



**HAL**  
open science

# Simulation stochastique des précipitations à fine échelle : application à l'observation en milieu urbain

Nawal Akrouf

► **To cite this version:**

Nawal Akrouf. Simulation stochastique des précipitations à fine échelle : application à l'observation en milieu urbain. Météorologie. Université Paris Saclay (COMUE), 2015. Français. NNT : 2015SACLV014 . tel-01357968

**HAL Id: tel-01357968**

**<https://theses.hal.science/tel-01357968v1>**

Submitted on 30 Aug 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

NNT 2015SACLV014

THESE DE DOCTORAT  
DE L'UNIVERSITE PARIS-SACLAY,

Préparée au LATMOS/IPSL « Laboratoire ATmosphères, Milieux, Observations Spatiales »

ÉCOLE DOCTORALE N° 129

Sciences de l'Environnement d'Ile de France

*Spécialité de doctorat : Océan Atmosphère, Climat et Observations spatiales*

Par

**Mme AKROUR Nawal**

Titre de la thèse

Simulation stochastique des précipitations à fine échelle: Application à  
l'observation en milieu urbain

**Thèse présentée et soutenue au LATMOS, le 27/11/2015 :**

**Composition du Jury :**

Mme Thiria Sylvie	Professeur / UVSQ	Président
Mr Bernardara Pietro	Ingénieur chercheur expert /EDF R&D	Rapporteur
Mr Delrieu Guy	Directeur de recherche/ LTHE	Rapporteur
Mr Kegl Balazs	Directeur de recherche/ LAL	Examineur
Mme Mallet Cécile	Maitre de conférences/ UVSQ	Directeur de thèse
Mr Chazottes Aymeric	Maitre de conférences/ UVSQ	Co-Encadrant de thèse
Mme Sandrine Vial	Maitre de conférences/ UVSQ	Invitée
Mr Sébastien Verrier	Maitre de conférences/Toulouse III	Invité

**Titre : Simulation stochastique des précipitations à fine échelle: Application à l'observation en milieu urbain.**

**Mots clés :** modélisation multi-échelle des précipitations, simulation stochastique, observation des précipitations, modélisation multifractal, invariance d'échelle.

**Résumé :**

Les précipitations ont une très grande variabilité sur une large gamme d'échelles tant spatiale que temporelle. Cette variabilité est une source importante d'incertitude pour la mesure, les applications et la modélisation, et au-delà pour la simulation et la prévision. De plus, les précipitations sont des processus extrêmement intermittents et possèdent plusieurs régimes d'invariance d'échelle.

Le générateur de champ précipitant développé au cours de la thèse est basé sur la modélisation statistique de l'hétérogénéité et de l'intermittence des précipitations à fine échelle. L'originalité de la modélisation repose en partie sur l'analyse de données observées par un disdromètre à très fine résolution. Cette modélisation qui diffère des modèles existants dont la résolution est plutôt de l'ordre de la minute, voire de l'heure ou du jour, permet d'obtenir des simulations dont les propriétés sont réalistes sur une large gamme d'échelle.

Ce simulateur permet de produire des séries chronologiques dont les caractéristiques statistiques sont similaires aux observations aussi bien à l'échelle de simulation (15s) qu'après dégradation (1h et 1 jour). Les propriétés multi-échelles du simulateur sont obtenues grâce à une approche hybride qui repose sur une simulation à fine échelle des événements de pluie par un générateur multifractal associé à une simulation du support basée sur une hypothèse de type Poissonienne. Une étape de re-normalisation des taux de pluie assure l'adaptation du générateur à la zone climatique considérée.

Le simulateur permet la génération de cartes 2D de lames d'eau. La méthodologie développée pour les séries chronologiques est étendue au cas 2D. Le simulateur stochastique multi-échelle 2D ainsi développé reproduit les caractéristiques géostatistiques et topologiques à la résolution de  $1 \times 1 \text{ km}^2$ .

Ce générateur est utilisé dans le cadre d'une étude de faisabilité d'un nouveau système d'observation des précipitations en milieu urbain. Le principe de ce système repose sur l'utilisation de mesures opportunistes de l'affaiblissement subi par les ondes radios émises par les satellites géostationnaires TV-SAT dans la bande 10.7-12.7 GHz. De façon plus spécifique on suppose que les terminaux de réception TV-SAT installés en ville chez les particuliers sont capables de mesurer de tels affaiblissements. A ce stade de l'étude nous ne disposons pas de telles observations. L'étude s'appuie donc sur des cartes de précipitations issues du générateur 2D et d'un réseau de capteur hypothétique. Le système d'observation envisagé permettra d'estimer les champs de précipitation ( $30 \times 30 \text{ Km}^2$ ) et avec une résolution spatiale de  $0.5 \times 0.5 \text{ Km}^2$ .

**Title: Stochastic simulation of precipitation at fine scales: observation application in urban environment**

**Key Words:** multiscale modeling of precipitation, stochastic simulation, rainfall measuring, multifractal modeling, scale invariance properties

**Abstract**

Precipitations are highly variable across a wide range of both spatial and temporal scales. This variability is a major source of uncertainty for the measurement and modeling, also for the simulation and prediction. Moreover, rainfall is an extremely intermittent process with multiple scale invariance regimes.

The rain-field generator developed during the thesis is based on the fine-scale statistic modeling of rain by the mean of its heterogeneity and intermittency. The modeling originality partially rest on the analysis of fine-scale disdrometer data. This model differs from other existing models whose resolution is roughly a minute or even an hour or a day. It provides simulations with realistic properties across a wide range of scales.

This simulator produces time series with statistical characteristics almost identical to the observations both at the 15s resolution and, after degradation, at hourly or daily resolutions. The multi-scale properties of our simulator are obtained through a hybrid approach that relies on a fine scale simulation of rain events using a multifractal generator associated with a rain support simulation based on a Poissonian-type hypothesis. A final re-normalization step of the rain rate is added in order to adapt the generator to the relevant climate area.

The simulator allows the generation of 2D water-sheets. The methodology developed in the first part is extended to the 2 Dimension case. The multi-scale 2D stochastic simulator thus developed can reproduce geostatistical and topological characteristics at the spatial resolution of  $1 \times 1 \text{ km}^2$ .

This generator is used in the scope of the feasibility study of a new observation system for urban area. The principle of this system is based on the opportunistic use of attenuation measurements provided by geostationary TV satellites which radio waves lay in the 10.7 to 12.7 GHz bandwidth. More specifically it is assumed that the SAT-TV reception terminals installed in private homes are able to measure such attenuations. At this stage of the study we do not have such observations. The study is therefore based on rainfall maps generated using the 2D generator in addition to a hypothetical sensor network. The considered observation system will allow to estimate precipitation fields ( $30 \times 30 \text{ km}^2$ ) with a spatial resolution of  $0.5 \times 0.5 \text{ km}^2$ .

*« J'ai appris que le succès ne doit pas être mesuré par la position qu'on atteint dans la vie, mais par les obstacles qu'on a dû surmonter, tout en essayant de réussir »*

*Booker T. Washington*



## Remerciements

En premier, je pense naturellement à mon directeur de thèse Cécile Mallet et mon co-directeur de thèse Laurent Barthès. Merci pour leur soutien, leur gentillesse, leur encadrement tout au long de cette thèse. Merci pour leur disponibilité sur tous les plans qui a été capitale pour l'aboutissement de ce travail ; je leur témoigne ma profonde gratitude. Merci également pour les corrections, la lecture et relecture de ce manuscrit. C'était pour moi une très belle expérience d'avoir travaillé avec eux, leur expérience et leurs qualités professionnelles étaient un grand bénéfice pour moi.

Un merci très particulier à Aymeric Chazottes, mon co-encadrant pour son encadrement. Il m'a apporté un soutien très précieux dans le cadre de ce travail. Ses encouragements incessants m'ont été d'une aide précieuse. Ses conseils avisés et ses qualités humaines m'ont été d'une très grande utilité pour mener à terme ce travail.

Merci à Sébastien Verrier pour ces suggestions, ces conseils et son aide. Merci pour la relecture de l'article. Merci aussi pour sa thèse et tous ces travaux qui ont été la base sur laquelle s'est appuyé mon travail.

Je tiens à remercier très sincèrement Pietro Bernardara d'avoir participé à mon comité de thèse. Merci à lui et Guy Delrieu d'avoir accepté d'être les rapporteurs de ce manuscrit. Merci pour l'intérêt qu'ils ont porté à mon travail.

