation de l'équation différentielle stochastique d'Itô. Les temps sont "subordonnés, tempérés, sélectionnés". En vert : la trajectoire moyenne. Paramètres : $\alpha = 0.95$, $V = \mu/\tau$, $\eta = 0.5396$, $D = 1$, $\lambda = 10^{-7}$.

P.D.F. des hauteurs de pluie

L'aboutissement de cette approche du processus recherché est, bien sûr, d'évaluer la P.D.F. des hauteurs de pluie pour un temps donné et d'estimer les cumuls de pluie dans les années à venir. Les trajectoires solutions de F.F.P.E., obtenues précédemment conduisent naturellement à ces P.D.F. par la méthode de Monte Carlo. Etant donné un certain temps t , on obtient la valeur atteinte par une trajectoire par interpolation entre les hauteurs de pluie atteintes avant cette valeur de t puis celle d'après. Il suffit donc par la suite de construire l'histogramme de toutes ces valeurs atteintes par toutes les trajectoires. L'allure de l'histogramme étant fonction du nombre de " boites " ou " bins ", il convient, comme pour le processus précédent des marches

aléatoires, de multiplier le choix de ces nombres de boîtes. La hauteur de pluie la plus probable sera alors estimée par simple lecture de l'histogramme dont un exemple est représenté sur la figure (3.2).

Les tableaux (3.1) et (3.2) ci-après résument les hauteurs de pluies les plus probables à diverses dates tant suivant un modèle de marche aléatoire que le modèle F.F.P.E. L'analyse des résultats montre que la marche aléatoire comme les solutions F.F.P.E approchent d'autant mieux les valeurs observées que le nombre d'observations est important. Ceci est dû au fait que tempérer les temps d'attente resserre plus les parties supérieures des trajectoires puisque limités par une distribution exponentielle. On note, de plus, l'amélioration des résultats due à la sélection des trajectoires par rapport aux distributions tempérées seulement. La figure (3.3) met en regard de la courbe expérimentale les hauteurs de pluie cumulées les plus probables, d'une part obtenues par une marche aléatoire "tempérée, triée, sélectionnée" et d'autre part, découlant de F.F.P.E. On remarque une très grande similitude entre les modèles, lesquels n'épousent pas parfaitement la courbe expérimentale. La différence entre les vitesses intervenant dans les deux modèles, marche aléatoire et équation différentielle stochastique d'Ito s'explique par le fait que l'advection $V = \mu/\tau$ s'exprime en fonction du pas de temps. Le changement de processus induit par la subordination induit un changement de pas de temps, donc une différence entre les vitesses. En effet, au "temps réel" en jours se substitue un temps "opérationnel" constitué du terme de la distribution α -stable directement supérieure au temps réel. Les pas de temps opérationnels étant moins nombreux que les réels, le saut moyen doit être, alors, plus important donc la vitesse plus grande.

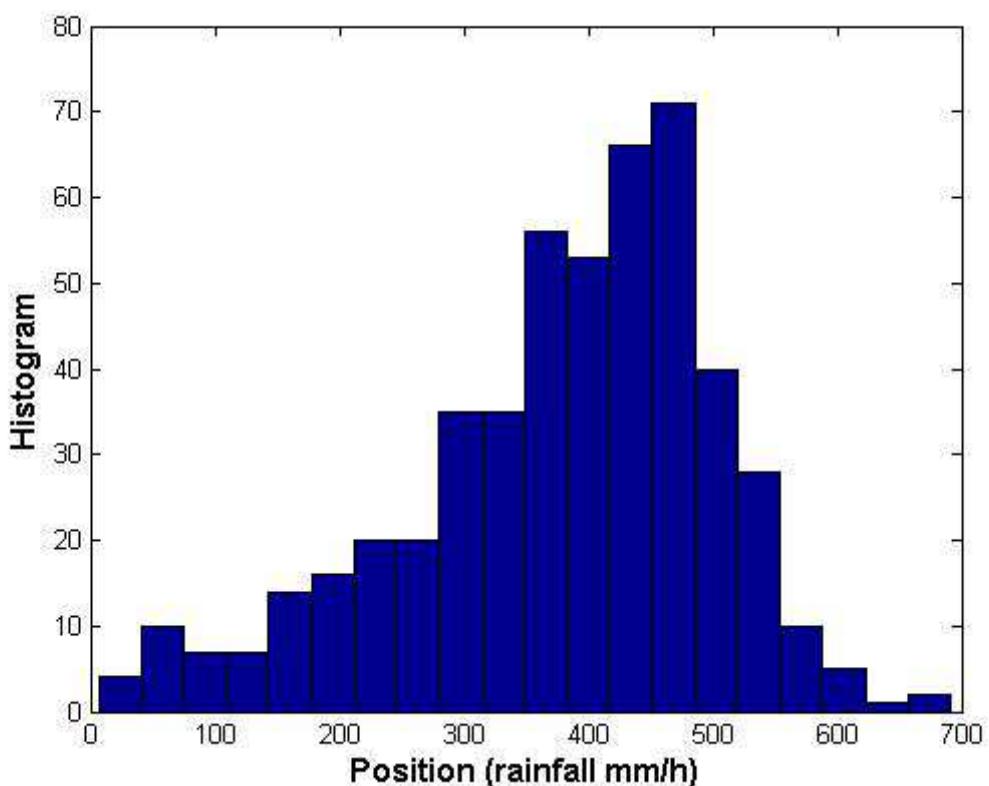


FIGURE 3.2 – Exemple de graphe permettant d’apprécier la hauteur de pluie cumulée la plus probable pour un temps donné, ici pour le 300^{ième} jour.

Num. du jour	100	500	1000	5000
Nb. de réalisations	121	950	1 707	9742
Marche aléatoire	140	960	1 800	9020
Ecarts	+ 15%	+1%	+5%	-7%
F.F.P.E.	290	1 310	2 670	8990
Ecarts	+139%	+37%	+57%	-8%

TABLE 3.1 – Comparaison des hauteurs de pluie les plus probables (en mm) obtenues avec la marche aléatoire et l'équation différentielle stochastique d'Itô. Les écarts par rapport aux valeurs mesurées sont en %

Num. du jour :	8000	10640	11600	12600	13 659
Nb. de réalisations	15000	20 918	21678	22283	22 593
Marche aléatoire	15000	18 700	20440	21100	22 650
Ecarts	0%	-10%	-5%	-5%	+0.3%
F.F.P.E.	15390	19 480	20560	22300	22 450
Ecarts	2.6%	8%	-5%	+1%	-2%

TABLE 3.2 – Comparaison des hauteurs de pluie les plus probables (en mm) obtenues avec la marche aléatoire et l'équation différentielle stochastique d'Itô. Les écarts par rapport aux valeurs mesurées sont en %

3.2 Apparaît-il un aspect d'un changement climatique en 1997 ?

La série des pluies cumulées étudiée au 1er chapitre, présente deux comportements graphiques bien distincts. Ceci correspond à deux portions de courbes notées naturellement "cumuls avant 1997" et "cumuls après 1997". La première, de 1968 à 1997 présente sur 31 ans une évolution à tendance linéaire suivie d'un "fléchissement" dû à des sécheresses plus accentuées de 1997 à 2008. Ce fléchissement est un changement de la dynamique des précipitations locales ce qui n'induit pas forcément un changement de climat car ce dernier dépend de bien d'autres paramètres mais peut en être une composante. En conséquence, une analyse séparée des données ces deux parties a été calquée sur la précédente étude. Parmi les divers paramètres qui ont été adaptés aux nouvelles données, l'index de stabilité est proposé ci-dessous au tableau (3.3), car on remarque que si l'index de la partie "cumuls avant 1997" est comparable à celui du temps d'attente dans son ensemble, celui du "cumuls après 1997" est très nettement inférieur donc révèle une "aile plus épaisse" donc des évènements extrêmes plus importants. Cela se vérifie bien ici puisque c'est après 1997 que les périodes de sécheresse ont augmenté. Cependant,

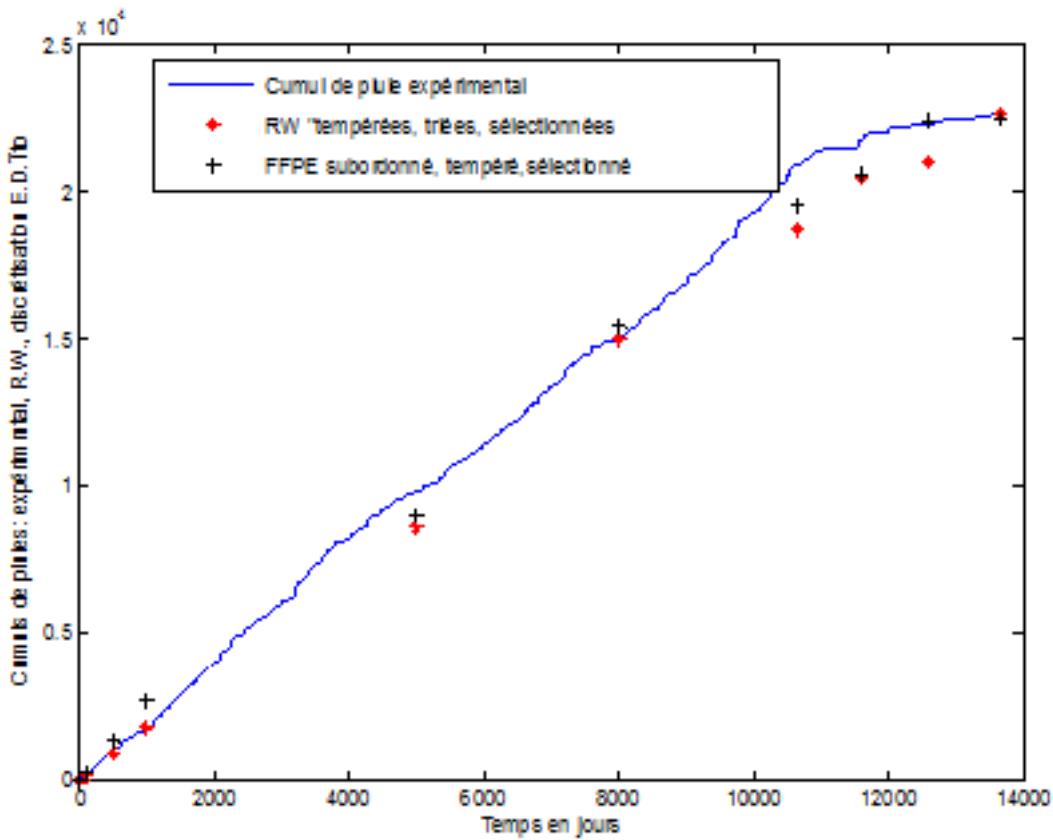


FIGURE 3.3 – Comparaison des hauteurs de pluie les plus probables par le "processus tronqué" : (*) par des simulations de marches aléatoires. Paramètres : $V = 3.4$; $D = 1$; $\alpha = 0.9$. (+) par discrétisation de l'E.D.P.Stochastique de Ito. Paramètres : $V = 5$; $D = 1$; $\alpha = 0.9$.

les deux périodes concernées ayant moins d'observations que la période entière, l'estimation est moins fiable.

Index de stabilité	cumuls avant 1997	cumuls après 1997
McCULLOCH	$\alpha = 0.91$	$\alpha = 0.62$
Koutrouvelis	$\alpha = 0.95$	$\alpha = 0.6$
Régression des queues	$\alpha = 0.92$	$\alpha = 0.6$

TABLE 3.3 – Estimation des exposants de stabilité de la loi stable des temps d'attente avant et après 1997.

Les résultats de cette étude séparée des données avant et après le fléchissement sont consignés dans les tableaux (3.4) et (3.5). Pour une modélisation par marche aléatoire, les hauteurs de pluies cumulées les plus probables, issues de la juxtaposition des deux processus correspondants est représentée en regard de la courbe expérimentale en figure (3.4). On remarque un assez bon accord entre les hauteurs cumulées observées et celles obtenues par juxtaposition des modèles. On note que les résultats obtenus en modélisant d'un seul trait la période de 1968-2008 apparaît

moins pertinent pour reproduire le fléchissement. Des résultats similaires s'obtiennent avec des simulations du processus d'Itô.

Marches aléatoires	α	D	V	λ	μ	σ
1968-2008	0.9	1	3.4	10^{-7}	1.3944	1.149
1968-1997	0.92	0.01	2.8	10^{-7}	1.3769	1.1541
1997-2008	0.6	0.03	4.4	10^{-7}	1.6288	1.077

TABLE 3.4 – Paramètres des simulations par marches aléatoires.

Equation d'Itô	α	D	V	λ	μ	σ	2K
1968-2008	0.9	1	5	10^{-7}	1.3944	1.149	0.0445
1968-1997	0.92	0.01	2.9	10^{-7}	1.3769	1.1541	0.0439
1997-2008	0.6	0.03	4.5	10^{-7}	1.6288	1.077	0.0497

TABLE 3.5 – Paramètres des simulations avec l'équation différentielle d'Itô.

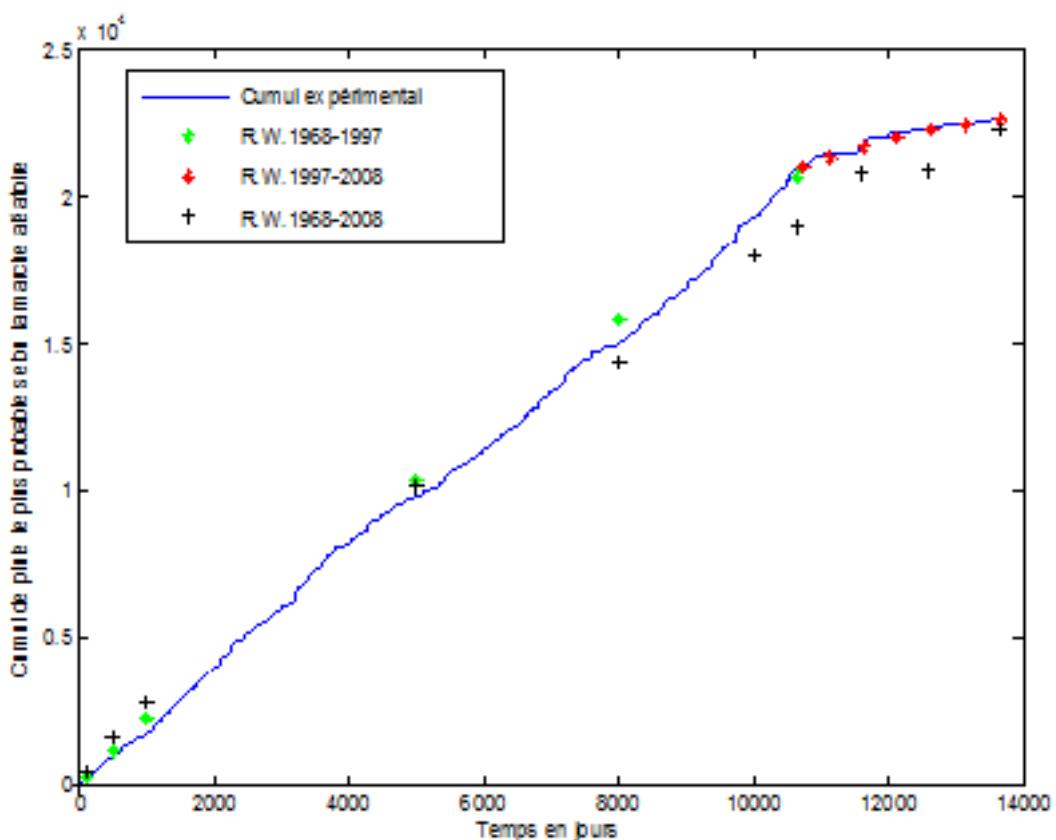


FIGURE 3.4 – Comparaisons des hauteurs de pluies cumulées les plus probables avec les observations. Les paramètres des simulations sont donnés dans le tableau (3.4)

3.3 Résolution numérique du modèle macroscopique

Les sections précédentes ont permis d'estimer au mieux les paramètres du processus. Ils peuvent être alors utilisés dans une discréttisation de F.F.P.E. réalisée en C++ sur un P.C. acceptant un nombre suffisamment grand de simulations donc en affinant les résultats. La discréttisation de la dérivée fractionnaire de Riemann-Liouville a été réalisée de manière implicite en temps et en utilisant la méthode de Grünwald-Letnikov en x. La figure (3.5) illustre les hauteurs de pluie cumulées les plus probables obtenues avec la solution numérique de FFPE. On remarque que la modélisation ne reproduit pas le fléchissement.

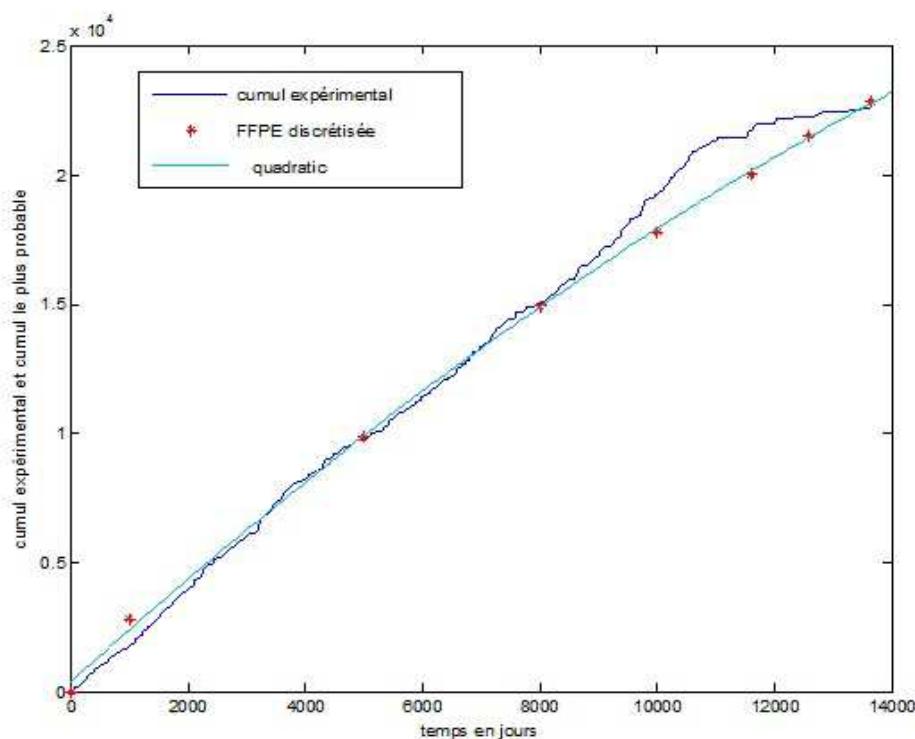


FIGURE 3.5 – Hauteurs de pluie cumulées les plus probables obtenues avec la résolution numérique de F.F.P.E. (Période 1968-2008)

Prévision de valeur de cumul de pluie en 2018

A titre de curiosité, nous avons réalisé une extrapolation de ces processus à 2018. Le tableau (3.6) ci-dessous compare les prévisions statistiques des cumuls de pluie la 50e année : Il apparaît un différentiel de 5 mètres et une moyenne de 24,6 mètres.

Processus tronqué	Marche aléatoire F.F.P.E.	25 mètres 26 mètres
Processus juxtaposés	Marche aléatoire F.F.P.E.	25 mètres 23 mètres
Processus global	Marche aléatoire F.F.P.E.	22 mètres 27 mètres

TABLE 3.6 – Tableau de prévision de la hauteur cumulée la plus probable en 2018

3.4 Conclusion du 3^{ème} chapitre

Dans ce chapitre trois, nous avons abordé avec deux démarches différentes la résolution de l'équation fractionnaire de Fokker-Planck (F.F.P.E.). Les paramètres de la F.F.P.E. sont donnés par les résultats de l'analyse des données réalisée au premier chapitre. La première démarche emprunte l'idée que la F.F.P.E. fournit aussi la fonction densité de probabilité d'un processus stochastique subordonné régi par l'équation différentielle stochastique d'Ito. Cette équation différentielle peut être aisément approximée par une discréétisation et un algorithme de simulation des trajectoires a été mis en oeuvre. Comme précédemment nous sommes restés dans le cadre d'un "processus tronqué" où les temps d'attente sont soumis aux mêmes restrictions que pour les marches aléatoires décrites au chapitre précédent. Les simulations obtenues permettent d'établir les hauteurs de pluie les plus probables pour un temps donné, lesquelles ont été comparées à celles obtenues par le modèle de marche aléatoire [22]. Il apparaît que les résultats fournissent des ordres de grandeurs tout à fait comparables aux observations. Afin d'éclairer le phénomène du changement de dynamique des précipitations cumulées apparu aux environs de 1997, nous avons scindé les données en deux "morceaux", avant et après 1997 et réalisé les mêmes démarches que précédemment sur chacun des morceaux. Il s'avère que dans ce cas, les hauteurs de pluie les plus probables suivent quasi parfaitement les valeurs expérimentales. Ce fait semble suggérer qu'il y a eu réellement un changement de dynamique dans les données à partir de 1997. Néanmoins, les raisons de ce changement restent une question ouverte.

La seconde partie du chapitre est consacrée à la résolution par une discréétisation numérique de la F.F.P.E. A la différence de la première démarche où nous avons simulé un "processus tronqué", aucune adaptation pour les temps d'attente ne pouvait être introduite car la solution numérique ne permet pas d'identifier facilement les paramètres. Il s'agit ici donc de travailler un "processus global".

En conclusion des différentes simulations produites, nous avons montré que le "processus global" donne un résultat légèrement plus linéaire et avec un fléchissement moins marqué que celui du "processus tronqué" qui lui-même est moins marqué aussi que celui constaté sur la courbe expérimentale. Les deux méthodes pourront cependant servir d'outils pour simuler les phénomènes de pluies dans le projet "sol virtuel". Ils serviront de paramètres d'entrées pour modéliser dans un cadre plus réaliste, la contribution du phénomène pluvieux dans le transfert

de masses vers la nappe phréatique.

Discussion sur la 1^{ère} partie

Cette première partie de notre travail a pour objectif de proposer une modélisation du phénomène pluvieux local en se basant sur les caractéristiques des observations effectuées localement sur une durée d'une quarantaine d'années. Compte tenu de cette ambition, le travail a nécessité la mise en oeuvre de diverses techniques appartenant à plusieurs domaines de recherche. Plus explicitement, cette première partie a nécessité la maîtrise de plusieurs méthodes de traitements de données combinées à la formulation et à la résolution d'une approche probabiliste. Cette approche stochastique a été justifiée et construite à partir de bases obtenues en analysant les observations. Conformément à l'objectif rappelé ci-dessus, le travail s'est déroulé en trois étapes à savoir l'étude des données mesurées, la formulation du modèle et enfin sa mise en oeuvre. Cet aspect articulé autour des données observées présente l'intérêt de la prise en compte des caractéristiques locales de la région et offre la possibilité de développer un outil plus réaliste. En contrepartie de cela, des adaptations "manuelles", pas facile à formaliser mathématiquement ont été introduites dans nos modèles afin d'approcher au mieux la tendance des observations qui rappelons le, constituent la seule vraie réalisation du processus recherché.

L'autre pilier de notre travail est l'approche stochastique du phénomène de pluie local. Ce choix a fait l'objet d'une recherche extensive depuis plusieurs décennies. Rappelons que la majorité partie des modèles de phénomène pluvieux publiés dans la littérature adopte cette approche stochastique en faisant appel à une modélisation par des processus de cascade multi-fractal introduits dans les années 70 par [30] [67] [50] et généralisé ensuite par [75]. Le succès de ces modèles de cascade repose sur leur facilité de mise en oeuvre et sur la théorie multi-fractale qui rappelons le implique l'observation de loi d'échelle dans les statistiques du phénomène étudié. Notons que cette approche par cascade a été remise en cause par plusieurs travaux publiés depuis une vingtaine d'année qui montrent que la dynamique observée dans un phénomène pluvieux ne suit pas toujours une loi d'échelle [65][11][66][51].

D'autre part, les progrès sur les techniques d'analyse de dynamique chaotique dans les années 80 ont permis le développement de certains travaux qui tendent à associer la dynamique des phénomènes pluvieux à des chaos déterministes de faibles dimensions, ce qui remet en cause le choix même de l'approche stochastique. Notons que Schertzer et al [76] ont étudié cette question et ont montré qu'en particulier l'obtention de faibles de dimension de corrélation associée aux

données de pluies ne peut pas constituer une preuve irréfutable de leur nature déterministe.

Notons aussi qu'une catégorie de modèles associés à des processus de diffusion existe. Nous citons en particulier, le modèle diffusif markovien de Veneziano et al [88] associé à un mouvement brownien géométrique et le modèle de diffusion fractionnaire de Kundu et Bell [34] et Kundu et Travis [35]. Rappelons enfin que notre modélisation fait appel à l'emploi d'un mouvement brownien géométrique subordonné dont la loi de probabilité est régie par l'équation fractionnaire de Fokker-Planck. Il se situe donc plus naturellement dans cette dernière catégorie.

Seconde partie

Estimation de l'impact des précipitations pluvieuses sur les fluctuations de la nappe phréatique.

Chapitre 4

Estimation de l'impact des précipitations pluvieuses sur les fluctuations de la nappe phréatique.

Dans cette partie du travail, nous nous attachons à analyser et à estimer l'impact des précipitations pluvieuses sur les fluctuations de la nappe phréatique. En accord avec l'idée principale de cette thèse, un des objectifs principaux de cette deuxième partie est de montrer sur la base d'observation expérimentale le transfert de l'eau pluvieuse au sein de la couche de sol vers la nappe. Les données expérimentales utilisées sont formées de mesures synchronisées de la variation du niveau de la nappe et des précipitations. A la différence de la mesure de pluie journalière analysée au premier chapitre, les données étudiées ici sont de type horaire. Elles ont été acquises sur le site atelier "Fontanille" de Centre de Recherche de l'INRA d'Avignon.

Mesurer cet impact des précipitations à partir des mesures du niveau de la nappe n'est pas très aisée car la nappe phréatique d'un bassin fait partie d'un système hydrologique. Elle est donc en lien avec différents processus physiques qui contribuent à ses variations en particulier nous remarquons que l'apport vertical d'eau vers la nappe est dû principalement aux précipitations car il n'y a pas de phénomène de ruissellement observé sur le site. Des apports latéraux liés aux rivières et au système d'irrigation contribuent aussi à la variation du niveau de la nappe. Il est important de noter ici que ces différentes contributions correspondent à des échelles de temps très différents. Ce constat est à la base de notre démarche car il permet de sélectionner les variations du niveau de la nappe correspondant à l'échelle de temps des précipitations. D'autrepart, il faut aussi prendre en compte les différents processus qui agissent sur la précipitation avant qu'elle n'atteigne la nappe. En particulier, le couvert végétal et le phénomène d'évapotranspiration à la surface doivent être pris en compte dans l'analyse.

4.1 Estimation des fluctuations de nappe corrélées avec les événements de pluie

Les données expérimentales

Les variations de niveau de la nappe ont été mesurées et enregistrées en trois points du site atelier (voir Annexe E) à l'aide de capteurs piézométriques. Chaque enregistrement correspond à une série temporelle horaire de 10.000 valeurs allant de la date du 02 Février 2005 au 27 Mars 2006. On note que les valeurs de la hauteur de nappe sont données avec la norme NGF qui a pour référence le niveau de la mer. Un bref résumé des statistiques obtenues avec les points de mesure est donné sur le tableau (4.1).

Piezometers	Mean value (m)	Std (m)
1C	26.552	0.258
2L	26.625	0.265
3C	26.256	0.244

TABLE 4.1 – Statistiques des niveaux de la nappe en trois points différents sur le site

Le tableau (4.1) montre que les trois séries temporelles présentent des caractéristiques assez similaires. Ceci suggère qu'il n'y a pas de variabilité spatiale importante sur le site atelier. Compte tenu des similarités obtenus sur les histogrammes des trois séries temporelles, par la suite, nous présenterons uniquement les résultats obtenus avec la série "3C" (Cf. Annexe E).

Pour la même période que celle citée plus haut et simultanément avec les mesures de données de niveau de la nappe, un pluviomètre à auget basculeur mesure les niveaux de précipitations à chaque heure. Des descriptions détaillées sur la précision des mesures de précipitations pluvieuses sont données par la note technique [38]. Le tableau (4.2) donne une répartition de l'échelle des événements de pluie recueillis. On remarque que presque 80% des événements de pluie durent moins de trois heures. On retrouve ainsi les caractéristiques du climat méditerranéen avec des précipitations inégalement réparties sur l'année, maximales en avril - mai et septembre - octobre.

La figure (4.1) montre une représentation en parallèle des variations de hauteur de nappe phréatique et de précipitations.

Filtrage et obtention des fluctuations de nappe corrélées avec les événements de pluie

On note sur les variations du niveau de la nappe de la figure (4.1) une "tendance générale à

Durée	Nombre	Proportion (%)
1-3 heures	92	79.31
4-6 heures	18	15.52
7 heures et plus	6	5.17

TABLE 4.2 – Statistiques des événements de pluies.

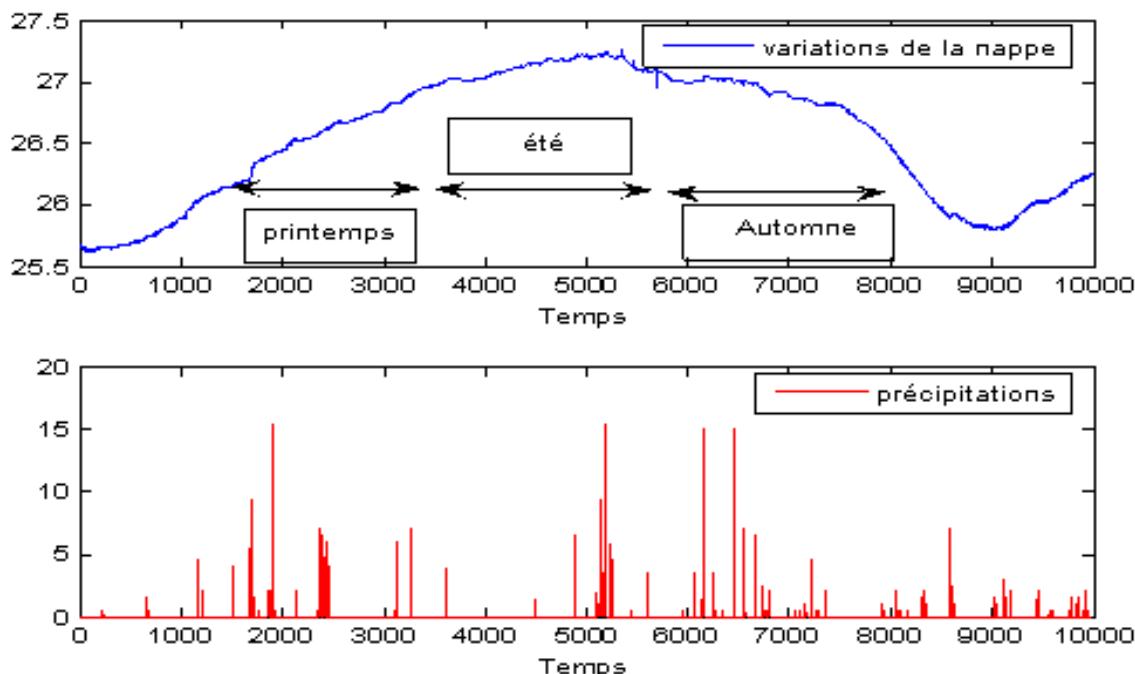


FIGURE 4.1 – Relevés horaires des variations de hauteur de nappe phréatique et de précipitations.

long terme” due à des phénomènes saisonniers tels que l’irrigation gravitaire, cours d’eau... et des fluctuations nombreuses sur des temps courts, de quelques heures qui coïncident parfois à des événements de pluie assez forts. En effet, on observe à la fois une évolution lente du niveau et des mouvements rapides de la surface piézométrique. Ceci signifie que le signal qui décrit les hauteurs de nappe présente de 2 composantes qui s’additionnent une variation saisonnière, qui correspond à une dynamique de grande échelle et des fluctuations, qui peuvent traduire une réponse locale de la nappe à une pluie. Il est ainsi évident que pour la caractérisation des transferts d’eau à l’échelle locale (c’est à dire à l’échelle de notre site atelier), ce sont les fluctuations piézométriques statistiquement correlées avec les événements de pluie qui peuvent nous intéresser. Ceci va nous amener à mettre en oeuvre une démarche de filtrage qui aura pour objectif d’éliminer les composantes du signal qui n’ont pas de lien avec les pluies.

Cohérence entre les hauteurs de nappe et les événements de pluie

Cette démarche de filtrage nécessite donc l’étude de la cohérence entre le signal de hauteur

de nappe et le signal de précipitations de pluies. Afin de détecter l'impact que pourrait avoir la pluie sur le niveau de la nappe phréatique, notre approche consiste à apprécier le degré de dépendance des deux séries par le calcul de leur corrélation croisée, il importe de préciser cette dépendance en spécifiant mieux la composante de la variation de la hauteur de la nappe qui réagit à l'apport de pluie. Les notations adoptées ici pour les précipitations et hauteurs de nappe en fonction du temps seront respectivement $\lambda(t)$ et $\eta(t)$.

Rappels sur les transformées de Fourier

Afin définir la fonction de cohérence entre deux signaux, nous allons introduire l'opérateur mathématique transformée de Fourier. De manière simple et sans souci de formalisme mathématique, nous décrivons la transformée de Fourier comme un opérateur qui permet de représenter une fonction du temps en une somme de combinaison d'amplitudes et de phases sous la forme

$$\hat{\eta}(f) = \int_{-\infty}^{\infty} \eta(t) e^{-2\pi i t f} dt \quad (4.1)$$

où $\eta(t)$ représente la fonction temporelle initiale et $\hat{\eta}(f)$ sa fonction transformée de Fourier. En d'autre termes, l'équation (4.1) permet de faire une décomposition harmonique de notre série temporelle. Remarquant que les données d'observation sont de nature discrète $\{\eta(t_1), \eta(t_2), \dots, \eta(t_i), \dots, \eta(t_n)\}$ alors il est plus convenable de remplacer la forme intégrale (4.1) par sa version discrète

$$\hat{\eta}(f_k) = \sum_j \eta(t_j) e^{-2\pi i t_j f_k} dt \quad (4.2)$$

En général, $\hat{\eta}(f)$ est une fonction complexe de forme $a(f) + ib(f)$. On peut écrire $\hat{\eta}(f)$ sous sa forme polaire :

$$\hat{\eta}(f_k) = |\hat{\eta}(f_k)| e^{i\Phi(f_k)} \quad (4.3)$$

où $|\hat{\eta}(f_k)| = \sqrt{a^2 + b^2}$ représente l'amplitude de Fourier à la fréquence f_k et $\Phi(f_k) = \tan^{-1}(\frac{b}{a})$ représente la phase. Sous l'hypothèse de stationnarité au second ordre, le théorème de Wiener-Kinchin [4] dit que la densité spectrale d'énergie de $\eta(t)$ est donnée par la transformée de Fourier inverse de la fonction d'autocorrélation $\rho_\eta(\tau)$

$$S_\eta(f) = \int_{-\infty}^{\infty} \rho_\eta(\tau) e^{-i2\pi f\tau} d\tau \quad (4.4)$$

La densité spectrale (4.4) donne la répartition en fréquence de l'énergie de $\eta(t)$. La forme

discrète de l'expression (4.4) s'écrit

$$S_\eta(f) = \sum_k C_{\eta k} e^{-i2\pi kf} \quad (4.5)$$

Par l'équation (4.3), il est simple de voir que $S_\eta(f) = |\hat{\eta}(f_k)|^2$. Ce qui montre que la phase du signal disparaît dans l'expression de la densité spectrale.

Définition de la fonction de cohérence

Désignons par $\lambda(t)$ et $\eta(t)$ deux séries temporelles qui représentent dans notre cas les mesures précipitations et les variations de hauteur de nappe. Nous appelerons fonction de corrélation croisée entre ces deux séries temporelle, la quantité

$$C_{\lambda\eta}(\tau) = \lim_{T \rightarrow +\infty} \frac{1}{2T} \int_{-T}^T \lambda(t + \tau) \eta(t) dt \quad (4.6)$$

qui généralise l'expression de la fonction d'autocorrélation définie dans le premier chapitre.

En appliquant la transformation de Fourier à l'expression (4.6), il est possible de montrer que la fonction de corrélation croisée $C_{\lambda\eta}(\tau)$ dépend du déphasage entre $\lambda(t)$ et $\eta(t)$, et de leurs amplitudes de Fourier. Dans l'espace de Fourier, il est plus convenable de décrire la fonction de corrélation croisée par une fonction cohérence normalisée entre 0 et 1, et un déphasage.

$$R_{\lambda\eta} = \frac{|S_{\lambda\eta}(f)|^2}{S_\lambda(f)S_\eta(f)} \quad (4.7)$$

$$\text{Ph}_{\lambda\eta}(f) = \arctan \left(\frac{-\text{Im}(S_{\lambda\eta}(f))}{\text{Re}(S_{\lambda\eta}(f))} \right) \quad (4.8)$$

avec $S_{\lambda\eta}(f)$ transformée de Fourier de $C_{\lambda\eta}(\tau)$, $S_\eta(f)$ et $S_\lambda(f)$ sont respectivement les transformées de Fourier de la fonction d'autocorrelation de chaque série temporelle suivant la définition (4.4). Ainsi, l'idée qu'une partie du signal de variation de hauteur de nappe soit corrélée avec la série temporelle de pluies constitue une hypothèse que nous allons maintenant vérifier. En d'autres termes, si notre hypothèse est significative alors il existe dans la transformée de Fourier du signal de nappe des amplitudes et des relations entre les phases qui permettent de produire la cohérence avec le signal de précipitations. Pour tester cette hypothèse nous allons construire un signal surrogat de $\eta(t)$. Rappelons que le signal surrogat est construit en remplaçant la phase $\Phi(f)$ de $\hat{\eta}(f)$ par une phase aléatoire uniformément distribuée dans l'expression (4.3). Ce qui a pour effet de détruire les relations de phases existantes. Le test de notre hypothèse de cohérence revient donc d'une part à estimer la cohérence entre $\lambda(t)$ et $\eta(t)$ d'autre part à

estimer aussi la cohérence entre $\lambda(t)$ et le signal surrogate de $\eta(t)$ [84]. La figure (4.2) montre les estimations des fonctions de cohérence et du déphasage entre $\lambda(t)$ et $\eta(t)$, et entre $\lambda(t)$ et le surrogate de $\eta(t)$. On note deux points intéressants dans les résultats que nous montrons ici : le premier point concerne le contraste très nette entre la cohérence obtenue en utilisant avec le signal de nappe et celle obtenue en le remplaçant par son surrogate. Plus précisément, les valeurs de cohérence obtenues avec le signal surrogate sont très faibles tandis que la cohérence avec le signal réel atteint des valeurs allant de 60 à 70 % sur un domaine de fréquence $f_l = 1e^{-4}Hz$ à $f_h = 5e^{-4}Hz$. Le deuxième point concerne l'évolution de la phase sur le même domaine de fréquence. Il apparaît que la phase obtenue avec le surrogate présente un comportement très désordonné à l'image de la phase aléatoire du surrogate lui-même tandis qu'une forme de déphasage plus structurée est obtenu avec le signal réel.

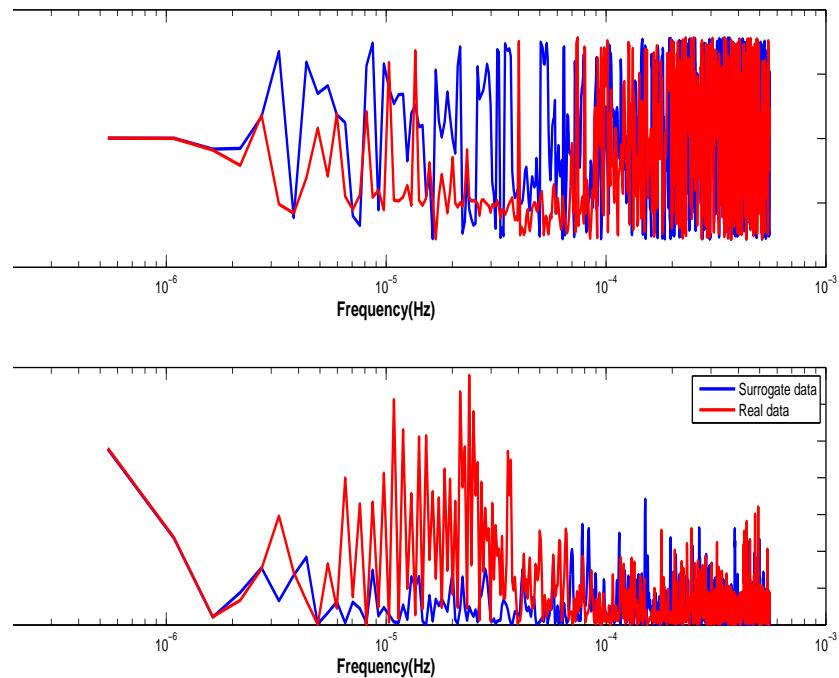


FIGURE 4.2 – Fonction de cohérence normalisée et déphasage entre λ , η et entre λ et le surrogate de η .

A des fréquences en dessous de f_l , on remarque que les résultats des signaux réel et surrogate présentent des comportements assez similaires. Ceci suggère directement l'idée que les basses fréquences qui correspondent à l'échelle des variations saisonnières ne sont pas corrélées avec le signal de pluie $\lambda(t)$. Les mêmes conclusions sont valables pour le domaine des fréquences supérieures à f_h . Cependant l'interprétation des résultats est moins aisée car en principe, on

doit tenir compte des erreurs induites par la limitation liée à la fréquence d'échantillonnage. Ces observations et la comparaison avec les résultats obtenus en remplaçant $\eta(t)$ par son signal surrogat indiquent que les composantes du signal de nappe corrélées avec le signal de pluie appartiennent à un domaine de fréquence supérieur à f_l . On peut rajouter à ce qui vient d'être dit qu'une simple observation du signal $\eta(t)$ (voir figure (4.3)) montre l'existence de deux échelles de temps (donc aussi de fréquences) bien distinctes. La première échelle correspondant à de très basse fréquence, liée aux variations saisonnières n'est pas corrélée avec les pluies. La deuxième échelle correspondante au reste du domaine fréquentiel peut contenir des composantes qui elles peuvent être corrélées avec la pluie.