J'adresse mes sincères remerciements à Sylvie Thiria et Balázs Kégl qui me font l'honneur de juger mon travail.

Un grand merci aussi à Sandrine Vial d'avoir accepté mon invitation et de m'avoir donné la chance d'enseigner dans le département informatique.

Je remercie Philippe Naveau d'avoir accepté de participer à mes deux réunions de comité de thèse, merci pour tous ses conseils, ses remarques et ses critiques qui ont été très utiles pour mener ce travail.

Cette thèse a été réalisée au sein de l'équipe SPACE du Laboratoire Atmosphères, Milieux, Observations Spatiales, je tiens donc à remercier l'équipe pour son accueil et sa bienveillance ainsi que les personnes que j'ai rencontrées durant les 3 années passées au LATMOS.

Je tiens à remercier ma petite famille pour son amour et patience particulièrement mon mari pour son soutien et son encouragement sans oublier ma fille Amel pour tous ses caprices et ses bêtises.

Je remercie également toute ma famille et ma belle-famille et plus particulièrement mon père et ma mère pour leur amour, leur soutien constant malgré la distance qui nous sépare.

Enfin, je dis grand merci à tous ceux qui, de près ou de loin, ont contribué à l'aboutissement et à la réussite de ce travail.

# Table des matières

Remerciements .....	5
Résumé .....	2
Abstract .....	2
Table des matières .....	7
Liste des figures .....	11
Liste des Tables .....	17
1 INTRODUCTION.....	19
1.1 Contexte général .....	19
1.2 Cadre de la thèse .....	24
1.3 Objectif de la thèse .....	26
1.4 Plan de la thèse .....	28
2 FORMALISME ET TECHNIQUES MULTIFRACTALES .....	33
2.1 Théorie des fractals.....	33
2.1.1 Objets fractal et multifractal.....	33
2.1.2 Fractals aléatoires et fractals déterministes .....	34
2.1.3 Résolution et échelle .....	34
2.1.4 Dimension et co-dimension fractales .....	34
2.1.5 Les Multifractals .....	36
2.1.5.1 Notion de singularité.....	36
2.1.5.2 Fonction d'échelle des moments.....	37
2.1.6 Les cascades multiplicatives .....	38
2.1.6.1 Cascade continue et discrète .....	38
2.1.7 Modèle multifractal universel : .....	40
2.1.8 Champs conservatifs / champs non conservatifs.....	41
2.2 Simulation et analyse multifractale : modèle FIF.....	42
2.2.1 Caractéristiques du générateur multifractal universel.....	42
2.2.1.1 Simulation des zéros (intermittence de la pluie).....	44
2.2.1.2 Produits de convolution .....	45
2.2.2 L'Analyse multifractale.....	47
2.2.2.1 Limitation de la taille de l'échantillon et observation des extrêmes.....	48
2.2.2.2 Divergence des moments .....	50
2.2.2.3 Estimation des moments .....	51



2.3	Restitution des paramètres multifractals en pratique.....	52
2.4	Généralisation de cartes en dimension 2 .....	54
2.5	Conclusion .....	54
3	GENERATEUR DE SERIES TEMPORELLES DE PRECIPITATIONS .....	59
3.1	Introduction .....	59
3.2	Séries temporelles de taux précipitants observées par le DBS .....	63
3.3	Périodes sèches et périodes pluvieuses.....	66
3.4	Hypothèses retenues .....	67
3.5	Analyse et modélisation du support de pluie des données observées par le DBS .....	68
3.5.1	Analyse des propriétés du support de pluie observé .....	68
3.5.2	Modélisation du support à partir d'une distribution GEV.....	72
3.5.3	Modélisation fine du support de pluie.....	76
3.6	Analyse et modélisation des séries de taux précipitants observés.....	80
3.6.1	Analyse des évènements observés et estimation des paramètres multifractals .....	80
3.6.2	Génération des séries temporelles de pluie .....	83
3.7	Générateur 1 .....	83
3.7.1	Introduction .....	83
3.7.2	Calibration quantile à quantile .....	84
3.7.3	Performances .....	87
3.7.3.1	Analyse spectrale .....	88
3.7.3.2	Analyse multifractale .....	90
3.7.3.3	Autocorrélation .....	91
3.7.4	Analyse critique.....	92
3.8	Générateur 2 .....	92
3.8.1	Introduction .....	92
3.8.2	Relation durée/intensité.....	94
3.8.3	Performances .....	97
3.8.3.1	Analyse spectrale .....	97
3.8.3.2	Analyse multifractale .....	98
3.8.3.3	Distribution de probabilité .....	99
3.9	Conclusion .....	101
4	GENERATION DE CARTES 2D DE PRECIPITATIONS .....	105
4.1	Introduction .....	105
4.2	Les données .....	107

4.3	Méthodologie.....	109
4.4	Analyse et modélisation du support de pluie.....	110
4.4.1	Quelques approches possibles pour la simulation du support.....	110
4.4.2	Méthode retenue (algorithme de simulation séquentielle) :Etape 1.....	110
4.4.2.1	Principe de l'analyse variographique.....	112
4.4.3	Analyse et modélisation des variogrammes empiriques des supports de pluie observés.....	116
4.4.3.1	Choix du modèle paramétrique.....	118
4.4.3.2	Estimation des paramètres des variogrammes observés.....	119
4.4.4	Exemple de supports simulés.....	122
4.5	Simulation des lames d'eau dans les zones de pluie.....	123
4.5.1	Simulation des zones de pluie: étape 2.....	123
4.5.2	Re-normalisation des lames d'eau simulées: Etape 3.....	125
4.5.2.1	Relation entre la surface et la lame d'eau moyenne des évènements.....	125
4.5.3	Traitement des transitions pluie/non pluie.....	130
4.6	Validation des cartes simulées.....	134
4.6.1	Propriétés multifractales des simulations.....	134
4.6.2	Propriétés spectrales.....	135
4.6.2.1	Analyse des propriétés spectrales des observations.....	135
4.6.2.2	Analyse des propriétés spectrales des simulations obtenues avec support conditionnel (réel).....	139
4.6.2.3	Analyse des propriétés spectrales des simulations obtenues avec support inconditionnel (synthétique).....	141
4.6.3	Distributions statistiques.....	142
4.6.3.1	Analyse statistiques des observations.....	142
4.6.3.2	Analyse statistique des simulations avec support conditionnel (réel).....	143
4.6.3.3	Analyse statistiques des simulations avec support inconditionnel (synthétique).....	144
4.7	Conclusion.....	145
5	APPLICATION A L'ESTIMATION DE CARTES DE PRECIPITATIONS A PARTIR DE LIAISONS MICRO-ONDES.....	149
5.1	Introduction :.....	149
5.2	Les données.....	152
5.2.1	Principe de simulation d'une liaison.....	153
5.2.2	Configurations retenues.....	156
5.3	Méthodologie pour l'estimation des champs précipitants.....	159

5.3.1	Assimilation variationnelle .....	159
5.3.1.1	3D-Var .....	160
5.3.1.2	4D-Var .....	160
5.3.2	Hypothèses retenues .....	162
5.3.2.1	Modèle advection pure .....	162
5.3.2.2	Vents .....	164
5.3.2.3	Bruit de mesure .....	164
5.3.2.4	Résolution temporelle .....	164
5.4	Résultats.....	164
5.4.1	Influence du nombre et de la position des antennes et de la direction du vent .....	165
5.4.2	Influence du bruit de mesure sur la qualité de la restitution .....	169
5.4.3	Echelles temporelles d'observation.....	171
5.4.4	Résultats en fonction de taux précipitants.....	174
5.5	Conclusion .....	177
6	CONCLUSION ET PERSPECTIVES .....	179
	Annexe .....	183
	Bibliographie .....	185
	Communication .....	195
	Article.....	197

## Liste des figures

Figure 1.1 Inondation à Montpellier et dans le département de l’Hérault en septembre 2014 .....	19
Figure 1.2 Hétérogénéité et caractère discret des précipitations à l’échelle microphysique, distribution des tailles de gouttes en fonction du temps.....	20
Figure 1.3 Exemple d’échelle d’observation du typhon Maysak approchant les Philippines le 3 Avril 2015 observé par la mission GPM. ....	21
Figure 1.4 Illustration des relations entre échelles et extrema : records de précipitation en France métropolitaine pour 5 séries de 50 ans pour des pas de temps de 6 mn à 1 jour selon une échelle logarithmique, figure extraite de (Soubeyroux, 2011) .....	22
Figure 1.5 Gammes de scaling des précipitations d’après Fabry (1996) .....	23
Figure 2.1 Exemple d’objets présentant des propriétés fractales dans la nature. ....	33
Figure 2.2 Calcul de la dimension fractal sur l’exemple de la poussière de cantor. ....	35
Figure 2.3 Schéma de principe des cascades multiplicatives discrètes 1D et 2D .....	39
Figure 2.4 Schéma de principe du générateur multifractal (D’après Verrier, 2011) .....	43
Figure 2.5 a et b Spectre obtenu sur une seule série simulée, c et d spectre moyen obtenu sur 100 séries simulées. Les deux figures de gauche sont obtenues avec le modèle originel et les deux figures de droite sont simulées avec le modèle modifié (N=8).....	47
Figure 2.6 Ordre maximum de la singularité (Schertzer et Lovejoy, 1993) .....	49
Figure 2.7 Moments estimés sur une seule réalisation (à gauche) et sur 100 réalisations (à droite) .....	52
Figure 2.8 Exemple de carte 2D 256x256 pixels générée à l’aide du modèle FIF avec les paramètres $\alpha=1.6$ , $C1=0.1$ et $H = 0.4$ . ....	54
Figure 3.1 Schéma de principe du Dual Beam Spectropluviometer .....	64
Figure 3.2 Fréquence des durées des périodes pluvieuse en fonction de la résolution des séries d’apprentissage et de validation accompagnée de leur co-dimension.....	70
Figure 3.3. Densités de probabilité empiriques des durées des périodes de pluie (rain duration) et des périodes de sécheresse (drought duration) de la série d’apprentissage (cf. section 3.3 pour les définitions de ces deux types de périodes.) .....	71
Figure 3.4 Densité de probabilité des périodes de pluie et des périodes de sécheresse de la série d’apprentissage et d’une série simulée.....	74
Figure 3.5. Fréquence de la pluie du support de la série apprentissage et du support simulé en fonction de la résolution .....	75
Figure 3.6. Période de retour des durées des évènements des séries apprentissages et validation et de la série simulée (même exemple de support utilisé pour tracer la Figure 3.4). ....	75
Figure 3.7 Pdfs normalisées des durées pluvieuses et des durées sèches de la série d’apprentissage et de la série simulée .....	78

Figure 3.8. Probabilité qu'il pleuve en fonction de la résolution pour les séries d'apprentissage et de validation (bleu) et pour 100 simulations (courbes en gris). La courbe en rouge représente la moyenne des 100 séries simulées.....	79
Figure 3.9. Période de retour des durées des évènements observés et synthétiques (même exemple de support synthétique utilisé pour tracer la Figure 3.7).....	80
Figure 3.10. Histogramme ainsi que les moments associés pour des valeurs des paramètres $\alpha$ (a) et $\mathbf{C1}$ (b) estimés sur les évènements de la série apprentissage .....	81
Figure 3.11. Estimation du paramètre H sur les évènements de pluie .....	82
Figure 3.12. Schéma de principe du générateur 1 .....	84
Figure 3.13. Fonction de calibration $\mathbf{RR}(x)$ : $\mathbf{R}$ (RR observé) versus X (x simulé) .....	86
Figure 3.14. Fonction de survie de la série temporelle issue du DBS, et celle de la série simulée avant et après calibration.....	86
Figure 3.15. Fonction de survie de la série DBS et ajustement de loi de Pareto tronquée à la résolution de 15 secondes (les bins sont en mm/h). .....	87
Figure 3.16. Fonction de survie de la série DBS et exemple d'une réalisation calibrée à l'aide de la loi de Pareto tronquée à une résolution dégradée de 5 minutes) .....	87
Figure 3.17. Spectre de puissance de la série DBS (bleu) ayant servi pour apprendre les paramètres du modèle, d'une série simulée avant (rouge) et après calibration (rose) .....	89
Figure 3.18. Spectre en loi de puissance de 100 séries simulées avec le générateur 1 et des séries apprentissage et validation .....	89
Figure 3.19. Distribution des paramètres multifractals obtenue sur 100 séries simulées. ....	91
Figure 3.20. Comparaison de l'autocorrélation entre les données DBS observées et les données synthétiques.....	91
Figure 3.21. Schéma de principe du générateur 2 .....	93
Figure 3.22. Intensité moyenne en fonction de la durée obtenue sur l'ensemble des périodes de pluie de la série apprentissage (en bleu). Intensité moyenne sur l'ensemble des périodes de même durée en rouge. ....	95
Figure 3.23. CDF des intensités moyenne des périodes de pluie de la série apprentissage et de la loi Alpha-stable correspondant.....	96
Figure 3.24. Spectre des séries apprentissage et validation, spectres des 100 séries simulées, spectre moyen des séries simulées .....	97
Figure 3.25. Distribution des paramètres multifractals estimés sur les évènements extraits de la série observée (histogramme en bleu) et la moyenne par bin des 100 séries simulées avec le générateur 2 (histogramme en rouge) avec l'écart type par bin (en vert) .....	99
Figure 3.26. De gauche à droite: Fonctions de survie empiriques des taux de pluie RR, des durées des périodes pluvieuses et de la hauteur d'eau correspondante. De haut en bas: les fonctions de survie calculées à différentes résolutions 15sec, 5 min, 1 heure et 1 journée. ....	101
Figure 4.1. Réseau Radar Météo France. ....	107
Figure 4.2. Exemple de variogramme empirique.....	114

Figure 4.3 Exemples de supports issus d'images radar, et des variogrammes empiriques correspondant à différents pourcentages de valeurs non nulles (exemple 1) 15%, (exemple 2) 86%, (exemple 3) 49%, (exemple 4) 54%.....	118
Figure 4.4 Exemples de variogrammes empirique et de modèles paramétriques ajustés (gaussien, exponentiel, stable et sphérique). .....	119
Figure 4.5. Valeurs des paramètres du modèle de variogramme sphérique et probabilité d'occurrence de pluie en fonction du numéro de la carte sur une sélection de 1277 cartes (Hiver 2012 ayant un pourcentage de pluie supérieur à 15%).....	120
Figure 4.6. Histogramme des paramètres des variogrammes empiriques et la probabilité d'occurrence de pluie calculés sur les 1277 cartes sélectionnées. ....	122
Figure 4.7. À gauche, exemple de support synthétique (128x128) généré avec les paramètres $\mu=0.25$ , $a=31.44$ , $c=0.17$ et $c_0=0$ . A droite, un support réel observé ayant des paramètres proches de ceux utilisé pour la simulation ( $\mu=0.29$ , $a=30.02$ , $c=0.16$ et $c_0=0.02$ ) .....	123
Figure 4.8. Variogramme empirique estimé sur la simulation présentée Figure 4.7 (courbe rouge). En bleu l'ajustement du variogramme par un modèle sphérique ( $\mu=0.22$ , $a=33.12$ , $c=0.18$ , $c_0=0.03$ ).....	123
Figure 4.9. Exemple d'extraction des évènements (cellules) de pluie : à gauche la carte radar des précipitations, à droite les évènements extraits entourés du vert .....	126
Figure 4.10. Moyenne des lames d'eau par cellule en fonction de la dimension des cellules de pluie. ....	127
Figure 4.11. Exemple d'ajustement sur nos données de la loi $\alpha$ -stable sur la distribution empirique des lames d'eau par évènement pour une longueur équivalente comprise entre 9 et 10 km.....	127
Figure 4.12. Paramètres de stabilité, d'asymétrie, d'échelle et de location estimé pour les lois $\alpha$ -stable ajustés sur les lames d'eau des évènements observés par classe de distance .....	129
Figure 4.13. Exemple de carte simulée avant l'application du traitement des transitions. A droite carte des transitions $F(d)$ calculée sur la carte simulée .....	131
Figure 4.14. Résultat final du processus de simulation.....	131
Figure 4.15 Méthodologie de simulation d'une carte de lame d'eau .....	133
Figure 4.16. Paramètres multifractals estimés sur une sélection de 50 évènements de 40x40 Km <sup>2</sup> extraits des cartes simulées. ....	134
Figure 4.17. Schéma de la structure spatiale des précipitations à l'intérieur de l'échelle synoptique (Gupta et Waymire 1979; 1981; 1993) .....	136
Figure 4.18. Variabilité spectrale de 42000 cartes radar Météo France de taille 130*130 km observées sur la région de Palaiseau. La moyenne de tous les spectres est présentée en rouge, les courbes noir, rose, jaune et bleu claire sont les spectres moyen des cartes ayant respectivement un pourcentage de pluie <25%, entre 25% et 50%, entre 50% et 75% et au de-là de 75% .....	136
Figure 4.19. Spectre annuel moyen des cartes précipitation en bleu et le spectre moyen annuel de leurs supports.....	139