Pour extraire ces composantes corrélées à la pluie, nous allons donc filtrer le signal $\eta(t)$ avec un filtre passe-haut dont la fréquence de coupure sera prise égale à la fréquence f_l . Un filtre d'ordre quatre de type IIR (Infinite Impulse Response) implémenté dans le toolbox de calcul "signal processing" de matlab a été mise en oeuvre pour cela. Le résultat du filtrage donne un nouveau signal que nous appellerons "signal de fluctuation de nappe" et que nous noterons ici $\psi(t)$. Une représentation comparative des signaux $\lambda(t)$, $\eta(t)$ et $\psi(t)$ est donnée sur la figure (4.3).

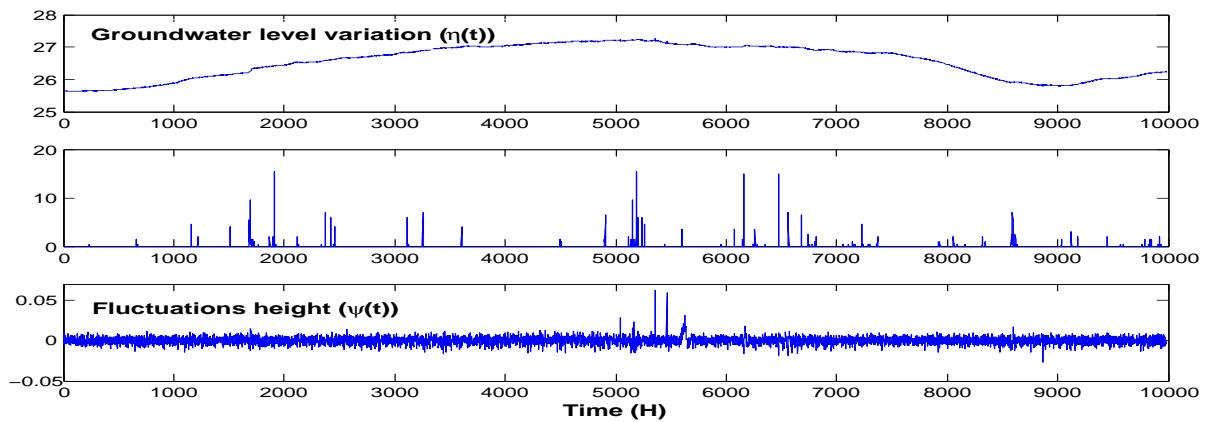


FIGURE 4.3 – Haut : Variation de hauteur de la nappe $\eta(t)$ (mesuré en m). Milieu : Intensité des précipitations $\lambda(t)$ (en mm/h). Bas : Fluctuations de la nappe $\psi(t)$ (in m)

Des vérifications sur d'éventuels déphasages introduits par la procédure de filtrage dans le signal $\psi(t)$ ont été effectué. La figure (4.4) montre que le signal de fluctuations ne présente

aucun déphasage par rapport à $\eta(t)$. Par ailleurs, les résultats de l'estimation de la fonction de cohérence entre le signal de pluie $\lambda(t)$ et les fluctuations $\psi(t)$ est représentée sur la figure (4.5) On y remarque entre autre que la fonction cohérence présente deux pics caractéristiques qui semblent suggérer l'existence de deux échelles de temps distinctes dans la réponse de la nappe aux événements de pluie [40].

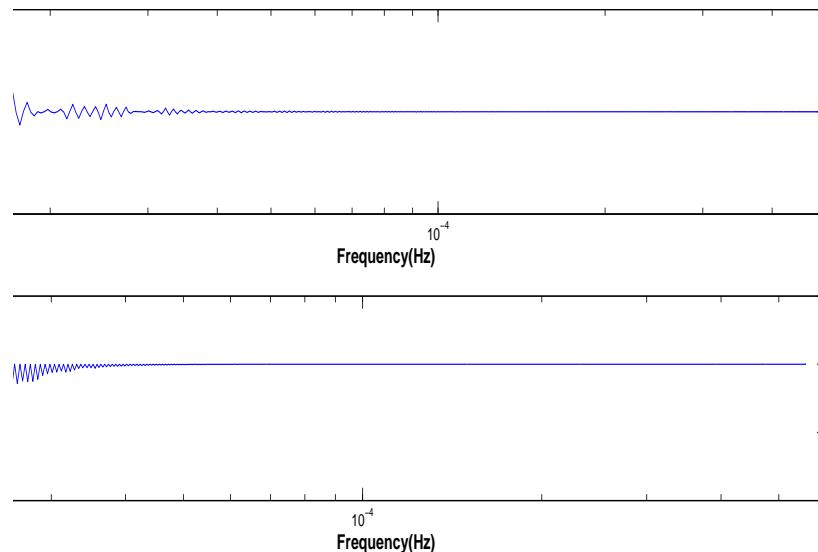


FIGURE 4.4 – Fonction de cohérence normalisée et déphasage entre le signal de hauteur de nappe $\eta(t)$ et le signal de fluctuation $\psi(t)$.

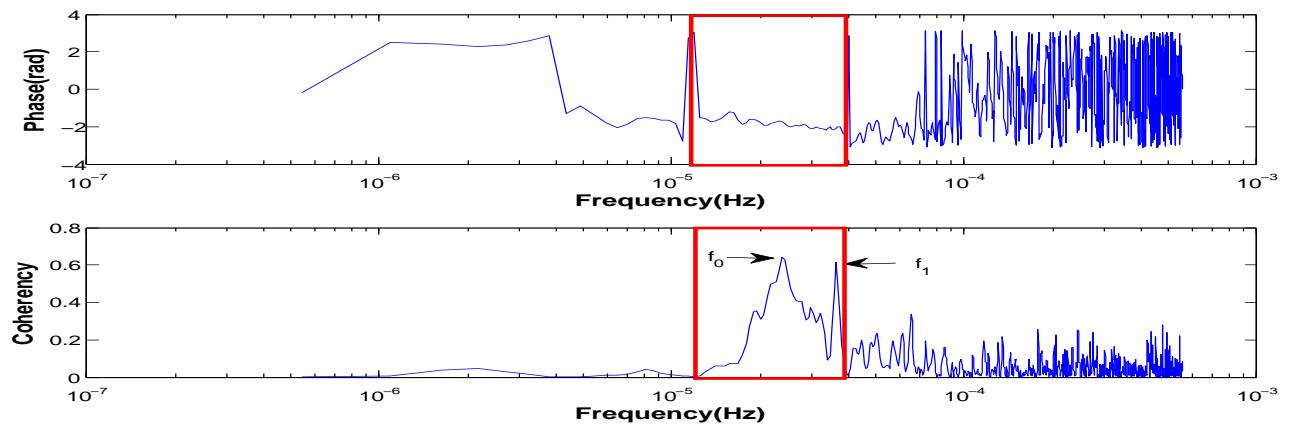


FIGURE 4.5 – Fonction de cohérence normalisée et déphasage entre le signal de pluie $\lambda(t)$ et le signal de fluctuation de nappe $\psi(t)$.

4.2 Caractéristiques du signal de fluctuations de nappe $\psi(t)$

Analyse spectrale des fluctuations de la nappe

La densité spectrale de puissance de $\psi(t)$ est représentée en coordonnées logarithmique sur la figure (4.6). Ce spectre de puissance montre une tendance linéaire entre le pic spectral et la fréquence $2 \cdot 10^{-5}$ Hz. La représentation étant en log-log, cette tendance linéaire traduit une régression formulée par une loi puissance du type :

$S_\psi(f) \approx \text{Constante} \times f^{-p}$ et par conséquent :

$$\log(S_\psi(f)) \approx -P \times \log(f) + \log(\text{Constante})$$

qui est l'équation d'une droite affine de coefficient directeur : $p = 0,602$, calculé à partir du graphique. Il est à noter que la fonction d'auto corrélation admet une régression du même type puisqu'elle est la transformée inverse de la densité spectrale. La relation entre leurs coefficient directeur est alors $p = 1 - \alpha$ si $C(\tau) \approx \tau^{-\alpha}$ soit ici $\alpha = 0,398$ (réf à mettre). Cette tendance linéaire suppose une invariance d'échelle et laisse supposer aussi une nature plutôt fractale des fluctuations de la nappe laquelle est confirmée par l'exposant de Hurst de 0,54 et par conséquent une dimension fractale de 1,46. Il est à noter que [95][96] trouve exactement la même valeur de H. La tendance linéaire qui apparaît entre $6,5 \cdot 10^{-6}$ et $2 \cdot 10^{-5}$ Hz ne se perpétue pas au-delà sur les hautes fréquence du spectre qui ne semble donc pas suivre une loi puissance. Ce comportement laisse présumer que se mêle au signal un bruit qui rend toujours difficile la mise en évidence du chaos, l'estimation de l'exposant de Hurst et donc la dimension fractale d'une série, bien que [78] prétende que dans certains cas, le bruit n'a pas d'effet sur sa nature fractale.

Estimation de la densité de probabilité des niveaux des fluctuations.

Au vu des résultats précédents, une estimation de la densité de probabilité de la série des fluctuations a été approchée à l'aide d'un programme idoine qui a donné les résultats suivants : Index de stabilité $\alpha = 1,9$ ce qui révèle un comportement α -stable et non pas gaussien, un paramètre de symétrie $\beta = -0,36$ donc une faible asymétrie, des paramètres d'échelle et de localisation respectivement de 0,6 et 0. Il est à noter que l'index de stabilité et l'exposant de Hurst [12] sont bien en accord avec la formule de Taqqu [82] qui relie ces deux paramètres pour les processus de variance infinie :

$$2H = 3 - \alpha \tag{4.9}$$

La figure (4.7) montre une comparaison entre l'histogramme issu des données et la loi de densité

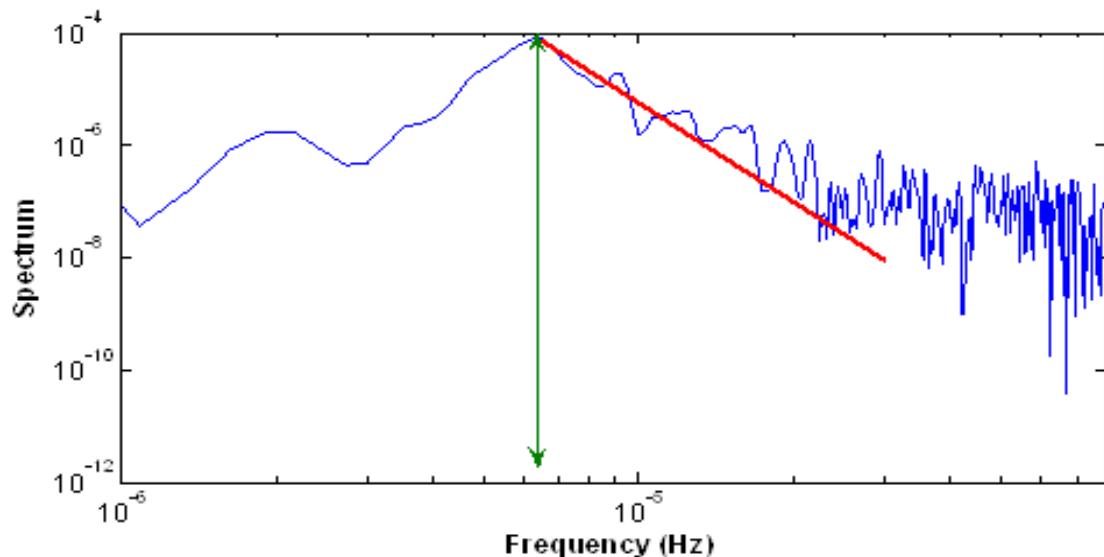


FIGURE 4.6 – Fonction densité spectrale de puissance des fluctuations de la nappe phréatique

de probabilité stable. On note une bonne adéquation entre les résultats expérimentaux et le modèle alpha-stable. La nature fractale, un comportement persistant et enfin une distribution α -stable tendent à traduire une mémoire à long terme dans la dynamique des fluctuations de la nappe à l'instar des caractéristiques des chroniques de pluie que nous avons analysées dans le premier chapitre de ce travail de thèse. L'asymétrie de la distribution d'une part et l'index de stabilité de 1,9 (donc des fluctuations non-gaussiennes) permet de penser que le transport des précipitations dans le sol pour contribuer aux fluctuations de la nappe relève d'une dispersion anormale (expression à prendre ici dans le sens non-linéaire du terme). Dans la section suivante, nous allons mettre en oeuvre cette idée en modélisant ce transport dans le sol par une équation non-linéaire de Richards. Nous allons ensuite étudier la corrélation statistique entre les sorties du modèle de transfert et les fluctuations de la nappe.

4.3 Modélisation du transport de la pluie dans le sol par le logiciel P.A.S.T.I.S.

Le logiciel P.A.S.T.I.S. (Predicting.Agricultural.Solute.Transport in Soils) permet de disposer d'une simulation des transferts de masse (eau, solutés, gaz) et d'énergie (chaleur) couplés à divers processus biogéochimiques dans la zone non saturée du sol. Il fait partie des outils intégrés dans le projet "Sol Virtuel" du Centre de Recherche INRA [37][52][6][16].

Le logiciel PASTIS est un code numérique écrit en langage fortran qui permet de simuler le transfert vertical dans le sol, de l'eau, de la chaleur, des solutés et des gaz. Le code de calcul

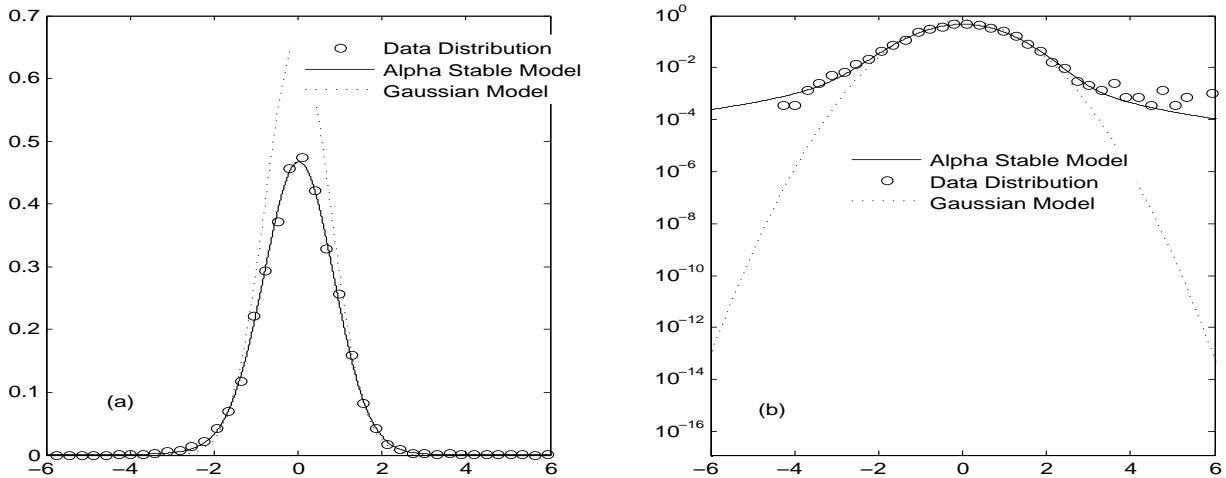


FIGURE 4.7 – Densité de probabilité du signal de fluctuations (symboles "o") et loi stable avec un exposant caractéristique $\alpha = 1.9$

PASTIS est modulaire en ce sens que l'on peut simuler des flux hydriques (version minimale) puis rajouter d'autres phénomènes connexes. Notons que pour notre utilisation ici seul le module "Flux hydrique" de P.A.S.T.I.S. sera utilisé. Les résultats du logiciel PASTIS sont écrits sous forme de fichier numérique de format ascii, ce qui permet de faire des post-traitements, de la visualisation avec le logiciel Matlab.

Le logiciel P.A.S.T.I.S.

Le modèle repose d'une part, sur l'équation de *Richards* qui combine la loi de *Darcy* et la loi de conservation de la masse et d'autre part, sur une modélisation du ruissellement et de son interaction avec l'infiltration. L'équation de Richards mise en oeuvre dans ce modèle s'écrit :

$$C(\psi, z) \frac{\partial H}{\partial t} = \frac{\partial}{\partial z} \left(K(\psi, z) \frac{\partial H}{\partial z} \right) + S(\psi, f(z)) \quad (4.10)$$

avec les définitions suivantes :

1. z [m] Profondeur orientée positivement vers le bas
2. ψ [m] Potentiel matriciel
3. $H = \psi - z$ [m] Charge hydraulique totale
4. $C(\psi, z) = \frac{d\theta}{d\psi}$ [m^{-1}] Capacité capillaire

-
5. $K(\psi, z)$ [ms^{-1}] Conductivité hydraulique.
 6. $S(\psi, f(z))$ Terme puits pour rendre compte de l'absorption racinaire [t^{-1}]
 7. $f(z)$ fonction pour la description du système racinaire
 8. θ [$cm^3 cm^{-3}$] Teneur en eau
 9. t [s] Temps

Le sol est caractérisé du point de vue hydrodynamique par les relations $\psi(\theta)$ et $K(\psi)$.

Ces relations peuvent dépendre de la profondeur. A la surface du sol, le code intègre des conditions aux limites de type Neuman (flux d'eau imposé) ou de type Dirichlet (potentiel matriciel imposé.). Dans notre cas, c'est le flux hydrique qui est imposé à la surface du sol, il correspond à la pluie ou à l'évaporation. L'évaporation (flux négatif) peut être une donnée ou peut provenir du bilan d'énergie quand celui-ci est pris en compte. Jusqu'à maintenant, dans l'utilisation du code, les précipitations (flux positif) sont introduites comme des données *intuitée* à partir de connaissances empiriques sur le régime des pluies. Il faut noter qu'ici, pour la première fois, nous introduisons la chronique des précipitations $\lambda(t)$ comme données à l'entrée du modèle. La condition de Neuman s'écrit :

$$\left[-K(\psi) \frac{\partial H}{\partial z} \right]_{z=0} = F_s(t) \quad (4.11)$$

$F_s(t)$ est un flux d'eau qui correspond à une intensité de pluie ou à une demande évaporative.

Mise en oeuvre de la modélisation du transport

La simulation des flux d'eau utilise une résolution par différences finies des équations décrites ci-dessus. Le pas de temps pour intégrer les équations différentielles résultant de la discréétisation est variable et est automatiquement géré par le code (Cf. Annexe F pour plus de détails sur les grandes lignes de la résolution de PASTIS). Les paramètres de simulation du "Flux hydrique" en entrée concernent

1. Les caractéristiques de la couverture végétale (permet une estimation de l'évaporation).
2. Les propriétés hydrodynamiques du sol du site.
3. La série des précipitations par heure.

Compte tenu de la spécificité de la formalisation de ces paramètres, nous nous sommes appuyés sur un jeu de paramètres "tout fait" dans la base de donnée disponible dans le laboratoire, correspondant à l'état du site atelier durant la période de campagne de mesure. Les résultats

des calculs faits par le code PASTIS donnent une chronique de l'état hydrique du sol (que nous noterons $\phi(t)$) pour une profondeur donnée dans la zone non-saturée sur la période concernée soit du 2 février 2005 au 27 mars 2006. On remarque que le code dans la formulation que nous utilisons (formulation en pression) ne permet pas une résolution jusqu'à la profondeur de battement de la nappe qui se trouve plus bas. Compte tenu de cette remarque et du fait que les caractéristiques hydrodynamiques du site n'étant disponible que sur une profondeur de 1.6 mètre, nous avons simulé l'état hydrique à cette profondeur. Les résultats obtenus dans cette simulation vont être comparés avec le signal de fluctuations de la nappe.

4.4 Analyse des corrélations entre les précipitations, l'état hydrique à 1.6m et les fluctuations de la nappe

Précipitations, fluctuations de la hauteur de la nappe phréatique et fluctuations de l'état hydrique à 1.6m sont examinées sur la même période de 14 mois, du 2 février 2005 au 27 mars 2006. Les planches ci-après, pour chaque couple :

1. Pluie-Etat hydrique à 1.6m.
2. Etat hydrique à 1.6m-fluctuations de la nappe,
3. Pluie-fluctuations de la nappe,

mettent en parallèle par la méthode des "transformées en Ondelettes" :

1. La corrélation croisée,
2. Les séries temporelles des couples concernés,
3. Les phases de la relation croisée.
4. Les spectres de puissance de chacun des éléments du couple.

Analyse des corrélations entre les signaux par la transformation en ondelette

La représentation graphique d'un signal temporel fini $f(t)$ permet de visualiser ses variations.

Faire apparaître les périodicités ou les fréquences de ce même signal, supposé stationnaire, relève de la transformation de Fourier qui le décompose dans le domaine fréquentiel. Une telle transformation informe sur la régularité globale du signal mais pas sur un comportement apparaissant uniquement sur un intervalle de temps donné. Ceci est le cas de nos événements de pluie. A contrario, la transformée en ondelette permet de mettre en évidence certaines

structures du signal bien localisées, dans le temps, des évènements qui présentent à certaines fréquences, une relation d'échelle. Parmi les variantes possibles de transformation en ondelette, nous utilisons la transformation en ondelette continue de Morlet, elle est réputée bien adaptée pour analyser des signaux de type géophysiques [45].