Figure 4.20. Spectre moyen des données observées en bleu (1 an soit 42 000 cartes) et des données simulées en rouge (4 ans soit 4*42000) dans le cas conditionnel (support réel) .....	140
Figure 4.21. Spectre moyen des données observées en bleu (42000 cartes) et des données simulées en rouge (soit 4*1000) dans le cas inconditionnel .....	141
Figure 4.22. Fonction de survie empirique des lames d'eau des 42000 cartes observées à la résolution 1x1 km <sup>2</sup> à Palaiseau en 2012, la ESF moyenne est présentée en rouge .....	142
Figure 4.23. Fonction de survie empirique des 4 simulations conditionnelles (4 courbes rouge correspondant à chacune à la moyenne de 42000 cartes) et des observations de 2012 (en bleu) 143	
Figure 4.24. . Fonction de survie empirique 4 moyennes annuelles des simulations inconditionnelles en rouge (4 courbes rouge correspondant à chacune à la moyenne de 1000 cartes) et des observations de 2012 (en bleu).....	144
Figure 5.1 image satellitaire de la région d'étude.....	152
Figure 5.2 Carte 1(simulée avec le modèle 2D utilisé pour toutes les configurations testées dans ce chapitre. ....	153
Figure 5.3. Cartes 2 & 3 simulées avec le modèle 2D utilisées pour la configuration « optimale ». ....	153
Figure 5.4 Principe du dispositif de mesure des affaiblissements des ondes de la bande KU .....	154
Figure 5.5 Grille de pixels d'une carte de lames d'eau avec un exemple de liaison présenté en bleu: les taux de pluie des pixels sont notés de $r_1$ à $r_n$ .....	155
Figure 5.6 Les quatre configurations étudiées dans ce chapitre.....	158
Figure 5.7 Exemple de série temporelles d'affaiblissement vues par les liaisons de la configuration 1 en advectant de l'est vers l'ouest le champ de pluie de la Figure 5.2 avec une vitesse =10m/s .....	159
Figure 5.8 Schéma présentant le principe de l'assimilation 4D-Var. 2 cycles de minimisation sont représentés ici (cycles c-1 et c). Dans chaque cas, on a représenté l'état d'ébauche (trait noir), les observations disponibles (ronds noirs) et leurs écarts à l'ébauche (traits noirs verticaux), l'incrément (flèches), c'est-à-dire la modification de l'état initial apportée par le minimiseur et l'état analysé qui en découle (pointillés) (Daget, 2007).....	161
Figure 5.9. Schéma du modèle (figure extraite de l'article de (Mercier et al., 2015)) .....	162
Figure 5.10 Résultats de l'assimilation lorsque la carte 1 est utilisée comme champ de pluie initial pour les 4 configurations géométriques et pour un vent Est – Ouest (colonne 1). Ecart pixel à pixel entre la carte 1 et les cartes obtenues (colonne 2). Colonnes 3 & 4 : Idem aux colonnes 1 et 2 mais pour un vent Nord - Sud .....	166
Figure 5.11. Boîtes à moustaches illustrant la différence entre la carte initiale et les résultats d'assimilation pour les 4 configurations avec le vent Est-Ouest et le vent Nord-Sud.....	167
Figure 5.12. RMSE des pixels en fonction de nombre de liaisons qui les traversent. ....	168
Figure 5.13 Colonnes 1 : Cartes estimées par l'assimilation à partir d'observations non bruitées (ligne 2), avec un bruit de 0,5 dB (ligne 3) et un bruit de 1 dB (ligne 4). Colonne 2 : erreurs correspondantes. Colonnes 3 et 4 : Idem mais à la résolution de 2 km .....	170

Figure 5.14. Boites à moustaches illustrant la différence entre la carte initiale et les résultats d'assimilation avec un bruit de 0.5 dB et 1 dB. ....	171
Figure 5.15 Cartes estimées pour trois résolutions temporelles (10s, 1 min et 5 min) à la résolution spatiale 0.5 km et 2 km.....	173
Figure 5.16. Boites à moustaches des erreurs entre la carte initiale et les cartes estimées par assimilation pour 3 résolutions temporelles (10s, 1min et 5 min) .....	174
Figure 5.17. Cartes restituées avec et sans bruit, en utilisant la configuration 3 avec un vent Est-Ouest.....	175
Figure 5.18. Boites à moustache présentant les résidus à la résolution de 0.5 km et 2 km, avec et sans bruit des 3 cartes utilisées.....	176
Figure 5.19. RMSE des pixels des trois cartes restituées avec et sans bruit en fonction de nombre de liaisons qui les traversent.....	177





## Liste des Tables

Tableau 2.1 paramètres multifractals estimés pour 8 jeux de paramètres appliqués au générateur multifractal original et au générateur multifractal modifié. Dans chaque cas de figure 100 séries à la résolution $\lambda=14$ ont été générées. ....	53
Tableau 3.1 Principales caractéristiques des séries chronologiques de taux précipitants à la résolution initiale de 15s utilisées pour mettre au point et valider le simulateur. Caractéristiques obtenues pour les séries agrégées avec un pas de temps horaire.....	65
Tableau 3.2 Probabilités empiriques des durées des périodes courtes et longues pour les périodes pluvieuses et sèches.....	76
Tableau 3.3 Paramètres de la loi de Pareto estimés sur les durées courtes et longues du support de la série apprentissage.....	77
Tableau 3.4 Caractéristiques du support (Co-dimension et probabilité d'occurrence) des séries deux observées( apprentissage et validation) et des 100 supports simulés avec les deux méthodes. ....	79
Tableau 3.5 Pentés des spectres de la série apprentissage et des séries simulées par le générateur 1 avant et après calibration.....	90
Tableau 3.6 Paramètres de la loi alpha stable estimés sur les intensités moyennes des évènements de durées courtes et longues.....	96
Tableau 3.7 Analyse des paramètres des 100 séries simulées avec le générateur 2. Pour chaque série simulée on considère la moyenne des paramètres estimés sur l'ensemble des périodes continues de pluie de durées supérieures à 15 mn.....	98
Tableau 4.1 Quelques exemple de modèles de variogramme (cf. Chiles et Delfiner 1999).....	116
Tableau 4.2 les pentes moyennes des spectres obtenues en fonction du pourcentage de pluie dans les cartes de précipitation .....	137
Tableau 4.3 Moyenne, écart type et quartiles de $\beta$ calculés sur l'ensemble de 42000 cartes observées dans les deux gammes d'échelles 2-7 km et 7-11 km .....	138
Tableau 4.4 Résumé des grandeurs statistiques calculées sur les pentes des spectres des observations et des simulations dans le cas conditionnel.....	140
Tableau 4.5 Caractéristiques statistiques des valeurs maximales observées sur les cartes radar de Palaiseau.....	143
Tableau 5.1 Caractéristiques des cartes simulées utilisées. ....	153
Tableau 5.2 Liste des satellites géostationnaire utilisés dans notre étude et position orbitale....	157
Tableau 5.3 Résumé des quatre configurations sélectionnées. ....	157
Tableau 5.4 Grandeurs statistiques sur la qualité des champs restitués par assimilation.....	167
Tableau 5.5 Biais, RMSE et intensité de pluie moyenne calculés à partir des cartes estimées par assimilation à la résolution pour des bruits de mesure de 0, 0.5 dB et 1 dB.....	171
Tableau 5.6 Biais, RMSE et intensité de pluie moyenne calculés à partir des cartes estimées par assimilation à la résolution de 10 secondes, 1 minute et 5 minutes .....	173

Tableau 5.7. Grandeurs statistiques sur la qualité des champs restitués en utilisant la configuration 3 pour les trois cartes de pluie avec et sans bruit. .... 175

# 1 INTRODUCTION

## 1.1 Contexte général

Les précipitations sont une partie importante et complexe du cycle d'eau de la Terre et du système climatique. Les processus liés au cycle de l'eau dans l'atmosphère peuvent se produire sur quelques minutes, quelques heures, quelques jours et plus, ainsi que sur des distances de quelques centaines de mètres à quelques dizaines de kilomètres ou plus encore. Ces processus climatiques peuvent dans certains cas induire des événements climatiques extrêmes (sécheresses intenses, pluies diluviennes, tempêtes tropicales et autres ouragans dévastateurs). Ainsi, la dernière décennie a connu un nombre record de précipitations extrêmes et dévastatrices, telles que « les inondations de juillet 2010 au Pakistan (20 millions de personnes affectées et au moins 3 000 morts) ou celles dans l'Est de l'Australie en décembre de la même année (l'épisode le plus intense dans le pays depuis 1900, avec 2,5 milliards de dollars de dégâts) »<sup>1</sup>. Pour terminer cette énumération, on peut noter les très fortes inondations qui ont frappé la France et l'Angleterre en 2014 et qui ont provoqués des dégâts considérables (Figure 1.1).