Spectre en ondelette

A l'instar de la définition du spectre de Fourier, nous appelerons densité spectrale en ondelette la quantité

$$E(s, t) = \frac{|g(s, \tau)|^2}{C_g s} \quad (4.12)$$

Dans cette équation, la quantité C_g constitue un facteur de normalisation tandis que τ est un paramètre de translation et "s" un paramètre de dilatation et $g(s, \tau)$ est la transformée en ondelette du signal à analyser (Cf. Annexe G). La représentation graphique de $E(f, t)$ porte aussi le nom de scalogramme. Par un changement de variable $s = \frac{f_0}{f}$, le scalogramme donne une représentation en temps-fréquence de la distribution d'énergie.

Corrélation croisée en ondelette

Considérons deux signaux $\lambda(t)$ et $\psi(t)$ qui ont respectivement comme transformée en ondelette $g(s, \tau)$ et $h(s, \tau)$. On définit la fonction de corrélation croisée en ondelette entre $\lambda(t)$ et $\psi(t)$ par l'expression

$$CWT_{\lambda\eta}(s, \tau) = \frac{g(f, \tau)^* h(f, \tau)}{s} \quad (4.13)$$

Compte tenu du caractère complexe de la quantité $CWT_{\lambda\eta}(\tau)$, il est possible de l'écrire comme des amplitudes en ondelettes et un déphasage. Notons que les définitions relatives aux équations (4.12) et (4.13) sont données et explicitées dans l'annexe G).

Dans ce qui suit, les corrélations entre les signaux sont estimées à l'aide des équations (4.12) et (4.13).

Résultats et interprétations

Les planches ci-dessous montrent, en fonction des couleurs, du bleu au rouge, du moins corrélé au plus corrélé, le degré de corrélation des couples aux dates auxquelles les évènements sont observés. La durée des évènements se lit en ordonnée sur le log à base 2 de l'échelle des temps. De même, les phases de la relation croisée sont évaluées par une échelle de couleurs de -3 à +3 (en radian) qui correspondent à $\pm \pi$. Les densités spectrales de puissance sont les carrés des modules des transformées en ondelette des signaux. Les résultats de l'estimation de la densité spectrale montrent la répartition d'énergie du signal dans le temps ainsi que leur durée

d'apparition. La même construction portant sur les modules et non plus les carrés, montrerait pour les fluctuations de l'état hydrique du sol, la répartition en volume d'eau. Notons qu'afin de mieux observer les résultats, l'échelle de temps a été rapportée en jour de la présentation des résultats.

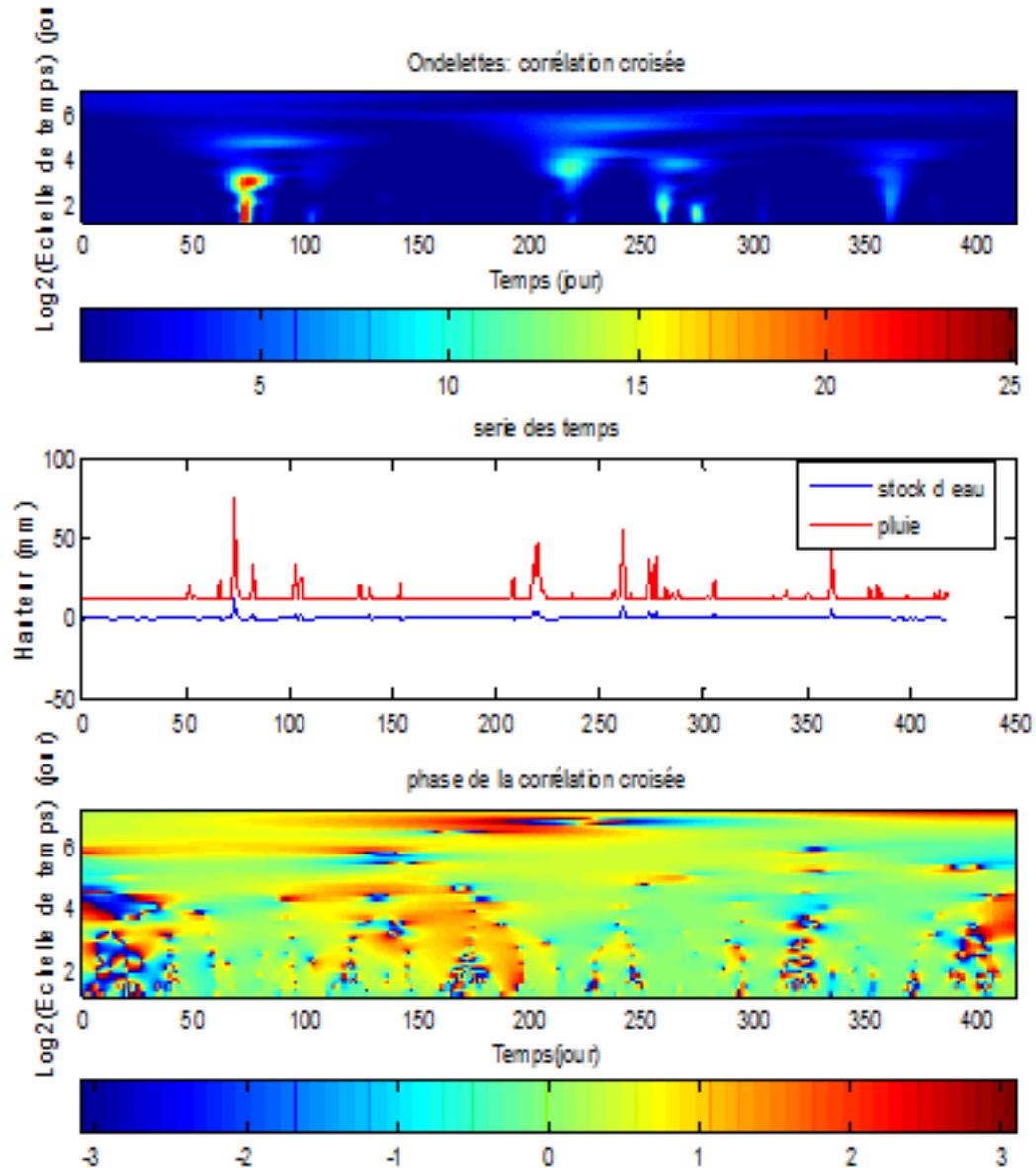


FIGURE 4.8 – Corrélation croisée entre précipitations et fluctuations de l'état hydrique

Corrélation croisée : Les pics importants des séries temporelles correspondent bien, (du bleu clair au rouge ou encore du degré 10 à 25), aux corrélations croisées : pluie-fluctuations de l'état hydrique à 1.6m, avec un certain décalage qui est le temps de transfert de l'eau de pluie sur 1,6 m de profondeur.

Phases de la corrélation croisée : L'ensemble de la figure présente peu de changements de phases car prédominent de larges plages de couleurs pales donc de faible corrélation. La plus grande

densité de ces changements, montrant une forte corrélation se situe sur l'échelle des temps pour des événements d'une durée de 0 à 2.

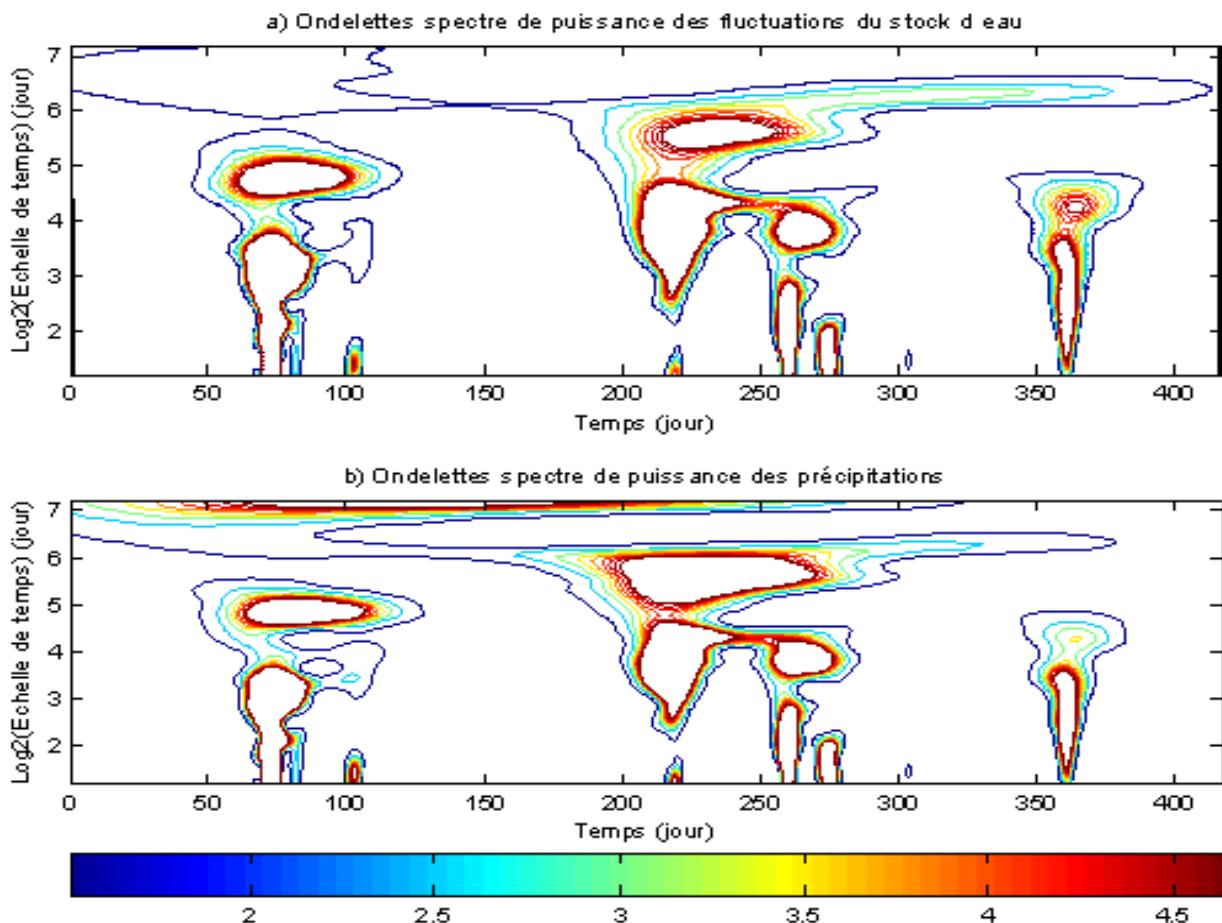


FIGURE 4.9 – Densités spectrales de puissance des précipitations et des fluctuations de l'état hydrique à 1.6m

Le spectre de puissance donne pour toute valeur de la fréquence, la puissance transportée par cette fréquence. Le terme “puissance” se rapporte à une énergie moyenne transportée par unité de temps, dans l’hypothèse où l’on peut associer un terme d’énergie au carré de l’amplitude du signal. Ici le terme consacré ”énergie” peut être remplacé par ”intensité du flux”.

Sur cette figure, apparaissent des flux d’eau quasiment simultanés, les 75^e, 250^e et 360^e jour, tant au niveau du sol qu’à 1,6 m de profondeur.

Corrélation croisée : Les pics importants des séries temporelles correspondent bien, (du bleu clair au rouge ou encore du degré 3 à 9), aux corrélations croisées : fluctuations de l’état hydrique à 1.6m-fluctuations de la nappe, avec un certain décalage qui est le temps de transfert du stock d’eau de 1,6 m de profondeur jusqu’à la nappe. On remarque que les corrélations sont plus fortes entre les fluctuations de l’état hydrique à 1.6m et la nappe qu’entre les précipitations et les fluctuations de l’état hydrique à 1.6m, ce qui tend à prouver que peu d’eau est retenu par

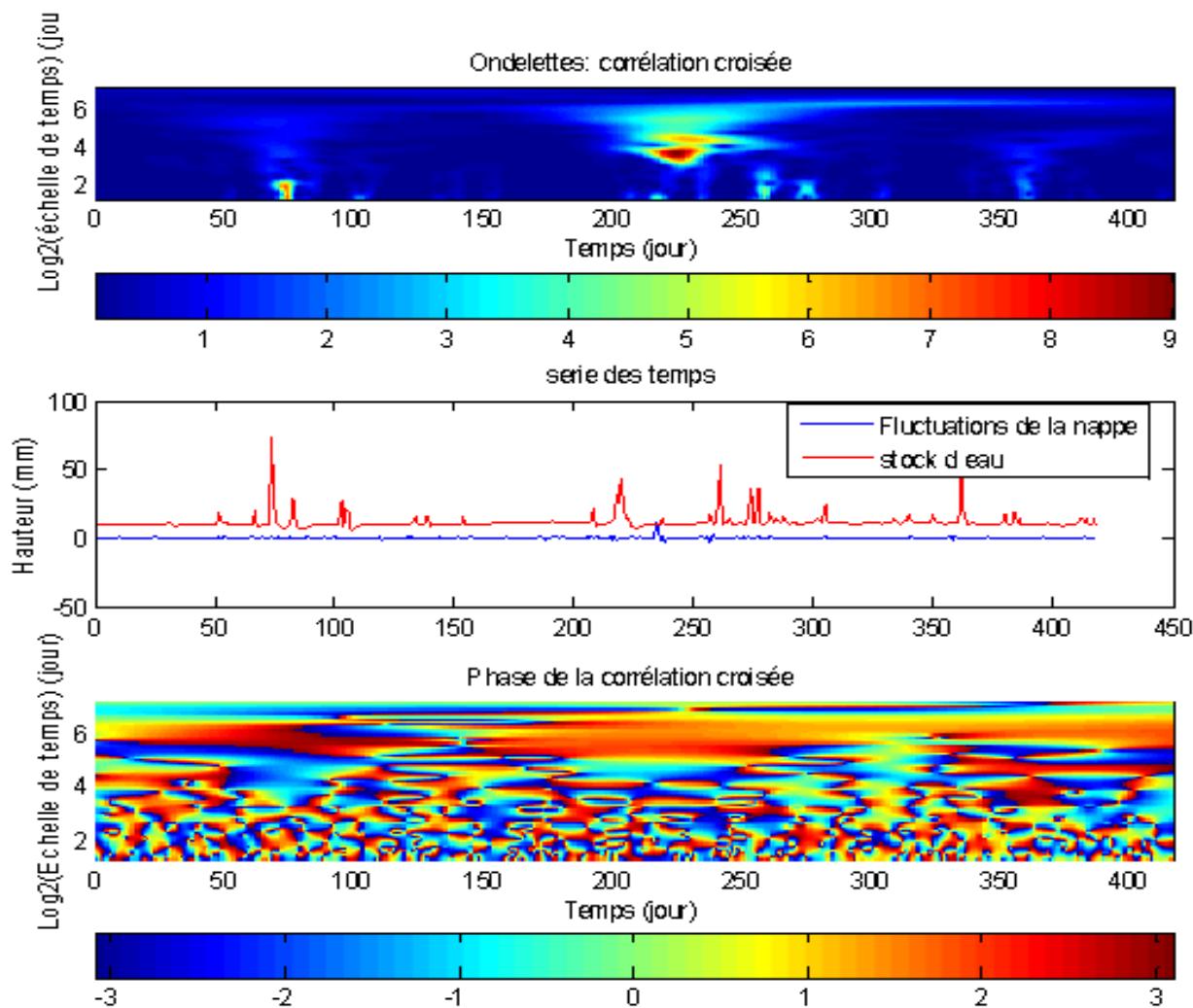


FIGURE 4.10 – Corrélation croisée entre fluctuations de l'état hydrique à 1.6m et fluctuations de la nappe

le sol entre 1,6 m de profondeur et la nappe.

Phases de la corrélation croisée : L'ensemble de la figure présente de nombreux changements de couleurs donc de phases pour des évènements d'une durée de 0 à 4 sur l'échelle des temps alors que les couleurs varient peu au-delà où persiste une forte stabilité de la phase. Les petits évènements présentent donc une alternance de forte et faible corrélation alors que les évènements de longue durée, 6 sur l'échelle, sont fortement corrélés entre stock d'eau et nappe. Ceci montre à l'évidence, que les évènements prolongés influent plus directement la hauteur de nappe que les petits.

Le spectre de puissance donne pour toute valeur de la fréquence la puissance transportée par cette fréquence. Le terme "puissance" se rapporte à une énergie moyenne transportée par unité de temps, dans l'hypothèse où l'on peut associer un terme d'énergie au carré de l'amplitude du signal. Ici le terme consacré "énergie" peut être remplacé par "intensité du flux".

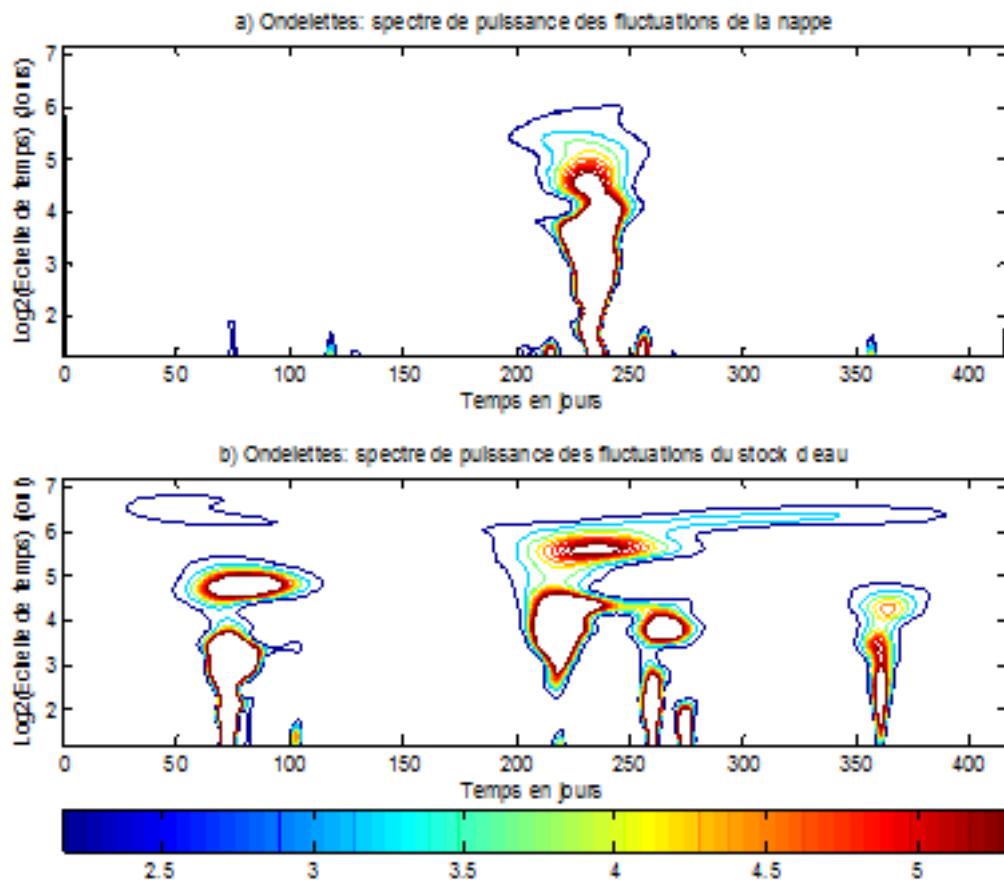


FIGURE 4.11 – Spectres de puissance des fluctuations de l'état hydrique à 1.6m et celles de la nappe

Sur cette figure, apparaissent avec un certain décalage, des fluctuations corrélées de l'état hydrique à 1.6m et de la nappe aux environs du 250^e jour alors que les fluctuations de l'état hydrique à 1.6m au 75^e jours ne laisse pas de trace sur celles de la nappe, ce qui laisse à penser que le flux d'eau correspondant repéré à 1,6 m de profondeur n'atteint pas la nappe.

Corrélation croisée : Les corrélations des étapes précédentes se retrouvent bien dans le transfert direct pluie-nappe, toujours avec une intensité comparable (échelle des couleurs de 0 à 8). Le décalage est toujours conservé.

Phases de la corrélation croisée : On retrouve ici, dans le transfert global pluie-nappe, les mêmes comportements que dans l'étape précédente. Ce qui tend à montrer que des quantités d'eau autres que celles du stock surviennent entre la cote 1,6 m et la nappe.

Comme précédemment, les fluctuations de la nappe qui ne semblaient pas affectées par les fluctuations de l'état hydrique à 1.6m de profondeur des 75^e et 360^e jour, ne sont pas non plus affectées par les précipitations de ces dates là.

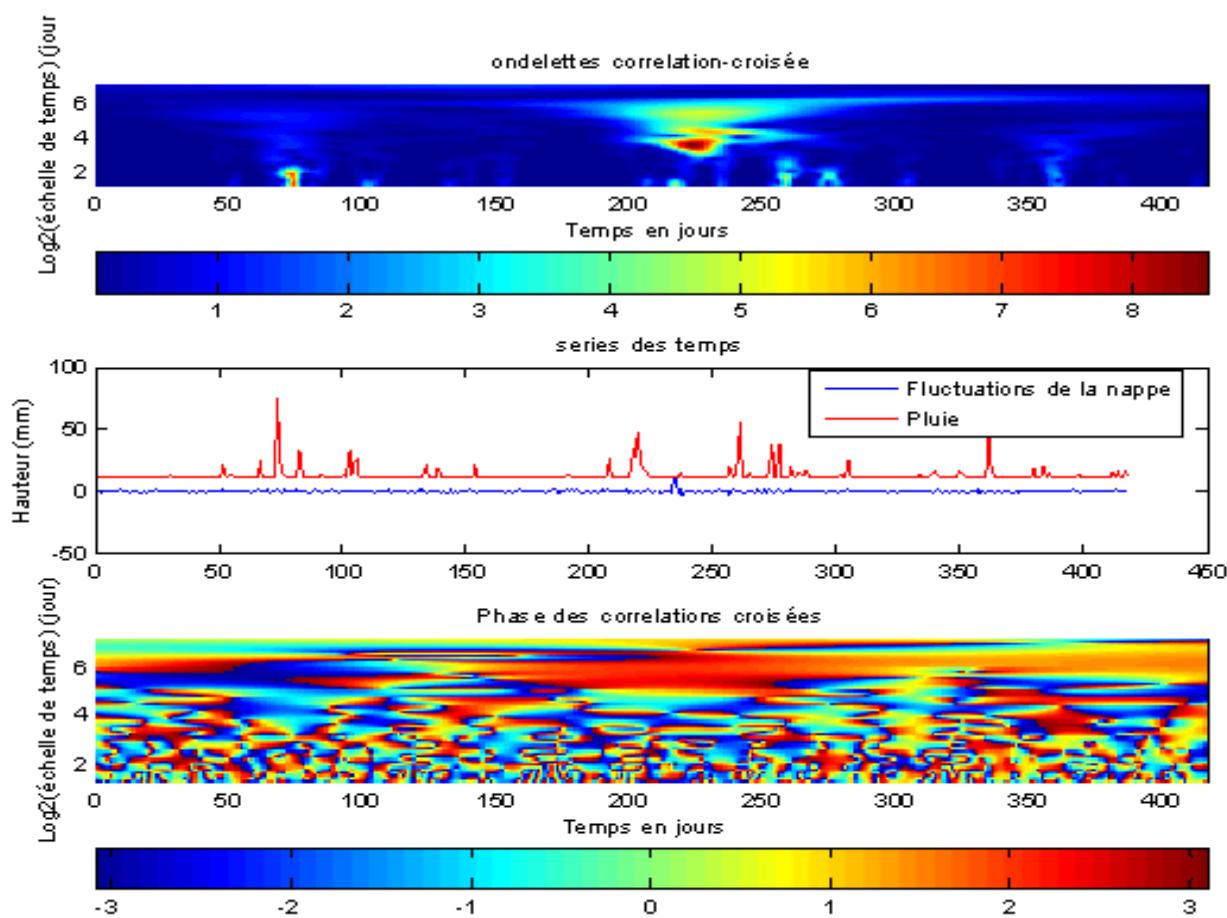


FIGURE 4.12 – Corrélation croisée entre fluctuations de la nappe et précipitations

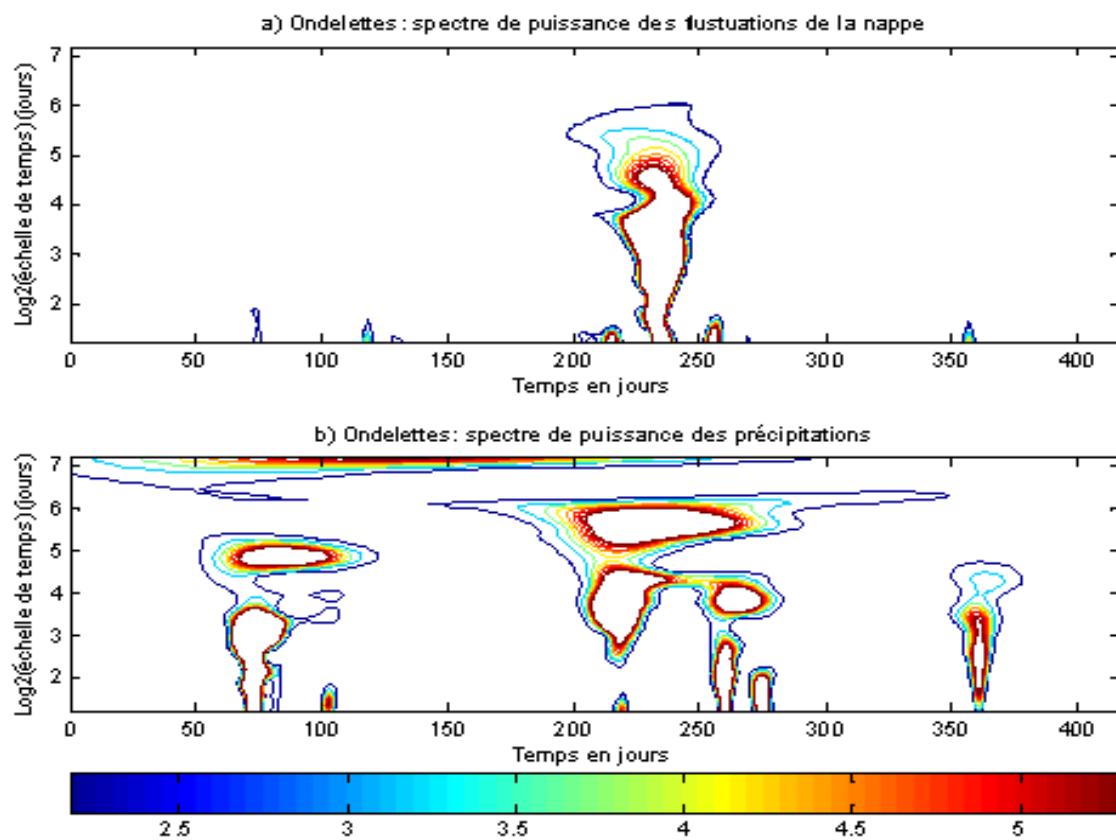


FIGURE 4.13 – Spectres de puissance des précipitations et des fluctuations de la nappe

4.5 Conclusion du 4ème chapitre.

Comme il a été annoncé à l'introduction de cette thèse, dans ce chapitre nous avons étudié sur la base d'expérience in situ le transfert de la pluie vers la nappe phréatique avec en filigrane l'idée que ce transfert constitue un vecteur principal de la pollution de la nappe par les polluants à la surface du sol. Cette étude a été possible grâce au site atelier "Fontanille" du laboratoire EMMAH qui à notre avis offre un cadre exceptionnel pour observer et comprendre les mécanismes de transfert des eaux de pluie vers la nappe. Rappelons que les données utilisées sont constituées des valeurs de précipitations et de la variation de la hauteur de nappe, acquises de manière synchronisée. Tout d'abord une simple observation de la variation de la hauteur de nappe nous a permis de nous apercevoir que la variation à l'échelle saisonnière de la nappe ne peut pas présenter une corrélation avec la pluie. Nous avons alors isolé les composantes de la variation de la hauteur de nappe en rapport avec l'échelle de temps de la pluie par un filtrage passe haut. Nous avons ensuite les propriétés statistiques des fluctuations ainsi obtenues. Nous avons montré que ces fluctuations présentent des caractéristiques comparables avec celles trouvées au premier chapitre dans les précipitations sur quarante ans, en particulier les fluctuations de nappe semblent aussi présenter des effets mémoires à long terme. D'autre part, l'histogramme des fluctuations suit assez bien une loi α -stable. Nous avons aussi trouvé que l'exposant de stabilité et l'exposant de Hurst de ces fluctuations vérifient un lien théorique connu [82]. Cela suggère qu'un des modèles candidats pour représenter la dynamique du processus sous-jacent à ces fluctuations pourraient être de type mouvement brownien fractionnaire. Notons que si cette idée peut-être mieux confirmée, une piste de modèle de dispersion anormale en temps pourra être retenue pour modéliser ces fluctuations. En effet les mouvements browniens fractionnaires sont connus pour être lié à des modèles de dispersion fractionnaire [59]. Une telle piste a déjà été observée et explorée dans d'autres domaines [53][74]. Après cette étape, nous avons utilisé le modèle PASTIS avec comme entrée les données observées de pluie. Une estimation des corrélations croisées en ondelette entre les fluctuations de nappe et la sortie du modèle PASTIS à la profondeur maximale accessible a été effectuée pour montrer les liens statistiques entre les deux grandeurs. D'autre part on note que la similitude entre ce qu'on observe comme corrélation en utilisant le signal de pluie mesuré et en utilisant le signal de sortie de PASTIS montre le potentiel de ce code de calcul comme outil pour modéliser le transfert de masse dans le sol.

Conclusion

Comprendre le devenir des éléments polluants à la surface constitue un des objectifs du laboratoire EMMAH. Cela nécessite de développer des pistes de recherche aussi bien expérimentale que théorique. L'inter-dépendance et les liens entre ces deux aspects sont indispensables pour créer un outil de modélisation efficace en rapport avec l'ambition de projet tel que la plateforme "Sol Virtuel". Ceci nécessite des observations en condition naturelle avec des contextes et des conditions de mesure plus difficiles par rapport aux expériences en laboratoire. En particulier, les mesures *in situ* sont plus marquées par les influences de phénomènes connexes qu'on ne peut pas facilement découplées des mesures. Celles-ci contribuent en général à faire apparaître des artefacts et des bruits plus importants dans les mesures *in situ*. Pour optimiser les conditions de mesure et mieux cerner les phénomènes à mesurer, le laboratoire EMMAH a mis en place le site atelier "Fontanille" dont un des caractéristiques est de se situer entre le cadre d'un laboratoire et celui du site naturel. Malgré cela, les observations ne peuvent pas être répétées à souhait, en particulier en ce qui concerne les phénomènes dont on ne maîtrise pas le déclenchement ou encore ceux qu'on doit mesurer sur une très longue période. Il a fallu donc à partir d'un seul jeu de données en tirer le maximum d'informations. Pour cela, nous avons élargi la gamme des outils d'analyse utilisés au détriment des démarches qui permettent d'aller plus en profondeur sur l'utilisation d'une méthode. Par ailleurs, des hypothèses et des contextes simplificateurs ont été adoptés pour analyser les données. L'étude faite dans cette thèse est articulée autour de relevés expérimentaux réalisés sur le site atelier, sur une zone de dimension limitée en considérant comme négligeable des paramètres tels que la dimension spatiale du bassin, les apports d'eau autres que la pluie d'un côté ou les déversements de l'autre, autant d'éléments dont il faudrait tenir compte pour une étude à plus grande échelle. Sur le plan théorique, les hypothèses permettant d'utiliser nos techniques d'analyse ont été retenues sans forcément trouver de justifications irréprochables à partir des données, de même en ce qui concerne les cadres mathématiques de nos travaux de développement. Néanmoins, notre étude aboutit à deux résultats permettant

d'une part de décrire un modèle de processus suivi par le cumul des précipitations et d'autre part, une caractérisation de la dynamique des fluctuations de la nappe phréatique en réponse à ces précipitations.

Le modèle de précipitation a été conçu, dans un premier temps, à partir d'une marche aléatoire dont les paramètres sont issus de l'analyse de la série expérimentale. Les sauts de la marche aléatoire suivent une loi de distribution log-normale et les temps d'attente une loi α -stable. Ce premier modèle donnant des simulations approchant valablement la courbe expérimentale tant par la position de celle-ci au sein des déciles que par son voisinage avec la courbe moyenne des simulations, nous sommes passés à la limite de cette marche aléatoire lorsque le pas de temps et celui de la marche tendent simultanément vers zéro en gardant un rapport constant, désigné habituellement sous le nom de coefficient de diffusion. Cette limite se traduit par une équation aux dérivées partielles de la fonction de densité de probabilité pour que le cumul des hauteurs de pluie atteigne une valeur x à un instant t . Cette équation ayant les mêmes solutions que l'équation différentielle stochastique de Ito associée, cette dernière permet de construire des trajectoires simulant les données. En appliquant la méthode de Monte-Carlo à grand nombre de simulations on peut en déduire, à un temps t donné, l'histogramme des valeurs possibles. Le pic de cette distribution correspondra à la hauteur de pluies cumulées la plus probable à une date donnée. En effet, compte tenu de notre approche stochastique, à la place d'une solution déterministe calculée, notre modèle donne une estimation de la valeur la plus probable à un instant donné. Rappelons aussi que les paramètres du modèle proviennent des lois de probabilités des précipitations locales mais que la procédure d'estimation de ces paramètres permettent d'envisager l'utilisation de notre approche pour d'autres situations climatiques. En ce qui concerne notre région, les lois à ailes épaisses et les mémoires à long terme retrouvés pour les données de précipitations ne sont pas sans lien avec le caractère prédominant des événements dits "extrêmes". Plus précisément, nous pouvons dire que la dynamique des occurrences de pluie et donc des temps d'attente portent la signature de ces événements "extrêmes" formés de longues périodes de sécheresse. Enfin si le fléchissement des données à partir de 1997 est accepté comme un phénomène naturel alors nous pouvons aussi confirmer que le rôle de ces événements "extrêmes" est amené à augmenter dans l'avenir.

Dans la deuxième partie du travail, nous avons d'abord montré que la dynamique des fluctuations des hauteurs de la nappe phréatique corrélées aux précipitations, présente une distri-

bution qui possède des "ailes épaisses" que l'on peut "fitter" assez bien avec une loi α -stable de variance divergente. D'autre part, ayant décelé aussi des effets de mémoire à long terme, nous suggérons que cette dynamique peut être représentée par un modèle de dispersion fractionnaire pouvant être associé au mouvement brownien fractionnaire. Nous avons ensuite estimé différentes corrélations statistiques avec les fluctuations de la nappe, en particulier la corrélation entre cette dernière et la sortie d'un modèle de transfert (PASTIS) quand on impose en entrée les précipitations mesurées. Les résultats de ces estimations de corrélations statistiques laissent entrevoir l'intérêt de la démarche en ondelette adoptée ici en vue de la confrontation entre modèle de transfert et donnée de nappe. Toutefois, notons qu'un approfondissement et une meilleure interprétation de ces résultats seront nécessaires pour cela.

Pour finir, deux points méritent d'être soulignés en perspective à ce travail de thèse. Tout d'abord, la suite naturelle de ce travail concernera le couplage du modèle de précipitation décrit dans la première partie de ce mémoire avec un modèle de transfert (PASTIS et/ou autres modèles). Ensuite, notons de manière générale que les résultats et les analyses obtenus dans ce travail méritent d'être repris dans un cadre plus large et plus général. L'utilisation des démarches proposées ici pour des études au-delà du cadre du site atelier constituent pour nous un intérêt primordial. En particulier dans le cadre des études *in situ* menées sur le site de Rustrel (un autre site d'étude du laboratoire EMMAH) concernant la réponse d'un système karstique aux précipitations ou encore pour l'étude des impacts de précipitations pour des régimes climatiques tropicaux dominés par des événements "extrêmes" de type cycloniques comme ce que l'on peut rencontrer à Madagascar.

En terme de valorisation du travail effectué, la première partie de cette thèse a fait l'objet d'une communication dans un congrès international et d'une publication en cours de révision dans un journal [22]. Tandis que la deuxième partie de la thèse fait l'objet d'une publication en cours de soumission [23]. Notons que nous avons aussi deux résultats publiés qui ne concernent pas directement le sujet de la présente thèse [21][36].

Annexe A

Dimension fractale et corrélation entre accroissements passés et futurs d'une série

Rappelons que la dimension fractale (notée ici D) qui caractérise la nature fractale d'une série [49], est liée à l'exposant H par la relation par

$$D = n + 1 - H \quad (\text{A.1})$$

avec $n=1$ (dimension physique des données). D'autre part, Feder [14] a proposé de calculer la corrélation entre accroissements passés et avenir à partir de la relation A.2.

$$C = 2(2^{2H-1} - 1) \quad (\text{A.2})$$

Notons que C n'est pas la corrélation classique, elle est croissante entre 0 (pour un signal dépourvu de mémoire) et 2 (lorsque $H = 1$). Si la corrélation classique détermine les dépendances à court terme [48], ici C est un indicateur de la dépendance des valeurs de la série à l'histoire.

Ces deux paramètres établis après l'évaluation de l'exposant de Hurst figurent dans le tableau ci-dessous. A titre de comparaison, y figurent aussi les dimensions de la "poussière de Cantor" et la courbe de Koch. Il apparaît que toutes les dimensions fractales sont inférieures à celles d'une série relevant d'une distribution normale, ce qui traduit une apparence graphique moins erratique que la normale. La dimension fractale diminuant lorsque l'exposant de Hurst ainsi que la persistance augmentent, on remarque que si les fractions de hauteurs de pluie sur 10 ans ont à peu près la même dimension, la hauteur de pluie sur 40 ans a une dimension bien inférieure. Ce qui peut se traduire par : "la persistance globale est la somme des persistances qui la composent". La partition des hauteurs de pluie avant et après 1997 montre bien

que la persistance est plus forte après 1997, la sécheresse imposant des temps d'attente plus persistants qu'avant 1997. Ce que confirme bien la dimension des temps d'attente. Les pluies cumulées présentent la plus petite dimension fractale du tableau, son allure graphique doit alors être plus régulière que les autres, ce qui se vérifie sur les diverses représentations.

La corrélation des accroissements passés-futurs est, elle, une fonction croissante de H , et là encore, les hauteurs de pluies globales semblent bien cumuler les corrélations des hauteurs de pluie partielles qui la composent. Pour les deux périodes avant et après 1997, les accroissements de la seconde sont les mieux corrélés du fait de l'importance des temps d'attente. Les précipitations ont des accroissements petitement corrélés du fait du climat méditerranéen où peuvent alterner des pluies d'importance très différentes. Les accroissements des pluies cumulées recèlent une forte corrélation induite par les accroissements qui restent nuls en l'absence de pluie.

Pour parachever l'"apparence" des signaux on calculera leur dimension fractale qui pourra être appréciée par rapport à des schémas bien connus comme la poussière de cantor ou la courbe de Koch. Enfin il peut exister une corrélation entre accroissements passés et avenir, on utilisera, alors, la formule établie par [14] qui quantifie cette corrélation sachant qu'une distribution normale admet une "Corrélation des accroissements passés-futur" nulle.

Série de données	Dimension fractale	Corrélation
Séries test	0,63	$C = 0$
Poussière de Cantor	1,26	Bruit blanc
Courbe de Koch		
Série Normale : randn-Matlab	Théorique $D = 1,5$ $D=1$	$C=2$
Rappel : Forte persistance		
Série temporelle totale 1968-2008	$D = 1,36$	$C = 0,43$
10 premières années : 1968-1977	$D = 1,49$	$C = 0,24$
10 années suivantes : 1978-1987	$D = 1,45$	$C = 0,15$
10 années suivantes : 1988-1997	$D = 1,43$	$C = 0,21$
10 années suivantes : 1998-2008	$D = 1,42$	$C = 0,245$
"Avant 1997" : 1968-1997	$D = 1,43$	$C = 0,19$
"Après 1997" : 1997-2008	$D = 1,39$	$C = 0,25$
Intensités des pluies	$D = 1,45$	$C = 0,15$
Temps d'attente	$D = 1,34$	$C = 0,48$
Pluies cumulées	$D = 1,17$	$C = 1,16$

TABLE A.1 – Les dimensions fractales sont inférieures à celles d'une série normale ce qui induit une certaine persistance, sauf pour les pluies cumulées où la très forte persistance est due à la construction de la série.

Annexe B

Pluies et sècheresses suivent-elles un processus markovien ?

Les paragraphes précédents montrent que les diverses séries qui servent de support à la détermination du processus appelé à décrire le phénomène pluvieux étudié, présentent une mémoire à long terme. En revanche, quelles observations peut-on faire si l'on ramène notre phénomène pluvieux à un système qui, chaque jour peut ne présenter que deux états : pluie ou sècheresse, en supposant que la probabilité d'être dans un état à un instant donné n'est conditionnée que par le seul fait d'être dans l'autre état à l'instant d'avant donc ne pas tenir compte des états antérieurs à celui-ci. Ce qui se traduit par "l'état futur ne dépend que de l'état présent". Un tel système constitue une Chaine de Markov en Temps Discret (C.M.T.D) et définit une variable aléatoire apériodique et évidemment irréductible dont on peut déterminer la matrice de transition si on connaît la répartition des temps où le système est dans un des états ou inversement[31][24]. On suppose, bien sûr que les données permettent de calculer les probabilités stationnaires, i-e les probabilités pour que la variable soit dans un état donné après un temps infini.

Ici le théorème utilisé est le suivant :

Théorème :

Soit un C.M.T.D. portant sur un ensemble d'états E et définissant une variable aléatoire X, apériodique et irréductible ;

Note : [X est irréductible s'il existe un chemin direct de tout état à un autre].