*Figure 1.1 Inondation à Montpellier et dans le département de l'Hérault en septembre 2014*

Les enjeux, concernant l'hétérogénéité des précipitations dans le temps ou dans l'espace, sont multiples. Les impacts sociétaux en termes de dégâts humains et matériels engendrés par l'hétérogénéité des précipitations concernent aussi bien le court terme (événements extrêmes) que le plus long terme (gestion des ressources en eau). D'un point de vue scientifique la diversité des

---

<sup>1</sup> Voir : <http://ecologie.blog.lemonde.fr/2012/03/27/les-evenements-climatiques-extremes-nouvelle-realite/>

domaines pour lesquels la variabilité des précipitations est problématique (climatologie, météorologie, hydrologie, télécommunication, télédétection) montre à quel point les échelles tant spatiales que temporelles concernées peuvent être différentes.

D'après les différents scénarios étudiés et présentés dans le 5<sup>ème</sup> rapport d'évaluation du GIEC, des changements concernant la variabilité des précipitations sont à prévoir. Les événements extrêmes de précipitation deviendront plus intenses et plus fréquents dans de nombreuses régions (comparaison 2081-2100 par rapport à 1986-2005). Les contrastes entre régions du globe ou entre saisons vont s'accroître. Dans certaines régions, des augmentations de l'ordre de 50 % pouvant se produire à certaines périodes de l'année alors que dans d'autres régions ou pour d'autres périodes les modèles ne prévoient que des changements minimes. D'après le 5<sup>ème</sup> rapport d'évaluation du GIEC ces changements aggraveront les problèmes sociétaux liés à la répartition des ressources hydriques.

L'amélioration des moyens d'observation a permis de quantifier la variabilité des précipitations de la plus petite échelle (micro-physique) à la plus grande (systèmes frontaux, cyclones) (Figure 1.2). Les observations fournies par les disdromètres ont permis d'observer l'hétérogénéité de la répartition granulométrique des gouttes à des échelles de temps de quelques secondes pour des surfaces de collectes très petites (50 à 100 cm<sup>2</sup>).

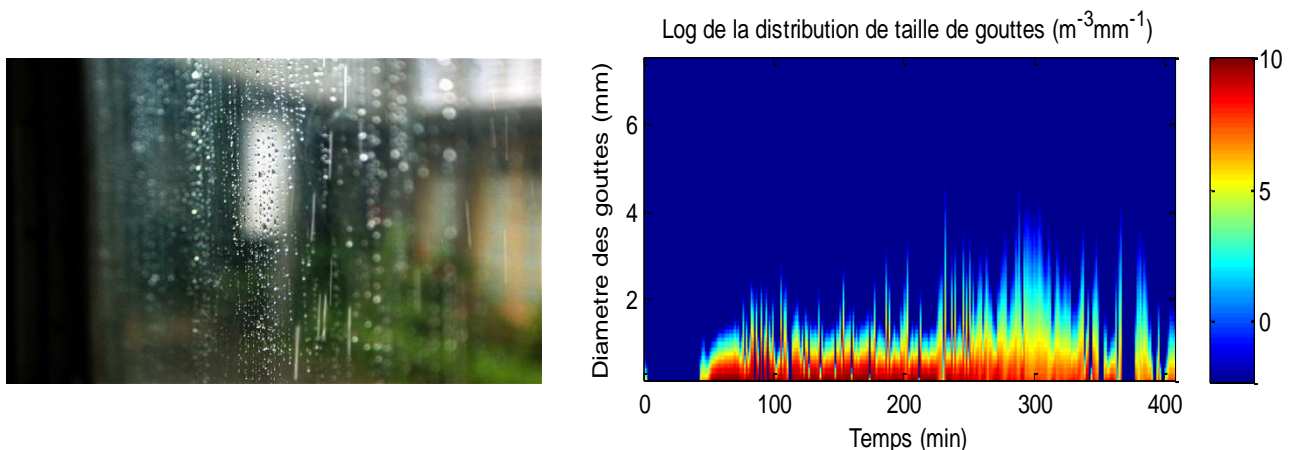
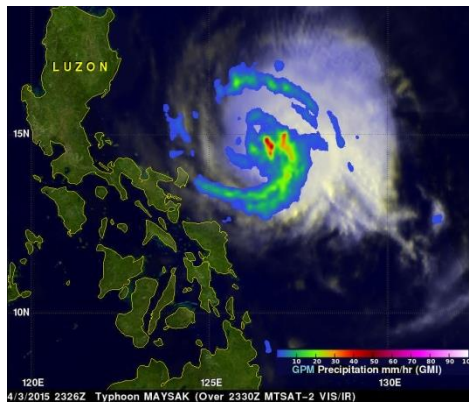


Figure 1.2 Hétérogénéité et caractère discret des précipitations à l'échelle microphysique, distribution des tailles de gouttes en fonction du temps

A l'opposé en termes d'échelle, l'observation satellitaire des précipitations a permis l'étude de la variabilité spatiale à l'échelle du globe avec une résolution de quelques km<sup>2</sup> (Figure 1.3).



*Figure 1.3 Exemple d'échelle d'observation du typhon Maysak approchant les Philippines le 3 Avril 2015 observé par la mission GPM.*

Entre ces deux gammes d'échelle, les radars terrestres et les réseaux de pluviomètres observent la variabilité des précipitations avec des résolutions d'échelle intermédiaires.

L'analyse des nombreuses observations dont la communauté scientifique dispose tant dans le domaine spatial que temporel a mis en évidence des relations statistiques entre les différentes échelles. La localisation des précipitations « intenses » ou « extrêmes » sont sensibles à l'échelle considérée. La hauteur d'eau observée est liée au pas de temps et à la surface de collecte considérée. La Figure 1.4 illustre pour la France métropolitaine les maxima obtenus pour 5 sites au cours des 50 dernières années. Il apparaît par exemple que le record métropolitain de précipitation sur un laps de temps d'une quinzaine de minutes correspond au record de précipitation sur 24 heures sur des villes comme Lille ou Rennes. De plus, pour un pas de temps fixé (1 h, 2h, 3h par exemple) un rapport 2 est observé entre les extrema observés à Lille ou à Marseille (Figure 1.4).

Les propriétés multi-échelles que l'on peut observer trouvent leur origine dans les processus physiques à l'origine des précipitations. L'évolution de l'atmosphère et des systèmes précipitants en particulier est caractérisée par le couplage d'un grand nombre de processus physiques agissant de l'échelle globale à une échelle plus locale. La combinaison de ces phénomènes peut être visualisée en termes de densité spectrale d'énergie (Figure 1.5, reprise de Fabry (1996)) : on observe que le processus de pluie est caractérisé par l'existence de gammes de fréquences

séparées à l'intérieur desquelles le spectre varie suivant une loi d'échelle (loi puissance décroissante).

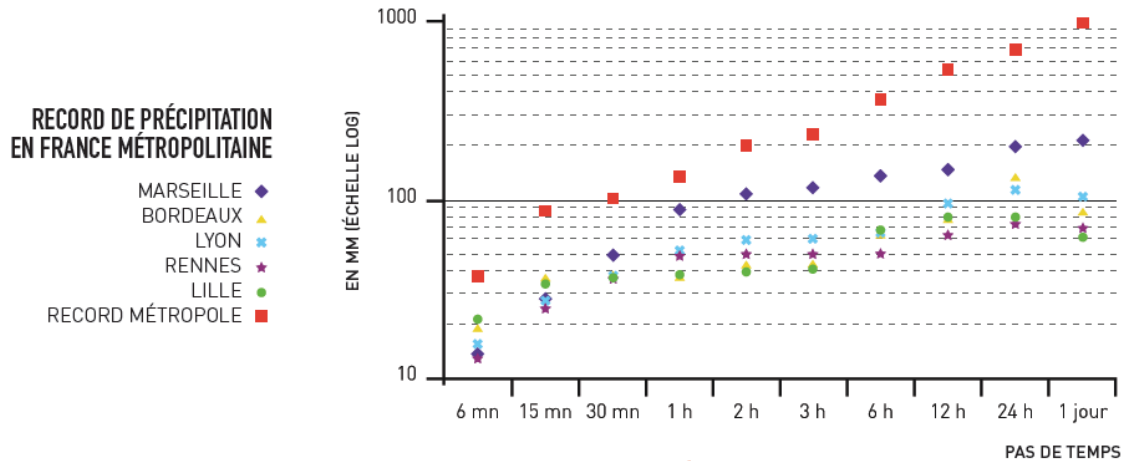


Figure 1.4 Illustration des relations entre échelles et extrema : records de précipitation en France métropolitaine pour 5 séries de 50 ans pour des pas de temps de 6 mn à 1 jour selon une échelle logarithmique, figure extraite de (Soubeyrou, 2011)

Comme indiqué sur la Figure 1.5 l'existence de ces lois d'échelle est en lien plus ou moins direct avec les processus météorologiques. En effet, l'origine physique de telles lois est liée au caractère turbulent des processus atmosphériques (Kolmogorov, 1941 ; Lovejoy et Schertzer, 2008 ; De Montera et al., 2010). De façon schématique l'énergie des grands tourbillons dont la durée de vie est longue (plusieurs jours) est peu à peu transférée à des tourbillons de plus en plus petits, d'énergie moindre et de durée de vie de plus en plus courte (quelques secondes ou moins). Cette turbulence joue un rôle fondamental sur les relations d'échelles. Dans certaines gammes d'échelle, on montre que les interactions entre deux échelles ne dépendent pas des échelles elle mêmes mais seulement de leur rapport. Pour cette raison, les observations des précipitations présentent des propriétés d'invariance d'échelle. Une partie aura donc des propriétés similaires au tout.

A la figure 1.5, les propriétés d'invariances d'échelle sont illustrées pour la pluie. On y voit les processus/régimes associés (i.e. qui prévalent) aux différentes gammes d'échelles. Ainsi selon la gamme d'échelle considérée, on caractérise plus spécifiquement les effets (de la plus grande

échelle à la plus petite) de la variabilité climatique, du forçage barocline, de la turbulence, du mélange des gouttes, de la coalescence, du break-up, ...

Différents modèles multifractals basés sur des cascades multiplicatives ont été développés pour reproduire mathématiquement ces lois d'échelles (Shertzer et Lovejoy 1997). L'utilisation d'un **modèle multifractal universel basé sur des cascades multiplicatives permet de caractériser chaque régime au moyen de seulement trois paramètres**. Ce jeu de paramètres multifractals (habituellement, deux ou trois paramètres caractérisant l'intermittence du processus) contient implicitement l'information des moments et donc de la distribution de probabilité, et ce, à toutes les échelles. Ces modèles permettent de simuler des champs précipitants pour lesquels on observe des propriétés d'invariance dans une gamme d'échelle donnée.

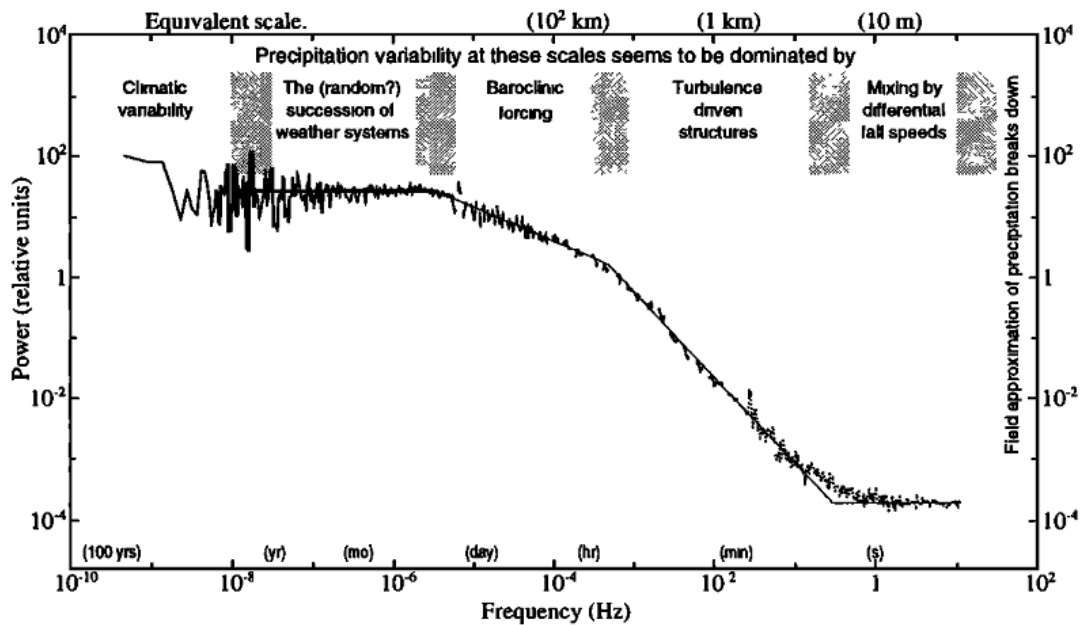


Figure 1.5 Gammes de scaling des précipitations d'après Fabry (1996)

Comme elle repose sur les observations, la majorité des modélisations proposées dans la littérature n'explore pas les fines échelles. En effet, les pluviomètres à auget basculant qui sont majoritairement utilisés pour l'observation dans le domaine temporel ne peuvent donner des résultats non biaisés aux échelles inférieures à 15 mn (Habib et al. 2001). Il devient ainsi difficile de les utiliser pour modéliser la variabilité interne des événements de pluie pour des durées de l'ordre de 30 minutes (voir de 1 heure).



La variabilité des précipitations est composée de deux aspects difficilement dissociables à partir des observations dont on dispose classiquement. D'une part la variabilité de l'intermittence ou de l'occurrence, c'est à dire la présence ou non de pluie qui varie suivant non seulement la région et la période considérées mais aussi suivant l'échelle d'observation : il est absolument certain qu'il pleut à l'échelle du globe à un instant donné et assez peu probable qu'il pleuve à ce même instant à l'échelle d'observation d'un pluviomètre. D'autre part, au sein même des événements de pluie il convient de considérer la variabilité des intensités de pluie. Dans le domaine temporel, les moyens récents d'observation à très fine échelle, tels que les disdromètres, ont permis de réaliser des travaux qui distinguent ces deux aspects (Molini et al. 2009). Dans le domaine spatial, l'étude est rendue plus complexe car les observations dont on dispose, réseaux radars ou observations satellitaires, correspondent à des résolutions grossières (temporelle ou spatiale) qui mélangent pluie et non pluie. Cela complexifie ainsi l'analyse de la structure intra événement.

### **1.2 Cadre de la thèse**

Les activités du LATMOS relatives aux précipitations concernent plus particulièrement le développement instrumental, l'observation, et la caractérisation des précipitations. Parmi les activités en physique de la mesure, on peut citer les comparaisons modèles/observations, la fusion de données issues de capteurs différents, le développement des algorithmes inverses pour les capteurs spatiaux, la conception de systèmes de mesure terrestres ou spatiaux (Mégha tropiques, DYCECT), l'évaluation des erreurs de mesures induites par les problèmes de beamfilling, la validation de grandeurs estimées par télédétection spatiale à partir d'observations effectuées par des capteurs in-situ.