Si π_n est la probabilité stationnaire de X en position n, n parcourant l'ensemble E des états :
i-e $\pi_k = \lim_{k \rightarrow \infty} \text{Pro}(X_k = n)$;

Si P est la matrice de transition d'un état à l'autre ;

Il existe une probabilité π telle que : 1) $\pi = \pi P$ et 2) $\sum_{i \in E} \pi_i = 1$

Appliqué à notre système pluvieux on obtient les résultats suivants :

Le système pluvieux étant 21% du temps en état "pluie" et 79% du temps en état sécheresse, ces valeurs peuvent être une approximation des états stationnaires, c'est-à-dire les probabilités π_p et π_s du théorème ci-dessus. (π_p étant la probabilité pour que le système soit en état "pluie" et respectivement π_s pour qu'il soit en état "sécheresse"). La matrice de transition d'un état à l'autre s'écrit alors :

$$\begin{pmatrix} \pi_p & \pi_s \end{pmatrix} = \begin{pmatrix} \pi_p & \pi_s \end{pmatrix} \begin{pmatrix} 1 - \alpha & \alpha \\ \beta & 1 - \beta \end{pmatrix} \quad (\text{B.1})$$

si α et β sont respectivement les probabilités de transition de l'état "pluie" à l'état "sécheresse" et réciproquement. Le calcul ne permet d'obtenir que le rapport $\frac{\beta}{\alpha} = \frac{21}{79}$ soit encore : la Probabilité de passer de l'état sécheresse à l'état pluie est égale à 27% de la probabilité de passer de l'état pluie à l'état sécheresse. Ce qui traduit bien le fait que, localement, on a plus de chances d'avoir du beau temps que de la pluie. Notre système pluvieux local, même s'il présente une mémoire à long terme et par conséquent ne relève pas d'une chaîne de Markov en temps discret, un tel résultat reste une approximation qui permettra, ultérieurement, lors de simulations de trajectoires du processus supposé gérer le phénomène pluvieux, de sélectionner les temps d'attente entre deux jours de pluie présentant des caractéristiques voisines. Une étude forte intéressante de ce type a déjà abouti à un modèle mettant en relation les précipitations, le déficit hydrique du sol et le ruissellement en utilisant une chaîne de Markov pour modéliser les occurrences de pluie et de sécheresse [26]. Plus généralement, les modèles de chaînes de Markov fournissent une représentation mathématique simple de précipitations journalières ou même seulement de leurs occurrences mais ceci dans des cas de phénomènes à forte persistance dans certaines régions du globe ou en certaines saisons.

Annexe C

Méthodes d'estimation de l'exposant de stabilité α

Méthode de McCULLOCH Cette méthode [54] [55] a pour principal inconvénient de n'être valable que pour $\alpha > 0,6$ et repose sur le rapport des divers quantiles d'une série.

Si on pose $V\alpha = \frac{X_{0.95} - X_{0.05}}{X_{0.75} - X_{0.25}}$ et $V\beta = \frac{X_{0.95} + X_{0.05} - 2 * X_{0.5}}{X_{0.95} - X_{0.05}}$

des tables donnent alors une valeur approchée de α de même pour les paramètres β , σ et μ , si alpha est supérieur à 0,6 .

La méthode de McCULLOCH appliquée aux simulations tests α -stable d'index connus 0,6 et 0,9 donne un résultat conforme aux attentes avec une erreur de 5% environ. On peut remarquer que la méthode n'a pas été appliquée à une simulation 0,4-stable puisque limitée à alpha supérieur à 6. La méthode de calcul de α pour les temps d'attentes est donc recevable. Les estimations des autres paramètres sont l'antisymétrie $\beta = 1$, échelle $\sigma = 0,653$ et localisation $\mu = -5,7954$. Le Tableau C.1 résume les résultats portant sur la série des temps d'attente telle qu'elle a été retenue et lui accorde une valeur de α comprise entre 0,87 et 0,94. Pour lever le doute sur la façon d'interpréter la notion de temps d'attente entre deux pluies exposée précédemment dans l'analyse du signal, figurent aussi les résultats de la seconde façon de considérer les temps d'attente dont l'index de stabilité est très inférieur. Pour confirmer, en première approche, la valeur trouvée de α , on compare les distributions de la série des temps d'attente avec celles de plusieurs simulations stables d'index 0,7 et 0,9 déterminés ci-dessus. Seule la distribution 0,9-stable montre une allure comparable à celle des temps d'attente, ce qui tend à conforter le choix de temps d'attente nulle entre deux jours de pluie consécutifs.

Rappel des dénominations :

ATNUL = Temps d'attente entre deux jours de pluie, le jour où il pleut est compté comme zéro temps d'attente.

	ATNUL	Série test 0,6-stable	Série test 0,9-stable
Valpha	5,5		
Vbeta	1		
α	0,9235 $0,87 < \alpha < 0,94$	0,6 avec une erreur de 5%	0,9 avec une erreur de 5%
β	1		
σ	0,653		
ν	-5,79		

TABLE C.1 – Méthode de McCulloch

Méthode des moments Press [69] a proposé la "méthode des moments" basée sur une transformation de la fonction caractéristique d'une série x_i :

$\widehat{\varphi}(t) = \frac{1}{n} \sum_{i=1}^n e^{itx_i}$. Celle-ci ayant une norme majorée par 1, tous les moments sont finis. $\widehat{\varphi}(t)$ est un estimateur convergent de la fonction caractéristique théorique. La norme de la fonction caractéristique prenant la forme : $|\varphi(t)| = \exp(-\sigma^\alpha |t|^\alpha)$ on obtient :

$\log |\varphi(t)| = -\sigma^\alpha |t|^\alpha$. Si $\alpha = 1$ on est conduit à résoudre un système de 2 équations qui se résolvent en prenant 2 valeurs de t particulières. Il est important alors de bien choisir ces valeurs là. Ce choix induit des résultats qui peuvent être aberrants. Les paramètres prennent la forme suivante :

$$\widehat{\alpha} = \frac{\log \left| \frac{\widehat{\varphi}(t_1)}{\widehat{\varphi}(t_2)} \right|}{\log \left| \frac{t_1}{t_2} \right|} \quad \text{et} \quad \log(\widehat{\sigma}) = \frac{\log \left| \widehat{\varphi}(t_1) \right|}{\log \left| \frac{\widehat{\varphi}(t_1)}{\widehat{\varphi}(t_2)} \right|}$$

Et de même avec 2 autres valeurs de t on obtient β et μ à partir de :

$u(t) = \text{Partie imaginaire de } \log(\varphi(t))$.

$$\widehat{\beta} = \frac{\frac{u(t_4)}{t_4} - \frac{u(t_3)}{t_3}}{\left[|t_4|^{\alpha-1} - |t_3|^{\alpha-1} \right] \sigma^\alpha \tan\left(\frac{\alpha\pi}{2}\right)} \quad \text{et} \quad \widehat{\mu} = \frac{|t_4|^{\alpha-1} \frac{u(t_3)}{t_3} - |t_3|^{\alpha-1} \frac{u(t_4)}{t_4}}{|t_4|^{\alpha-1} - |t_3|^{\alpha-1}}$$

Si $\alpha = 1$, seuls les autres paramètres sont à estimer de la même façon que précédemment.

Cette méthode appliquée à la série des temps d'attente donne des résultats très différents en fonction des valeurs particulières choisies. Ce manque de précision vient du fait que la série des temps d'attente n'est pas stationnaire alors que les séries tests donnent des valeurs proches de la valeur nominale de α .

Conclusion : Cette méthode conduit donc à des valeurs trop disparates pour être retenues bien que la méthode soit validée sur les simulations tests.

Méthode de Koutrouvelis [32] : Cette méthode porte aussi sur le comportement de la fonction caractéristique dont le log du carré du module, à une constante près, se comporte comme une puissance de la variable. La représentation en log-log doit alors, localement, approcher une droite dont la pente est une estimation de α . Les formules qui régissent cette méthode sont les suivantes :

$$\log?(-\log|\varphi(t)|^2) = \log(2\sigma^\alpha) + \alpha \log|t| \quad (\text{C.1})$$

et

$$\arctan\left(\frac{\operatorname{Im}\varphi(t)}{\operatorname{Re}\varphi(t)}\right) = \mu t + \beta\sigma^\alpha \tan\frac{\alpha\pi}{2} \operatorname{signe}(t)|t|^\alpha \quad (\text{C.2})$$

Pour l'équation (C.1) on peut procéder à une régression linéaire de deux façons différentes, l'une à partir de la variable t qui est la pulsation dans la fonction caractéristique avec $t = 1, 2, \dots, K$, l'autre, comme le propose Koutrouvelis, à partir de la variable $t_k = \frac{\pi t}{25}$ et K entre 9 et 134. La première façon donnant des résultats trop disparates, c'est la 2^{nde} qui est retenue. La régression prend la forme :

$$y_k = \log?(-\log|\varphi(t_k)|^2) = \log(2\sigma^\alpha) + \alpha \log|t_k| + \varepsilon_k$$

La fonction est donc représentée relativement à la variable $\log|t_k|$ et l'on choisit la partie la plus linéaire. L'ordonnée à l'origine $\log(2\sigma^\alpha)$ permet de calculer σ .

μ et β sont obtenus à partir de la 2^{nde} équation. Ces valeurs de μ et β dépendent trop du choix de t_k pour être recevables. Pour apprécier le résultat nous avons appliqué la méthode à des séries connues positives 0,4-stable et 0,9-stable. Le choix de la partie linéaire du graphe est faîte sur 20 simulations des séries connues.

Les estimations graphiques relatives à cette méthode sont effectuées respectivement pour la série des temps et celles des séries-types 0,4 et 0,9 α -stables. Les résultats sont consignés dans le Tableau C.2 où sont mis en parallèle l'estimateur de Koutrouvelis pour les simulations tests. Seul l'estimateur de la série 0,9-stable présente une erreur acceptable de 2%. Il est à noter qu'entre les deux méthodes de Koutrouvelis et McCULLOCH les intervalles de confiance ne se recoupent pas mais restent très voisins quand même.

Méthode de Hill L'estimateur de Hill [25], tend à mesurer l'épaisseur de la queue de distribution par la différence entre les statistiques d'ordre. Il prend la forme :

$$\alpha_{\text{Hill}}(k) = k \left[\sum_{i=1}^k \log \frac{x_{n+1-i}}{x_{(n-k)}} \right]^{-1}$$

Temps d'attente	Simulations 0,4-stables	Simulations 0,9-stables
$\alpha = 0.97$	$\alpha = 0.35$	$\alpha = 0.89$
$0.95 < \alpha < 0.99$	Erreur 12%	Erreur 2%

TABLE C.2 – Méthode de Koutrouvelis.

n étant le nombre de variables aléatoires indépendantes et identiquement distribuées de la distribution considérée et k le nombre de points choisis dans la queue de la série. K ne doit être ni trop grand pour ne pas disposer d'observations hors de la queue elle-même, ni trop petit pour disposer de suffisamment d'observations. Les calculs menés ci-après portent sur une queue constituée par le dernier décile de la série. Les éléments de la queue sont ordonnés de façon à ce que le logarithme du rapport soit positif donc que la série soit croissante. Par cette méthode, seule la série test 0,4-stable montre un résultat recevable, celui sur la série 0,9-stable présentant un écart de 50% entre l'index simulé et l'index mesuré. Ceci jette un doute sur l'index de la série des temps d'attente qui ne sera donc pas retenu. D'autre part, le graphe de l'estimateur des temps d'attente montre une croissance qui suppose une divergence lorsque la taille de la série augmente, ce qui n'est pas le cas de la série 0,4-stable positive.

Méthode de régression des queues : Cet estimateur [1] porte sur les queues des lois α -stables qui adoptent un comportement asymptotique en puissance α soit :

$$P(X > x) \sim C_\alpha (1 + \beta) \sigma^\alpha x^{-\alpha}$$

Ce qui prend la forme linéaire suivante dont ($-\alpha$) est la pente :

$$\log P(X > x) \sim \log [C_\alpha (1 + \beta) \sigma^\alpha] - \alpha \log(x)$$

La régression linéaire consistera alors à calculer les fréquences relatives d'une partie de k valeurs du comportement asymptotique ci-dessus données par la formule :

$$y_i = \frac{1}{n} \sum_{j=1}^k 1[X_j > x_i] \sim P(X > x_i).$$

Puis d'ajuster la régression linéaire :

$$\log y_i = constante - \alpha \log x_i.$$

nous adopterons trois façons d'aborder la régression linéaire des queues. Puisque pour de grandes valeurs de la variable leurs logs et ceux de la fonction de survie sont supposés suivre une fonction affine de coefficient directeur α , on peut :

1. Rechercher graphiquement la partie ou encore "la plage" du graphe de longueur k la plus linéaire possible puis ajuster une régression linéaire par la méthode des moindres carrés. La variable sera $\log(x_i)$ pour $1 < i < k$ et $y_i = \log P(X > x_i)$.
2. Reprendre l'ajustement ci-dessus en considérant que les écarts ($y_i - y_k$) restent proportionnels aux écarts ($x_i - x_k$). D'où $\alpha = \frac{\overline{y_i} - y_k}{\overline{x_i} - x_k}$.
3. Reprendre encore l'ajustement ci-dessus avec $k = n$ (longueur de la série) donc $y_k = 0$.

Pratiquement : Partant de la série des temps d'attente on détermine la fonction de répartition empirique puis la fonction de survie i-e la fonction donnant la probabilité pour que la variable soit supérieure à une grandeur donnée. Le choix des k valeurs de queue se fait alors sur la représentation de cette fonction de survie en log-log en prenant la partie la plus en fin de courbe possible et la plus linéaire possible. C'est-à-dire la partie droite la plus régulière possible.

Pour le point 1 ci-dessus : l'estimateur prend la forme :

$$\alpha = - \frac{\sum_{i=1}^k (\log x_i \log y_i) - \overline{\log x_i} \overline{\log y_i}}{\sum_{i=1}^k (\log x_i)^2 - \overline{\log x_i}^2}$$

Cette méthode appliquée à la série des temps d'attente pour une sélection de la queue entre 70 et 95 % de l'ensemble des valeurs donne un index de stabilité α compris entre 0,91 et 0,95. Pour évaluer l'influence du nombre des observations sur le résultat, la même méthode a été testée sur des nombres d'observations bien moins importants et qui n'offrent pas de longues tendances linéaires. Les index s'échelonnent entre 0.7 et 0.8, ce qui est assez loin des résultats obtenus par les autres méthodes. Testée sur des simulations 0,9-stables, la méthode donne, en général, des valeurs entre 0.86 et 0.95, soit une marge d'erreur de 5%. L'étude ne peut porter sur un grand nombre de simulations car pour chacune d'elle il faut choisir graphiquement la partie adéquate de la représentation, ce qui est très variable d'une simulation à l'autre. Sauf, bien sûr, à automatiser le calcul en imposant une "tranche de courbe" identique pour toutes les simulations ce qui n'est pas recevable.

Pour le point 2 : l'estimateur prend la forme : $\alpha = \frac{(1/k) \sum_{i=1}^k (\log y_i) - y_k}{(1/k) \sum_{i=1}^k (\log x_i) - x_k}$

Pour le point 3 : l'estimateur prend la forme : $\alpha = \frac{\sum_{i=1}^n (\log y_i)}{\sum_{i=1}^n (\log x_i) - x_n}$

Sur la "plage 20% – 60%" les deux premiers estimateurs donnent le même résultat pour l'index des temps d'attente alors que le 3ième estimateur donne un résultat aberrant vu que la partie finale de la queue n'est pas linéaire.

Annexe D

Processus stochastique

Définition

Un processus stochastique peut être défini simplement comme une famille indéxée de variable aléatoire X_t où l'indexe t est paramètre appartenant à un ensemble I lui même pouvant être un sous-ensemble de R . Par exemple si :

1) $I = (-\infty, \infty)$ or $I = [0, \infty]$. Dans ce cas X_t (qu'on notera aussi $X(t)$) est un processus stochastique continu du temps t.

2) $I = \{0, \pm 1, \pm 2, \dots\}$ or $I = \{0, 1, 2, \dots\}$. Ici, X_t est un processus temporel discret.

Pour faire la différence, nous notons $\{X_t\}$ le processus stochastique et $\{x_t\}$ sa réalisation ou encore son échantillon. Dans ce travail de thèse nous avons utilisé la réalisation temporelle d'un processus pour le caractériser. De manière exacte, rappelons qu'un processus stochastique est défini par la donnée de sa distribution de probabilité marginale. La distribution de probabilité marginale d'ordre k d'un processus stochastique est donnée par

$$F_{X_{t_1}, \dots, X_{t_k}}(x_1, \dots, x_k) = P(X_{t_1} \leq x_1, \dots, X_{t_k} \leq x_k) \quad (\text{D.1})$$

avec $t_1, \dots, t_k \in I$ et x_1, \dots, x_k des nombres réels. La connaissance de cette distribution de probabilité pour tout ordre k définit complètement le processus $\{X_t\}$. Dans la pratique cette condition n'est généralement pas satisfaite. Ceci est en particulier vrai quand on veut caractériser un processus stochastique en partant d'observation expérimentale.

Moments d'un processus stochastique

Il est possible de caractériser un processus stochastique à partir de ses moments statistiques, $E(X_t)$, $E(X_t^2)$, $E(X_t \cdot X_s)$ etc.. Dans la pratique compte tenu des erreurs d'estimation qu'on peut rencontrer dans les estimations de ces moments, on s'arrete généralement aux deux premiers ordres. Notons $Var(X_t) = \sigma_t^2$ et $E(X_t) = \mu_t$ respectivement les fonctions variance

et moyenne du processus stochastique. On définit alors la fonction de covariance du processus entre deux instants X_t , X_s par

$$\text{Cov}(X_t, X_s) = E((X_t - \mu_t)(X_s - \mu_s)) \quad (\text{D.2})$$

De même, il est possible de définir la fonction d'autocorrélation comme

$$\rho(X_t, X_s) = \frac{\text{Cov}(X_t, X_s)}{\sqrt{\sigma_t^2} \sqrt{\sigma_s^2}} \quad (\text{D.3})$$

Elle décrit le lien statistique entre les deux variables aléatoires X_t , X_s .

Stationarité

Un processus stochastique (avec $E(X_t^2) < \infty$) est dit stationnaire si pour tout ordre k , $(X_{t_1}, \dots, X_{t_k})$ et $(X_{t_1+\Delta}, \dots, X_{t_k+\Delta})$ possèdent la même distribution de probabilité donnée par (D.1) quelque soit l'incrément temporel Δ . En d'autres termes :

$$F_{X_{t_1}, \dots, X_{t_k}}(x_1, \dots, x_k) = F_{X_{t_1+\Delta}, \dots, X_{t_k+\Delta}}(x_1, \dots, x_k)$$

Dans ce cas, le processus $\{X_t\}$ admet une moyenne et une variance constantes. De plus les fonctions de covariance et d'auto-correlation (D.2 et D.3) dépendent uniquement de l'écart s . En pratique, la condition de stationnarité reste très difficile à prouver en partant d'une réalisation x_1, \dots, x_k . A la place, on fait plus souvent appel à la stationnarité faible qui ne concerne que les moments du processus stochastique. Un processus est dit stationnaire d'ordre n si tous ses moments jusqu'à cet ordre existent et sont invariants par un changement d'origine du temps.

Estimation des moments

Dans le cas où nous ne possédons qu'une réalisation $\{x_t\}$ du processus $\{X_t\}$, les estimateurs naturels de ses moments statistiques sont données par

$$\bar{x} = \frac{1}{n} \sum_{t=1}^n x_t \quad (\text{D.4})$$

$$\hat{\gamma}(k) = \frac{1}{n} \sum_{t=1}^{n-k} (x_t - \bar{x})(x_{t+k} - \bar{x}) \quad (\text{D.5})$$

$$\hat{\rho}(k) = \frac{\sum_{t=1}^{n-k} (x_t - \bar{x})(x_{t+k} - \bar{x})}{\sum_{t=1}^n (x_t - \bar{x})^2} = \frac{\hat{\gamma}(k)}{\hat{\gamma}(0)} \quad (\text{D.6})$$

Les équations (D.4, D.5 et D.6) donnent des estimateurs respectivement de la moyenne, de la fonction de covariance et de la fonction de corrélation. D'un point de vue plus théorique, la validité de ces estimateurs repose sur les hypothèses d'ergodicité et de stationnarité annoncées.

au premier chapitre de ce mémoire. Sous ces hypothèses, ces estimateurs ont été utilisés pour analyser notre série temporelle de pluie qui représente une réalisation du processus stochastique recherché.

Le mouvement brownien

Un processus stochastique B_t qui est un mouvement brownien peut être défini par son incrément qu'on appelle aussi processus de Wiener, s'écrit

$$dB_t = \epsilon \sqrt{dt} \quad (\text{D.7})$$

où ϵ est une variable aléatoire de loi normale centrée réduite et t représente le temps. De plus le processus $dB(t)$ est de moyenne nulle et sa variance vaut dt .

Processus d'Itô

Les calculs stochastiques d'Itô permettent de définir une famille de processus stochastiques dits processus d'Itô à partir du mouvement brownien. La formule d'Itô a été démontrée pour la première fois par le mathématicien japonais Kiyoshi Itô dans les années 1940. Un processus X_t est un processus d'Itô s'il est solution de l'équation différentielle stochastique suivante

$$dX_t = \mu_t dt + \sigma_t dB_t \quad (\text{D.8})$$

En d'autres termes, les processus d'Itô possèdent des moyennes linéaires en temps $E(X_t) = \mu_t t$ et une variance de la forme $E(X_t^2) = \sigma_t^2 t$.