La caractérisation des précipitations repose sur des analyses de données à différentes échelles spatiales (locales, régionales ou globales). L'analyse de séries longues observées de manière intermittente ou ponctuelle est particulièrement délicate en particulier dans le contexte du changement climatique. Des méthodes d'analyse adéquates doivent être développées pour étudier : l'impact des caractéristiques de l'observation par satellite (résolution spatiale, temporelle, répétitivité faible) sur la caractérisation globale des précipitations, la possibilité de détecter des tendances correspondant au changement climatique. L'ensemble de ces travaux liés

au développement de systèmes d'observation ou à l'interprétation des observations réalisées se heurtent constamment au manque de connaissances relatif à l'hétérogénéité et à l'intermittence du champ précipitant à petite échelle et aux relations inter-échelles.

Les développements de systèmes de mesure, que ce soit les développements d'instruments ou des méthodes d'inversion associées nécessitent de disposer de données simulées (ou synthétiques). La mise au point ou la validation d'indicateurs fiables pour caractériser l'évolution de la fréquence ou de l'intensité des précipitations à partir d'observations nécessairement partielles et entachées d'erreur, nécessite de disposer de données synthétiques suffisamment représentatives de la réalité. Les travaux présentés dans ce manuscrit, destinés à l'amélioration des connaissances relatives à la variabilité des précipitations, abordent la modélisation statistique, la génération de champs synthétiques multi-échelles, la faisabilité d'un système d'observation.

Il n'existe pas à notre connaissance de simulateurs de champs précipitants apte à reproduire la variabilité des précipitations à fine échelle, inférieurs à 1 minute, et valable sur une vaste gamme d'échelle spatiale et temporelle, incluant les échelles internes aux événements de pluie, mais aussi l'intermittence du processus. Les générateurs développés dans le cadre de cette thèse sont destinés à être mis en ligne et utilisés par l'ensemble de la communauté scientifique pour tout type d'application.

Une bonne connaissance des pluviométries sur une zone donnée nécessite un système de mesure fournissant des observations denses sur cette zone. Bien qu'il existe des réseaux opérationnels, déployer un réseau dense de pluviomètres ou un radar coûte cher, en particulier en entretien. De telles données ne sont disponibles que sur des zones limitées de la planète pour des raisons géographiques (relief) ou économiques. D'après Overeem et al. (2013), il apparaît, de plus, un déclin de 50% du nombre de pluviomètres dans la période 1989-2006 pour le réseau GPCC en Afrique et une baisse de 50% du nombre de relevés journaliers valides dans la période 2000-2007 en Asie. L'étude de faisabilité concernant un système de mesure hyperfréquence, réalisée dans le cadre de cette thèse est destinée à remédier à cette difficulté en proposant un système d'observation reposant sur la mesure opportuniste des affaiblissements atmosphériques.

Les travaux réalisés dans le cadre de cette thèse s'appuient sur des études récentes menées par l'équipe SPACE du LATMOS. Dans la thèse qu'il a soutenue en 2011, Sébastien Verrier a étudié les propriétés statistiques des champs précipitants dans le cadre des cascades multiplicatives multifractales. Les travaux de simulation réalisés dans le cadre de cette thèse s'appuient en partie sur les conclusions qu'il a obtenues, à savoir:

- les différentes gammes d'invariance d'échelle observées dans le spectre d'énergie (cf Figure 1.5) sont liées à l'alternance de périodes pluvieuses et de périodes sèches (Verrier 2011)
- en restreignant l'analyse à des périodes de pluie sans intermittence, les trois paramètres du modèle multifractal obtenus sont identiques en temps et en espace.

Les outils développés par Sébastien Verrier pour réaliser les analyses multifractales ou pour générer des évènements de pluie ont servi de point de départ à cette thèse.

Un système d'observation basé sur la mesure opportuniste de différents affaiblissements issus de liaisons micro-ondes Terre – Satellites en bande Ku (Barthes et Mallet, 2013) a été développé au LATMOS. Dans le cadre de son doctorat, François Mercier a mis au point une méthode de traitement permettant d'estimer le champ de pluie à partir de ces liaisons. La méthode a montré sa pertinence lors de deux campagnes dont la campagne HYMEX (Mercier et al., 2015). L'intérêt de ce type d'observations, outre sa fine résolution temporelle (10 s), est de réaliser des mesures en incidence oblique permettant une vision dans le plan horizontal de quelques dizaines de km<sup>2</sup> à partir d'un site unique. La méthode d'assimilation variationnelle de type 4D VAR développée permet de relier entre elles les observations successives et restitue ainsi le champ de pluie en le forçant à présenter une cohérence spatiale et temporelle. L'ensemble de l'étude réalisée par Mercier et al. (2015) a servi de point de départ pour généraliser le concept à un ensemble de capteurs.

### **1.3 Objectif de la thèse**

Les travaux présentés dans ce manuscrit concernent l'étude de la variabilité des précipitations dans une vaste gamme d'échelle spatiale et temporelle. Une attention particulière est portée sur

les échelles les plus fines, c'est à dire sur la variabilité interne des évènements de pluie et sur le caractère intermittent du processus de pluie

Dans la première partie de notre étude, nous avons développé un simulateur stochastique de séries temporelles de taux précipitants. Il existe de nombreuses études consacrées à la modélisation des précipitations, cependant les échelles de connaissances sont liées aux échelles d'observations. Or dans la plupart des cas les échelles d'observations sont de l'ordre de l'heure et du jour avec dans de rares cas des observations à la minute. Pour cette raison, la plupart des modèles existants permettent de décrire la variabilité de la pluie à ces échelles d'observation (heure, jour). Dans notre étude et grâce à l'utilisation d'un disdromètre à haute résolution temporelle, nous avons pu mettre en œuvre des méthodologies et développer des modèles à des résolutions plus fines, nous permettant ainsi de distinguer les propriétés du support de pluie (inter évènements) de celles de la variabilité intra événement. Ici l'objectif est de produire des séries chronologiques de plusieurs années à la résolution de 15 secondes. Les caractéristiques des séries simulées doivent être similaires à celles observées aussi bien à l'échelle de simulation (15s) qu'après dégradation à des échelles plus grossières (horaire et journalière) garantissant ainsi le bon respect des lois d'échelle.

La méthodologie développée pour la réalisation du simulateur stochastique de séries temporelles de taux précipitants est ensuite étendue à la génération de cartes 2D de lames d'eau. Des observations à la résolution de  $1 \times 1 \text{ km}^2$  issues des radars de Météo France ont été analysées. Le simulateur stochastique multi-échelle 2D ainsi développé permet de reproduire les caractéristiques géostatistiques et topologiques à cette même échelle.

Les cartes de lames d'eau simulées sont utilisées pour la mise au point d'un nouveau système d'observation des précipitations en milieu urbain. Le principe de ce système repose sur l'utilisation de mesures opportunistes de l'affaiblissement subit par les ondes radios émises par les satellites géostationnaires TV-SAT dans la bande 10.7-12.7 GHz (bande Ku). Une campagne de mesure (Mercier et al. 2015) a permis de valider le principe de fonctionnement du système. Il s'agit de quantifier l'apport de mesures multi-sites. De façon plus spécifique on suppose que les terminaux de réception TV-SAT installés en ville chez les particuliers sont capables de mesurer de tels affaiblissements (cf. section 5.1 pour plus de détails). Il s'agit alors d'utiliser ces

affaiblissements pour restituer les champs de pluie à partir d'une méthode d'assimilation. Les observations d'affaiblissements atmosphériques en bande Ku n'étant pas disponibles, des données synthétiques sont produites en utilisant le simulateur stochastique de cartes 2D de lames d'eau. On valide ainsi le mode de fonctionnement multi-site. Ces données synthétiques permettent la mise au point des méthodes de traitement et l'évaluation des performances potentiellement obtenues par différentes configurations du système d'observation et de conditions météorologiques (préparant ainsi une campagne de mesure à venir).

### **1.4 Plan de la thèse**

Le chapitre 2 présente des aspects théoriques, le formalisme multifractal (la définition des multifractals, les techniques de simulation et d'analyse multifractales, ...) ainsi que le générateur multifractal universel (Schertzer et Lovejoy, 1987) qui a été utilisé par la suite. L'analyse de simulations produites par ce générateur dans différentes conditions d'utilisation permet de cerner ses performances et ses limitations.