Mouvement brownien géométrique

La formule d'Itô (D.8) se généralisent par l'introduction d'une fonction de transformation f sous la forme

$$d(f(X_t, t)) = \frac{\partial f}{\partial t}(X_t, t)dt + \frac{\partial f}{\partial x}(X_t, t)dX_t + \frac{1}{2} \frac{\partial^2 f}{\partial x^2}(X_t, t)\sigma_t^2 dt \quad (\text{D.9})$$

En posant $f(X_t, t) = \ln X_t$, on obtient grâce à la formule d'Itô l'expression

$$\begin{aligned} d(\ln X_t) &= 0dt + \frac{1}{X_t}dX_t + \frac{1}{2} \left(-\frac{1}{X_t^2} \right) (\sigma X_t)^2 dt \\ &= \frac{1}{X_t}(\mu X_t dt + \sigma X_t dB_t) - \frac{1}{2}\sigma^2 dt \\ &= \left(\mu - \frac{1}{2}\sigma^2 \right) dt + \sigma dB_t \end{aligned} \quad (\text{D.10})$$

On peut alors intégrer et il en découle

$$X_t = X_0 \exp \left(\sigma B_t + \mu t - \frac{1}{2} \sigma^2 t \right) \quad (\text{D.11})$$

où X_0 est une valeur initiale aléatoire. L'équation (D.11) montre que le processus X_t suit une loi géométrique compte tenu de la fonction exponentielle. Par ailleurs, on voit aussi que le mouvement brownien géométrique est lié naturellement à la loi de distribution de probabilité log-normale à cause du caractère gaussien du processus B_t . Le mouvement brownien géométrique porte aussi le nom de modèle de Black-Scholes et constitue un modèle basique pour étudier les fluctuations de valeurs financières. Son extension au mouvement brownien géométrique fractionnaire est utilisé dans le chapitre trois de ce mémoire en lien avec l'équation fractionnaire de Fokker Planck (FFPE).

Annexe E

Données et mesures expérimentales

Les données employées dans cette étude ont été obtenues sur le site expérimental "Fontanille" appartenant à l'Inra. Les données sont constituées de mesures horaires, locales et synchronisées du niveau de la nappe phréatique et des précipitations pluvieuses. Ces mesures correspondent à la période du 4/02/2005 au 27/03/2006.

Le site atelier "Fontanille" est un site expérimental visant à une meilleure prédiction de la pollution des nappes souterraines à l'échelle de la parcelle agricole [13].

La géologie du site est caractéristique d'une plaine alluviale, d'une dizaine de mètres d'épaisseur avec un substratum marneux. La parcelle présente une très légère pente, constante, avec un point haut au SE et un point bas au NO. Cependant, aucun ruissellement n'a été constaté depuis la création du site.

En surface, on trouve un sol argilo-limoneux qui fissure facilement sous les effets de la sécheresse, de sorte à former des fentes de dessiccation. La nappe sous-jacente est appelée nappe alluviale de la basse Durance. On est en présence d'une nappe libre qui oscille entre 4 et 6 m de hauteur. Au niveau du site expérimental, l'écoulement local suit l'écoulement général de la nappe, du SE vers le NO, avec un gradient hydraulique qui varie entre 1,4 % et 3 %.

Un dispositif spécifique constitué d'une géomembrane entourant le site jusqu'à 2m de profondeur a été mis en place pour limiter les apports latéraux venant de l'extérieur de la parcelle.

Ce site atelier est entre autre équipé de plusieurs piézomètres munis de sondes piézométriques de type Diver, de marque Schlumberger Water Services, qui permettent de mesurer et d'enregistrer de manière autonome et continue le niveau et la température de la nappe phréatique. La précision des mesures obtenues est 1mm.

A proximité du site, en appui aux expériences menées in situ, les données climatiques sont

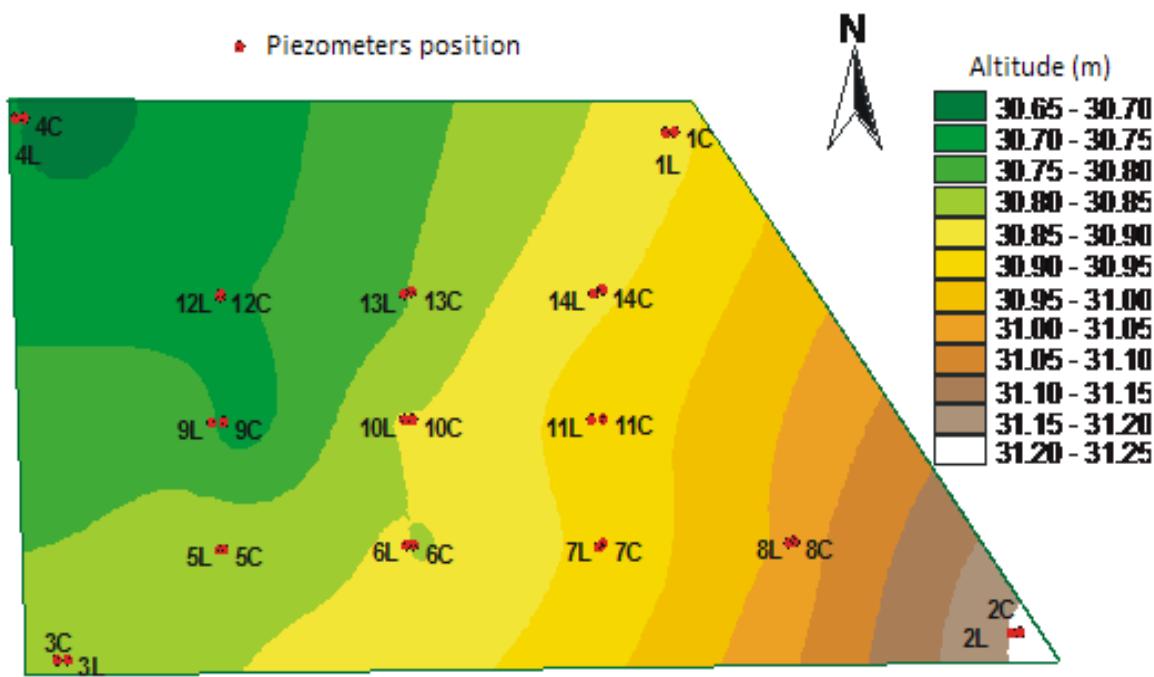


FIGURE E.1 – Schéma de l'instrumentation du site atelier "Fontanille". Les données issues du site 3C ont été utilisées dans ce mémoire. Elles sont équivalentes à celles obtenues sur 1C et 2L.

recueillies par une station agrométéorologique. Les données pluviométriques nécessaires à notre étude ont été acquises par la station agrométéorologique de Montfavet (n°87007004) gérée par l'unité expérimentale agroclim de l'INRA. Toutes les stations météorologiques AgroClim mesurent la pluviométrie à l'aide d'un dispositif normalisé (également utilisé par Météo France), et régulièrement étalonné : le pluviomètre à auget basculeur située à une centaine de mètre du site atelier "Fontanille". La résolution du pluviomètre à auget est de 0,5 mm.

Annexe F

Code de calcul PASTIS

Résolution par différences finies de l'équation (4.10)

On note $\{x_i; i = 1, \dots, N\}$ les coordonnées des noeuds du maillage. On définit les pas d'espace par $\{dx_i = x_{i+1} - x_i; i = 1, \dots, N-1\}$. On définit aussi des pas d'espaces centrés sur les noeuds par $\{dxi_i = (dx_{i+1} + dx_i)/2; i = 2, \dots, N-1\}$ avec $dxi_1 = dx_1$ et $dxi_N = dx_{N-1}$.

La discrétisation des dérivés spatiales se fait selon le schéma suivant qui permet d'avoir une approximation quadratique. En un noeud i du maillage on a :

$$\left[\frac{\partial}{\partial z} \left(K(\psi) \frac{\partial H}{\partial z} \right) \right]_i \approx \frac{1}{dxi_i} \left[K_{i+\frac{1}{2}} \frac{H_{i+1} - H_i}{dx_i} - K_{i-\frac{1}{2}} \frac{H_i - H_{i-1}}{dx_{i-1}} \right] \quad (\text{F.1})$$

où $K_{i-\frac{1}{2}}$ et $K_{i+\frac{1}{2}}$ sont les conductivités internodales qui peuvent être calculées par diverses moyennes (arithmétique, géométrique, etc...)

La moyenne arithmétique et la moyenne géométrique peuvent être utilisée dans le code et sont respectivement définies par les équations suivantes :

$$K_{i+\frac{1}{2}} = \frac{K_i + K_{i+1}}{2} \quad (\text{F.2})$$

$$K_{i-\frac{1}{2}} = \frac{K_i + K_{i-1}}{2} \quad (\text{F.3})$$

$$K_{i+\frac{1}{2}} = \sqrt{K_i K_{i+1}} \quad (\text{F.4})$$

$$K_{i-\frac{1}{2}} = \sqrt{K_i K_{i-1}} \quad (\text{F.5})$$

On obtient donc en chaque noeud i une équation différentielle du temps qui dans le cas de la formulation en potentiel s'écrit :

$$\begin{aligned}
C_i \frac{dh_i}{dt} \approx & h_{i-1} \frac{K_{i-\frac{1}{2}}}{dx_{i-1} dx_i} \\
& - h_i \left[\frac{K_{i-\frac{1}{2}}}{dx_{i-1} dx_i} + \frac{K_{i+\frac{1}{2}}}{dx_i dx_i} \right] \\
& + h_{i+1} \frac{K_{i+\frac{1}{2}}}{dx_i dx_i} + S(\psi_i, f(x_i))
\end{aligned} \tag{F.6}$$

On a donc un système de N équations différentielles couplées. Ces équations sont parfaitement définies pour tous les noeuds internes au domaine d'intégration ($i = 2, \dots, N - 1$). La discrétisation des conditions aux limites permet de traiter les équations pour les noeuds aux frontières $i = 1$ et $i = N$. Le traitement fait habituellement intervenir des noeuds fictifs $i = 0$ et $i = N + 1$ extérieurs au domaine. Le noeud $i = 0$ est symétrique du noeud $i = 2$ par rapport au noeud frontière $i = 1$ et le noeud $i = N + 1$ est symétrique du noeud $i = N - 1$ par rapport au noeud frontière $i = N$.

Traitement des Conditions aux Limites

Condition de gradient de potentiel matriciel nul à la base du profil

C'est la condition qui est utilisée quand on ne sait rien sur ce qui se passe à la limite inférieure du profil. La condition de gradient de potentiel nul $d\psi/dx = 0$ conduit donc à la condition $dH/dx = -1$. Elle traduit donc que le flux d'eau à la base du profil de sol se fait uniquement sous l'influence de la gravité avec un gradient unitaire. Le flux d'eau est donc égal à la conductivité hydraulique du sol. Cette condition à la limite est donc bien particulière et doit être réservée à des cas pour lesquels on est sûr que l'écoulement est toujours orienté vers le bas du profil ; au moins pour ce qui est de la partie inférieure du profil. Cette condition à la limite se traduit par :

$$H_{N+1} = H_{N-1} - 2dx_{N-1} \tag{F.7}$$

Elle correspond à l'équation précédente dans laquelle on a supposé l'égalité entre K_N et F_b .

On combine l'expression ci-dessus avec la dernière équation différentielle du système pour éliminer H_{N+1} . Il vient :

$$\begin{aligned} C_N \frac{dH_N}{dt} &\approx h_{N-1} \left[\frac{K_{N-\frac{1}{2}}}{dx_{N-1} dx i_N} + \frac{K_{N+\frac{1}{2}}}{dx_{N-1} dx i_N} \right] \\ &- h_N \left[\frac{K_{N-\frac{1}{2}}}{dx_{N-1} dx i_N} + \frac{K_{N+\frac{1}{2}}}{dx_{N-1} dx i_N} \right] \\ &- \frac{2K_{N+\frac{1}{2}}}{dx i_N} + S(\psi_N, f(x_N)) \end{aligned}$$

Condition de Neumann à la surface

En utilisant un noeud fictif hors du domaine, la discréétisation de la condition à la surface (eq. 4.11) donne la relation suivante

$$H_0 = H_2 + 2dx_1 F_s(t)/K_1 \quad (\text{F.8})$$

où H_0 est la pression au noeud fictif 0 symétrique du noeud 2 par rapport à la surface du sol. On combine l'expression ci-dessus avec la première équation différentielle du système pour éliminer H_0 . Il vient :

$$\begin{aligned} C_1 \frac{dH_1}{dt} &\approx -h_1 \left[\frac{K_{\frac{3}{2}}}{dx_1 dx i_1} + \frac{K_{\frac{1}{2}}}{dx_1 dx i_1} \right] \\ &+ h_2 \left[\frac{K_{\frac{3}{2}}}{dx_1 dx i_1} + \frac{K_{\frac{1}{2}}}{dx_1 dx i_1} \right] \\ &+ \frac{2K_{\frac{1}{2}} F_s(t)}{dx i_1 K_1} + S(\psi_1, f(x_1)) \end{aligned}$$

Pour améliorer les résultats, le code propose quand le flux $F_s(t)$ est négatif d'utiliser directement une loi de conservation de la masse appliquée à la "cellule" $[0, dx_1/2]$. On a donc :

$$C_1 \frac{dH_1}{dt} \approx -\frac{2}{dx_1} \left(-K_{\frac{3}{2}} \frac{H_2 - H_1}{dx_1} - F_s(t) \right) \quad (\text{F.9})$$

l'équation différentielle au premier noeud devient donc :

$$\begin{aligned} C_1 \frac{dH_1}{dt} &\approx -H_1 \frac{2K_{\frac{3}{2}}}{dx_1 dx_1} + H_2 \frac{2K_{\frac{3}{2}}}{dx_1 dx_1} \\ &+ \frac{2}{dx_1} F_s(t) + S(\psi_1, f(x_1)) \end{aligned} \quad (\text{F.10})$$

Intégration Temporelle

Formulation en pression

Une fois les conditions aux limites incluses, on a donc un système tridiagonal d'équations différentielles non-linéaires. Ce système est de dimension N avec les conditions de flux aux deux extrémités. Il peut se mettre sous la forme

$$[C] \frac{dh}{dt} = -[K]h + S \quad (\text{F.11})$$

où $[C]$ est la matrice diagonale des capacités capillaires, et $[K]$ est la matrice tridiagonale dite de *rigidité*. S est le vecteur des sollicitations. Il contient les termes puits et les termes provenant des conditions aux limites.

L'intégration en temps se fait avec un schème stable à deux niveaux

$$\begin{aligned} [C]^{j+1} \frac{h^{j+1} - h^j}{dt} &= \theta \left(-[K]^{j+1} h^{j+1} + S^{j+1} \right) \\ &\quad + (1 - \theta) \left(-[K]^j h^j + S^j \right) \end{aligned} \quad (\text{F.12})$$

Les indices j indiquent la solution connue au temps t et les indices $j + 1$ indiquent la solution cherchée au temps $t + dt$. Si $\theta \geq \frac{1}{2}$ le schème est stable. En prenant $\theta = 1$ on a un schème implicite pur qui est inconditionnellement stable. En prenant $\theta = \frac{1}{2}$ on a un schème de *Crank-Nicholson*.

Le schème ci-dessus conduit au système non-linéaire de solution H^{j+1} :

$$\begin{aligned} \left([C]^{j+1} + \theta dt [K]^{j+1} \right) H^{j+1} &= \\ \left([C]^{j+1} - (1 - \theta) dt [K]^j \right) H^j & \\ + \theta dt S^{j+1} + (1 - \theta) dt S^j & \end{aligned} \quad (\text{F.13})$$

La non-linéarité provient du fait que les matrices $[K]$ et $[C]$ et le vecteur S dépendent implicitement de la solution recherchée. Ce système doit donc être résolu par une méthode itérative. Une méthode des approximations successives, souvent appelée *schème de Picard* est mise en oeuvre dans le code. La méthode consiste à construire une suite d'approximations qui converge vers la solution. Pour ce faire, il est commode d'introduire le résidu $\omega_k = H_{k+1}^{j+1} - H_k^{j+1}$ de la suite (H_k^{j+1}) qui converge lors de l'itération vers la solution H^{j+1} au temps $t + dt$. Le

premier terme de cette suite est donné par la solution H^j au temps t . On obtient alors le système :

$$\begin{aligned} & \left([C]_k^{j+1} + \theta dt [K]_k^{j+1} \right) \omega_k^{j+1} = \\ & - \left([C]_k^{j+1} + \theta dt [K]_k^{j+1} \right) H_k^{j+1} \\ & + \left([C]_k^{j+1} - (1 - \theta) dt [K]^j \right) H^j \\ & + dt [\theta S_k^{j+1} + (1 - \theta) S^j] \end{aligned}$$

Si on note :

$$\begin{aligned} [A]_k^{j+1} &= [C]_k^{j+1} + \theta dt [K]_k^{j+1} \\ SML^j &= (1 - \theta) dt \left(-[K]^j H^j + S^j \right) \end{aligned}$$

On obtient le système linéaire suivant d'inconnue ω_k^{j+1} :

$$\begin{aligned} [A]_k^{j+1} \omega_k^{j+1} &= -[A]_k^{j+1} H_k^{j+1} + [C]_k^{j+1} H_k^j \\ &+ SML^j + dt \theta S_k^{j+1} \end{aligned} \tag{F.14}$$

La résolution de ce système linéaire est faite par un algorithme de THOMAS. Le processus itératif s'arrête quand un critère de convergence sur ω est atteint. Le critère utilisé est un critère relatif :

$$\frac{\|\omega\|_\infty}{\|H_k^{j+1}\|_\infty} \leq \epsilon \tag{F.15}$$

Dans le code, le critère ϵ est spécifié dans un fichier de données. Dans le code numérique, on peut utiliser des critères différents selon que l'on est en phase d'infiltration (phénomène brutal) ou en phase de redistribution ou évaporation (phénomènes plus doux). La stratégie de gestion du pas de temps dans le code est très liée à la vitesse de convergence des itérations ci-dessus. Classiquement, si on part d'une estimation *raisonnable* de H^{j+1} , et H^j en est une si dt n'est pas trop grand, on doit avoir convergence en 2 à 3 itérations. Si tel n'est pas le cas, on peut penser que le pas de temps est trop grand pour l'intensité des phénomènes qui se déroulent. Il faut bien garder à l'esprit que la notion de taille du pas de temps n'est pas indépendante de la

- Si la convergence est atteinte en 3 itérations on ne change pas le dt
- Si la convergence n'est pas atteinte en 3 itérations, on réduit le pas de temps (division par 2) et on reprend la résolution sur cette base.
- Si la convergence est atteinte avec au plus 2 itérations, on augmente le pas de temps (multiplications par 1.05). Celui-ci sera utilisé pour le pas de temps suivant.

Annexe G

Transformation en ondelette

Définition

Les ondelettes ont vu le jour lorsque certains sujets d'étude ont nécessité une analyse en fréquence et en temps. En effet, l'apparition dans un signal d'une propriété localisée en temps et de manière transitoire reste assez mal reproduite par la transformée de Fourier qui ne permet qu'une analyse en fréquence. Ce qui est typiquement le cas des pics d'événements de pluies et de fluctuations de nappe observés dans notre cas. En d'autres termes, la transformation en ondelette est une technique qui permet de faire une analyse fréquentielle multi-échelle en temps c'est à dire qu'elle permet la décomposition en fréquence des signaux par niveau d'échelle de temps.

Il existe deux familles d'ondelettes : les ondelettes continues et les ondelettes discrètes. Le choix de l'ondelette la plus adaptée pour l'analyse d'un jeu de données dépend principalement de la nature même des données. La transformée en ondelettes discrètes est préférée lorsqu'un petit nombre de données est présent en entrée. Généralement, l'ondelette discrète est plus adaptée pour la compression de données et les modèles de prédiction. Alors que les ondelettes continues sont plus adaptées lorsqu'on a besoin de bonnes résolutions temporelles, à toutes échelles, et pour obtenir des informations sur la phase [45]. Cela vient du fait que les ondelettes continues possèdent des propriétés de similarité lorsqu'une translation et / ou une contraction leurs sont appliquées. Dans cette étude, nous n'utiliserons que la transformée en ondelettes continues (en anglais CWT pour Continuous Wavelet Transform).

De manière simple et sans se préoccuper des aspects mathématiques de la théorie des ondelettes, nous pouvons dire pour qu'une fonction ζ soit une base de décomposition en ondelettes (ou tout simplement une ondelette) elle doit avoir une énergie finie et être de moyenne nulle.

$$\begin{cases} E = \int_{-\infty}^{\infty} |\zeta|^2 dt < \infty \\ C_g = \int_0^{\infty} \frac{|\hat{\zeta}(f)|^2}{f} df < \infty \end{cases} \quad (\text{G.1})$$

avec $\hat{\zeta}(f)$ la transformée de Fourier de $\zeta(t)$. La quantité C_g constitue un facteur de normalisation dont nous verrons ci-dessous le rôle. La transformée en ondelettes consiste donc à décomposer un signal en une famille de fonctions localisées en temps et en fréquence.

Compte tenu de la nature de nos signaux nous avons utilisé les ondelettes de Morlet. Ce type d'ondelettes est bien connu car elles sont très utilisées pour l'analyse de jeux de données présentant des oscillations et des fluctuations, quand on recherche à localiser temporellement des structures isolées dans un environnement bruité. La fonction ondelette de Morlet consiste en une onde plane modulée par une cloche Gaussienne. L'ondelette mère de Morlet s'écrit

$$\zeta_0(t) = \pi^{-\frac{1}{4}} \exp(-t^2/2) \exp(2\pi f_0 t i) \quad (\text{G.2})$$

où f_0 est la fréquence centrale de l'ondelette mère. Pour des raisons liées à la condition d'admissibilité (G.1) et au fait que les ondelettes de Morlet ne sont pas orthogonales, nous prenons la fréquence $f_0 = \frac{1}{\sqrt{2 \log(2)}}$ Hz.

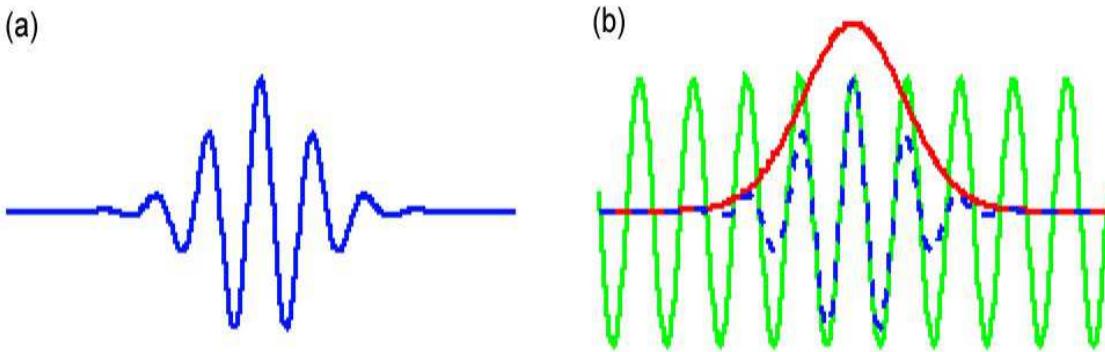


FIGURE G.1 – Ondelette de Morlet - (a) Forme de sa partie réelle (b) Ondelette de Morlet (en pointillé bleu), résultante d'une cloche Gaussienne (en rouge) et d'une onde plane (en vert)

Une famille d'ondelettes est ensuite construite en dilatant (ou contractant) et en translatant l'ondelette-mère.

$$\zeta_{s,\tau}(t) = \zeta_0\left(\frac{t-\tau}{s}\right) \quad (\text{G.3})$$

où τ est le facteur de translation et "s" le facteur de dilatation. Analyser une fonction de carré sommable en ondelettes consiste ainsi à calculer l'ensemble de ses produits scalaires avec les ondelettes de la famille. Les nombres obtenus sont appelés coefficients d'ondelettes, et l'opération associant à une fonction ses coefficients d'ondelettes définit la transformée en ondelettes.

Mathématiquement, on définit la transformée en ondelette continue d'une fonction $f \in L^2(\mathbb{R})$ par

$$g(s, \tau) = \int_{-\infty}^{\infty} f(t) \zeta_{s,\tau}(t)^* dt \quad (\text{G.4})$$

où $\zeta_{s,\tau}$ est une ondelette de la famille d'ondelettes, le symbole "*" désigne le complexe conjugué. Pour retrouver le signal "f" d'origine on utilise la transformée en ondelette continue inverse donnée par

$$f(t) = \frac{1}{C_g} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{|s|^2} g(s, \tau) \zeta_{s,\tau}(t) ds d\tau \quad (\text{G.5})$$

Bien évidemment, pour l'analyse d'un signal discret il y a lieu de remplacer l'intégrale dans (G.4) par une somme discrète.

Bibliographie

- [1] A. Alvarez et P. Olivares. *Méthode d'estimation pour des lois alpha-stable avec des applications en finance*. Journal de la Société Française de Statistique, Tome 146, Num. 4. 2005
- [2] S.Asmussen. *Applied probability and queues*. Ed. Berlin-Springer. 2003.
- [3] E. Barkai, R. Metzler and J. Klafter. *From continuous time random walks to the fractional Fokker-Planck equation*. Physical Review E, Volume 61, Num. 1, pp. 132-138. DOI : 10.1103/PhysRevE.61.132 2000
- [4] A. Blanc-Lapierre *Modèles statistiques pour l'étude de phénomènes de fluctuations*, P. 148. Ed. Masson et Cie. 1963.
- [5] F.C. Blondeau, *Autocorrelation versus entropy-based auto information for measuring dependence in random signal*. Physica A 380, 1-18. 2007
- [6] P. Cannavo, F. Lafolie , B. Nicolardot and P. Renault. *Modelling seasonal variations in CO₂ and N₂O concentrations with a model describing C and N behaviour in the vadose zone*. Vadoze Zone Journal. Vol. 5. P. 990-1004. 2006
- [7] A. Chakrabarty and M.M. Meerschaert. *Tempered stable laws as random walk limits*. Statistique and Probability letters. Volume 81, Num. 8, pp. 989-997. DOI : 10.1016/j.spl.2011.01.019. 2011
- [8] J.M. Chambers, C.L. Mallows and B.W. Stuck *A method for simulating stable random variables*. Journal of American Statistical Association. Vol.71. Num. 354. P. 340-344. 1976
- [9] H. Cramer and M.R. Leadbetter. *Stationary and Related Processes* P. 348, Ed. John Wiley and Sons. 1966
- [10] J. De Reffye. *Etude phénoménologique des précipitations pluvieuses. Modélisation mathématique des intensités de pluie en un point du sol*. Revue de Statistique Appliquée. Tome 30, Num.3, P. 39-63. 1982

- [11] M.I.P. De Lima and J. Grassman *Multifractal analysis of 15-minute and daily rainfall from a semi-arid region of Portugal.* Journal of hydrology. Vol.220. P. 1-11. 1999
- [12] J.E. Dickinson, R.T. Hanson, T.P.A. Ferr and S.A. Leake. *Inferring time-varying recharge from analysis of long-term water level* Water Resources Research, Volume 40 W07403. Doi : 10.1029/2003WR002650. 2004
- [13] L. Di Pietro, C. Doussan, S. Ruy, J.C. Gaudu, R. Pallut, D. Renard. *Site ateliers "Fontanille", infiltrations, recharges, écoulements préférentiels et transport de polluants vers la nappe.* Rapport Technique, Centre de Recherche INRA PACA. 2004
- [14] J. Feder *Fractals.* New York : Plenum Press, 1988.
- [15] W. Feller. *An introduction to probability theory and its applications. Vol. II.* Ed. John Wiley and Sons, New York. 1966.
- [16] A. Findeling, P. Garnier, F. Coppens, F. Lafolie and S. Recous. *Modelling water, carbon and nitrogen dynamics in soils covered with decomposing mulch.* European J. Soil Sci. Vol. 58. P. 196-206. 2007
- [17] A.M. Fraser and H.L. Swinney. *Independent coordinates for strange attractors from mutual information.* Physic rev.A, 33,1134, 1986
- [18] J. Gajda and M. Magdziarz *Fractional Fokker-Plank equation with tempered alpha-stable waiting times : Langevin picture and computer simulation.* Physical Review E, Volume 82, 011117. 2010
- [19] E. Gaume, N. Mouhous and H. Andrieu. *Rainfall stochastic disaggregation models : Calibration and validation of a multiplicative cascade model.* Advances in Water Resources, Volume 30, Issue 5, p. 1301-1319. 2007
- [20] B.V. Gnedenko and A.N. Kolmogorov. *Limit distributions for sums of independent variables.* Ed. Addison-Wesley. 1968
- [21] J. Golder, M. Joelson, M.C. Neel. *Mass transport with sorption in porous media.* Journal Mathematics and Computers in Simulation, Volume 81, Issue 10, pp. 2181-2189. 2010.
- [22] J. Golder, M. Joelson, M.C. Neel. and L. Di Pietro *A time fractional model to represent rainfall process* article en révision au journal Water Science and Engineering. 2013

- [23] J. Golder, M. Joelson, M.C. Neel. and L. Di Pietro *Non-Gaussian Stable distribution of groundwater level fluctuations* article en cours de soumission au journal Chaos Solitons and Fractals. 2013
- [24] P. Gordon. *Théorie des chaîne de Markov finies et ses applications*. Ed. Dunod. P. 132. 1965
- [25] B. Hill. *A simple approach to inference about the tail of a distribution*. Annals of Statistics. Vol. 3. P. 1163-1174. 1975.
- [26] J.R.M. Hosking and R.T. Clarke *Rainfall-Runoff relations derived from the probability theory of storage*. Water Ressources Research, Volume 26, Issue 7, pp. 1455-1463. 1990
- [27] H.E. Hurst. *The long-term storage capacity of reservoirs*. Transactions of the American Society of Civil Engineers. Vol.116, pp. 770-799. 1951
- [28] M. Joelson, D. Dewit, P. Dussouliez, A. Ramamonjiarisoa. *Searching for chaotic and deterministic features in laboratory surface waves*. Nonlinear process in géophysics. Volume 7, pp 37-48. 2000
- [29] M. Joelson 1, M.-C. Néel *On alpha stable distribution of wind driven water surface wave slope*. Chaos : An Interdisciplinary Journal of Nonlinear Science 18, 3 , 033117-1. 2008
- [30] J. P. Kahane. *Le mouvement brownien - Un essai sur les origines de la théorie mathématique* Séminaires et Congrès de la SMF. Vol 3., p. 123-155. 1998
- [31] L. Kleinrock. *Queuing systems. Volume1 : theory*. Ed. Wiley. New-York. 1975
- [32] I.A. Koutrouvelis. *Regression-Type Estimation of the Parameters of Stable Laws*, Journal of The American Statistical Association 75, 918-928. 1980
- [33] D. Koutsoyiannis, A. Paschalis, N. Theodoratos *Two-dimensional Hurst-Kolmogorov process and its application to rainfall fields*. 2011
- [34] P. Kundu and T.L. Bell *Space-time scaling behavior of rain statistics in a stochastic fractional diffusion model*. Journal of Hydrology. Vol.322. P. 49-58. 2006
- [35] P. Kundu and J. Travis. *A Stochastic Fractional Dynamics Model of Rainfall Statistics*. Geophysical Research Abstracts, Vol. 15, EGU 2013-1491, 2013.
- [36] F. Lalbat, J. Golder, M. Joelson and L. Bruckler *Modélisation stochastique simplifiée du transfert de pesticide dans le sol*. Bulletin du GFHN. Num. 54. P. 119-124. 2008

- [37] F. Lafolie. *Modeling water flow, nitrogen transport and root uptake including physical non-equilibrium and optimization of the root water potential.* Fertilizer Research, Vol. 27. P. 215-231. 1991
- [38] M. Leroy, *Estimation de l'incertitude de mesure des précipitations.* Note interne DSO N 42, Météo France. 2000
- [39] P. Lévy. 1925, *Calcul des probabilités*, Paris, Ed. Gauthier Villars.
- [40] Z. Li, Y- K. Zhang, *Quantifying fractal dynamics of groundwater systems with detrended fluctuation analysis.* Journal of Hydrology, Volume 336, Issue 1-2, pp 139-146. 2007
- [41] A.W. Lo, *Long-Term Memory in Stock Market Prices.* Econometrica, Vol.59. Num. 5. P. 1279-1313. 1991
- [42] S. Lovejoy, B. Mandelbrot : *Fractal properties of rain and a fractal model.* Tellus,ser.A, vol 37, pp. 209-232. 1985
- [43] M. Magdziarz, A. Weron, K.Burnec. *FBM versus the CTRW : a simple test for subdiffusive dynamics.* Physical Review Letters. Volume 103(18), pp.180602. 2009
- [44] M. Magdziarz, A. Weron. *Fractional Fokker Planck dynamics : stochastic representation and computer simulation.* Phys. Rev. E, Volume 75, Issue 1, pp. 016708. Doi : 10.1103/PhysRevE.75.016708. 2007
- [45] S. Mallat *A Wavelet Tour of Signal Processing.* Ed. Elsevier. 3th Edition. 2009
- [46] B.B. Mandelbrot and J.R. Wallis. *Noah, Joseph, and Operational Hydrology* Water Resources Research, 4, 5, 909-918.1968
- [47] B.B. Mandelbrot, Van Ness. *Fractional Brownian Motion, fractional noises and Applications.* SIAM REW., 10, 422-437. 1968
- [48] B.B. Mandelbrot *Statistical Methodology for Nonperiodic Cycles from Covariance to R/S Analysis* Annals of Economic and Social Measurement. Vol.1. P.259-290. 1972
- [49] B.B. Mandelbrot, *The Fractal Geometry of Nature.* Ed. W. H. Freeman, New York. 1982.
- [50] B.B.MandelBrot *Intermittent turbulence in self similar cascades : Divergence of higher moments and dimension of the carrier.* Journal of Fluid Mechanics. Vol. 62. P. 331-358. 1974
- [51] M. Marani *Non-Power law scale properties of rainfall in space and time.* Water Resources Research. Vol. 41(5). W08413. Doi : 10.1029/2004WR003822. 2005

- [52] F. Maraux and F. Lafolie. *Modelling soil water balance of a maize-sorghum sequence*. Soil Sci. Soc. Am. J. Vol. 62. P. 75-82. 1998
- [53] T.T. Marquez-Lago, A. Leier and K. Burrage *Anomalous diffusion and multifractional Brownian motion : simulating molecular crowding and physical obstacles in systems biology*. Systems Biology, IET. Vol. 6 , Issue : 4. P. 134-142. DOI : 10.1049/iet-syb.2011.0049. 2012
- [54] J.H. McCulloch. *Simple consistent estimators of stable distribution parameters*. Communications in Statistics-Simulation and Computation, Volume 15, Issue 4, pp. 1109-1136. DOI :10.1080/03610918608812563. 1986.
- [55] J.H. McCulloch. *Measuring tail thickness in order to estimate the stable index α : a critique*. Bussiness and Economic Statistics. Vol. 15. P. 74-81. 1997.
- [56] M. M. Meerschaert and H. P. Scheffler *Limit Distributions for Sums of Independent Random Vectors*. John Wiley, New York. 2001
- [57] M. Meerschaert, Y. Zhang, B. Baeumer. *Tempered anomalous diffusion in heterogeneous systems*. Geophysical Research Letters, Volume 35, Issue 17. DOI : 10.1029/2008GL034899. 2008
- [58] R. Metzler, A. Chechkinb, V. Goncharb, J. Klafter. *Some fondamental aspect of Lévy flight*. Chaos Solitons and Fractals, 34 DOI : 10.1016/j.chaos.2007.01.055. 2007
- [59] R. METZLER and J. KLAFTER *The random walk's guide to anomalous diffusion : A fractional dynamics approach*. Physics Reports. Vol. 339 (2000) P. 1-77. 2000
- [60] S. Mittnik and M.S. Paolella. "Prediction of financial downside-risk with heavy-tailed conditional distributions". Handbook of Heavy Tailed Distributions in Finance. P. 385-404. 2003. Ed. North-Holland, Amsterdam.
- [61] E. W. Montroll, G. H. Weiss *Random walks on lattices. II*. Journal of Math. Phys. Volume 6, pp. 167-181. 1965
- [62] S. Naranayan and F. Hussain. *Measurements of spatiotemporal dynamics in a forced plane mixing layer*. J. Fluid Mech. Vol.320, p. 71-115. 1996
- [63] M.C. Néel, A. Abdennadher and M. Joelson. *Fractional Fick's law : the direct way*. Journal of Physics A : Math and Theor.,
- [64] C.L. Nikias and M.Shao. *Signal Processing with Alpha-Stable Distributions and Applications*. Ed. Prentice Hall. 1995.

- [65] J. Olsson *Limits and characteristics of the multifractal behaviour of a high-resolution rainfall time series*. Nonlin. Processes Geophys. Vol. 2. P. 23-29. 1995
- [66] J. Olsson and P. Burlando *Reproduction of temporal scaling by a rectangular pulses rainfall model*. Hydrol. Process. Vol. 16. P. 611-630. 2002
- [67] J. Peyriere *Turbulence et dimension de Hausdorff*. C.R. Séances Académies des Sciences Ser.A. 278A. 567-569. 1974
- [68] R.D. Pierce *Application of the positive alpha stable distribution*. Proc. IEEE Signal Processing Workshop on Higher-Order Statistics, SPW-HOS'97, pp. 420-424. 1997
- [69] S.J. Press. *Estimation in univariate and multivariate stable distributions*. Journal of the American Statistical Association. Vol. 67. P. 842-846. 1972
- [70] W.J. Reed, and B.D. Hughes *From gene families and genera to incomes and internet file sizes : Why power laws are so common in nature*. Phys. Rev. E 66, 067103. 2002
- [71] S.G. Roux, V. Venugopal, K. Fienberg, A. Arneodo and E. Foufoula-Georgiou. *Evidence for inherent nonlinearity in temporal rainfall* Advances in Water Resources. Vol. 32. P. 41-48. 2009
- [72] G. Samorodnitsky, M.S. Taqqu. *Stable Non-Gaussian Random Processes*. Chapman and Hall. 1994
- [73] G. Samorodnitsky. *Long Range Dependence, Heavy Tails and Rare Events*. MaPhySto, Centre for Mathematical Physics and Stochastics, Aarhus. Lecture Notes. 2002.
- [74] A. L. Sellerio, M. Daniele and G. Gremaud. *Fractional Brownian motion and anomalous diffusion in vibrated granular materials* J. Stat. Mech. P01002. DOI :10.1088/1742-5468/2012/01/P01002. 2012
- [75] D. Schertzer and S. Lovejoy *Physical modeling and analysis of rain and clouds by anisotropic scaling multiplicative processes*. Journal of Geophysical Research : Atmospheres. Vol. 92(D8). P. 9693-9714. 1987
- [76] D. Schertzer, I. Tchiguirinskaia, S. Lovejoy, P. Hubert, H. Bendjoudi and M. Larchevêque. *Which chaos in the rainfall-runoff process* Hydrological Science. Vol.47. P. 139-148. 2002
- [77] H. Sheng, Y. Q. Chen and T. Qiu. *Heavy-tailed distribution and local long memory in time series of molecular motion on the cell membrane* Fluctuation and Noise Letters. Vol. 10. No. 01. P. 93-119, (doi : 10.1142/S0219477511000429). 2011

- [78] D. Serletis. *Effect of noise on fractal structure*. Chaos Solitons and Fractals, Volume 38, Issue 4, P. 921-924. Doi 10.1016/j.chaos.2007.01.031. 2008
- [79] K. Sigman. *A primer on heavy-tailed distributions*. Queueing Systems, Vol.33, 261-275. 1999
- [80] F. Takens F. Detecting strange attractors in turbulence In :Lecture notes in mathematics. Rand DA , Young LS (Eds), Vol 898, Springer Verlag, Berlin, pp. 366 -381. 1981
- [81] C.D. Tankersley, W.D. Graham and K. Hatfield *Comparison of univariate and transfert function models of groundwater fluctuations*. Water Ressources Research, Volume 29, Issue 10, pp 3517-3533. 1993.
- [82] M.S. Taqqu, W. Willinger, R. Cherman. *Proof of fundamental result in self-similar traffic modelling*. Computer Communication Review, Volume 27, PP 5-23. 1997.
- [83] L. Telesca, V.Lapenna, E. Scalcione, D. Summa *Searching for time-scaling features in rainfall sequences*. Chaos Soliton and Fractal Volume 32. 2007, pp. 35-41. 2005
- [84] J. Theiler, S. Eubank, A. longlin, J.D. Farmer. *Testing for nonlinearity in times series : the methode of surrogate*. Physica D, Vol 58, pp 77-94. 1992
- [85] A.A. Tsonis. *Chaos from theory to Applications*. Plenum Press New York. 1992
- [86] L. Turgeman, S. Carmi, E. Barkai. *Fractional Feynman-Kac Equation for non-Brownian functional*. Physical Review Letters, Volume 103(19). pp.190201. 2009
- [87] B. Venema, S. Bachner, H.W. Rust and C. Simmer. *Statistical characteristics of surrogate data based on geophysical measurements* Non-linear Processes in Geophysics. Vol. 13. P. 449-466. 2006
- [88] D. Veneziano D., R.L. Bras and J.L. Niemann. *Nonlinearity and self-similarity of rainfall in time and a stochastic model*. Journal of Geophysical Research. Vol. 101. P. 371-392. 1996.
- [89] D. Veneziano, A. Langousis, P. Furcolo *Multifractality and rainfall extremes*. Water resources research, Volume 42, W06D15. Doi :10.1029/2005WR004716. 2006
- [90] D. Veneziano, P. Furcolo, V. Iacobellis. *Imperfect scaling of time and space-time rainfall*. Journal of Hydrology, Volume 322 (2006), pp. 105-119. 2005
- [91] H.VENTSEL "Théories des probabilités", Editions MIR MOSCOU, 1973

- [92] R.Weron. *Computationally intensive Value at Risk calculations*. In "Handbook of Computational Statistics : Concepts and Methods", Eds. J.E. Gentle, W. Haerdle, Y. Mori, Springer, Berlin, 911-950. 2004
- [93] A. Weron, M. Magdziarz, and K. Weron, *Modeling of subdiffusion in space-time dependent force fields beyond the fractional Fokker-Planck equation*. Physical Review E, Volume 77, 036704. 2008
- [94] YAGLOM *Introduction to the Theory of stationary Random Processes*. Prentice-Hall, Englewood Cliffe, 1962
- [95] Y.K. Zhang, K. Schilling *Temporal scaling of hydraulic head and river base flow and its implication for groundwater recharge*. Water Resources Research Volume 40, Issue 3. DOI : 10.1029/2003WR002094. 2004
- [96] Y.K. Zhang, Z. Li. *Temporal scaling of hydraulic head fluctuations : Nonstationary spectral analyses and numeral simulations*. Water Ressources Research, Volume 41, W07031. Doi : 10.1029 :2004WR003797. 2004