Le chapitre 3 présente une analyse détaillée des séries temporelles de précipitation observées par un disdromètre utilisées pour développer notre modèle. Nous exposons ensuite la méthodologie que nous avons proposée pour la génération des séries temporelles de précipitations réalistes. Nous présentons ainsi deux versions du générateur de séries chronologiques avec pour chacun leurs avantages et leurs inconvénients respectifs. Différentes propriétés des séries observées et simulées sont considérées pour valider le simulateur (invariance d'échelle, occurrences, régime de scaling, spectre, distribution des taux de pluie, relation intensité/durée des événements...) et définir ses limites.

Le chapitre 4 présente l'extension du simulateur stochastique de séries temporelles de taux précipitants à la génération de cartes 2D de lames d'eau. Une analyse détaillée des cartes de lames d'eau issues des radars météo-France est un préalable à la modélisation. De même que pour les simulations en 1D, les propriétés statistiques des cartes simulées sont considérées pour la validation du simulateur (structure spatiale, occurrence, relation surface d'évènement-intensité moyenne, spectre, ...) ainsi que pour définir ses limites d'utilisation.

Le chapitre 5 est dédié à une étude de faisabilité d'un système innovant d'observation des précipitations en milieu urbain à l'aide des liaisons satellite - récepteur TV. La méthodologie et les résultats obtenus pour différentes configurations réalistes sont détaillés.

Le chapitre 6 présente les conclusions et les perspectives de ce travail.



# Chapitre 2

## FORMALISME ET TECHNIQUES MULTIFRACTALES

### Sommaire

---

2	FORMALISME ET TECHNIQUES MULTIFRACTALES .....	33
2.1	Théorie des fractals .....	33
2.1.1	Objets fractal et multifractal.....	33
2.1.2	Fractals aléatoires et fractals déterministes .....	34
2.1.3	Résolution et échelle .....	34
2.1.4	Dimension et co-dimension fractales .....	34
2.1.5	Les Multifractals .....	36
2.1.5.1	Notion de singularité .....	36
2.1.5.2	Fonction d'échelle des moments .....	37
2.1.6	Les cascades multiplicatives .....	38
2.1.6.1	Cascade continue et discrète.....	38
2.1.7	Modèle multifractal universel : .....	40
2.1.8	Champs conservatifs / champs non conservatifs .....	41
2.2	Simulation et analyse multifractale : modèle FIF .....	42
2.2.1	Caractéristiques du générateur multifractal universel .....	42
2.2.1.1	Simulation des zéros (intermittence de la pluie).....	44
2.2.1.2	Produits de convolution .....	45
2.2.2	L'Analyse multifractale .....	47
2.2.2.1	Limitation de la taille de l'échantillon et observation des extrêmes.....	48
2.2.2.2	Divergence des moments.....	50
2.2.2.3	Estimation des moments.....	51
2.3	Restitution des paramètres multifractals en pratique .....	52
2.4	Généralisation de cartes en dimension 2.....	54
2.5	Conclusion .....	54

---



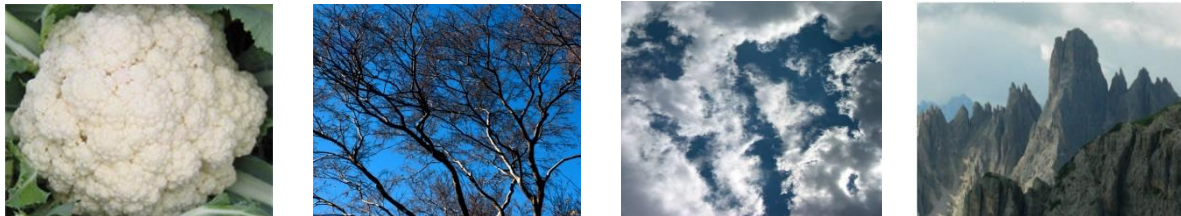


## 2 FORMALISME ET TECHNIQUES MULTIFRACTALES

### 2.1 Théorie des fractals

#### 2.1.1 Objets fractal et multifractal

Le concept de fractal a été introduit par Benoît Mandelbrot en 1975 dans le but d'étudier les processus et les formes irrégulières ou / et fragmentées que l'on trouve dans la nature ou dans le domaine des mathématiques. On dit qu'un objet est fractal si sa structure est invariante par changement d'échelle, autrement dit un objet fractal a des détails similaires à des échelles arbitrairement petites ou grandes ; d'où la propriété la plus connue des fractales qui est leur autosimilarité (Mandelbrot, 1967; 1974 ; 1983 ; 2006). Ceci signifie, qu'à différentes échelles, une partie de la fractale sera similaire à la fractale dans son intégralité et ce, aussi loin que l'on puisse « zoomer » : on parle donc de propriété d'invariance d'échelle (Falconer, 1990). D'où la définition d'objet fractal : « un objet fractal est un objet mathématique généré par un processus itératif, et qui présente un comportement invariant d'échelle ». Des formes fractales (sur une certaine gamme d'échelle) sont facilement observables dans la nature (Figure 2.1).



*Figure 2.1 Exemple d'objets présentant des propriétés fractales dans la nature.*

Les modèles (multi/mono) fractals répondent à des problématiques liées aux extrêmes et aux relations inter échelles, ils sont de ce fait utilisés en géophysique. Avant d'aller plus loin, il convient donc de définir plus précisément les concepts de fractal et de multifractals et leurs liens avec la notion d'invariance d'échelle.

### **2.1.2 Fractals aléatoires et fractals déterministes**

Deux classes de fractals peuvent être distinguées : les fractals déterministes et les fractals aléatoires (Falconer, 1990 ; De Lima, 1998). Les fractals déterministes sont construits à partir de règles déterministes, l'exemple le plus classique est celui de la poussière de Cantor présentée à la Figure 2.2 dans laquelle on observe que l'objet fractal est semblable à chacune de ses parties à une contraction près (autosimilarité).

La notion de fractal déterministe peut être étendue à des processus stochastiques. Dans ce cas, l'autosimilarité s'interprète au sens statistique. Les fractals aléatoires présentent généralement un aspect moins régulier plus en adéquation avec les phénomènes naturels. Ce sont donc ces derniers qui sont mis en œuvre dans les modélisations en géophysique.

### **2.1.3 Résolution et échelle**

Un objet fractal représenté dans un espace de dimension  $D$  est en général défini à l'aide d'une grille dont les arêtes ont une longueur  $l$ . Il est possible d'étudier cet objet à des échelles plus ou moins fines en fonction de la taille de la grille. Si l'espace contenant l'objet est de dimension géométrique  $D$ , on peut découper l'espace en  $\lambda^D$  cases, soit  $\lambda$  intervalles le long d'une dimension de l'espace. En notant  $L$  la longueur de l'objet le long d'une arête on définit la résolution  $\lambda$  à laquelle est vue l'objet :  $\lambda = \frac{L}{l}$

On définira par le terme « échelle » l'inverse de la résolution. Avec cette notation, un objet à la résolution 1 est seulement composé d'un seul pixel (case) alors qu'un objet à la résolution  $\lambda$  est composé de  $\lambda^D$  pixels.

Bien que nous ayons présenté ici le concept de résolution à partir de grandeurs spatiales, cette notion s'applique bien évidemment à des grandeurs temporelles :  $\lambda = \frac{T}{t}$

Dans lequel  $T$  désigne la durée totale de la série considérée et  $t$  la période d'échantillonnage.

### **2.1.4 Dimension et co-dimension fractales**

Les objets fractals sont trop irréguliers pour être décrits simplement au moyen de la géométrie usuelle : ils sont caractérisés par une dimension non entière appelée dimension fractale. Le terme

«fractale» (du latin « fractus » et qui veut dire brisé) fut introduit par Mandelbrot pour désigner un type d'objet dont l'irrégularité le distingue des figures géométriques euclidiennes telles qu'une droite, un carré, un cercle, etc. En géométrie euclidienne, les figures ont une dimension entière : 0 pour un point, 1 pour une courbe, 2 pour une surface, etc. En revanche, la dimension d'une fractale peut prendre des valeurs non entières.

L'exemple de la poussière de Cantor (Figure 2.2) permet de mieux comprendre la notion de dimension fractale. On notera par la suite  $D_f$  cette dimension. La Figure 2.2 est construite de manière itérative : à chaque itération, on découpe chaque segment en trois parties de même longueur et on enlève celle du milieu. Le segment initial à l'étape  $n = 0$  étant de dimension 1, on lui enlève une infinité de ses parties lorsque  $n$  tend vers l'infini. Ainsi, on a bien l'intuition que l'on obtiendra un objet de dimension fractale inférieure à 1.

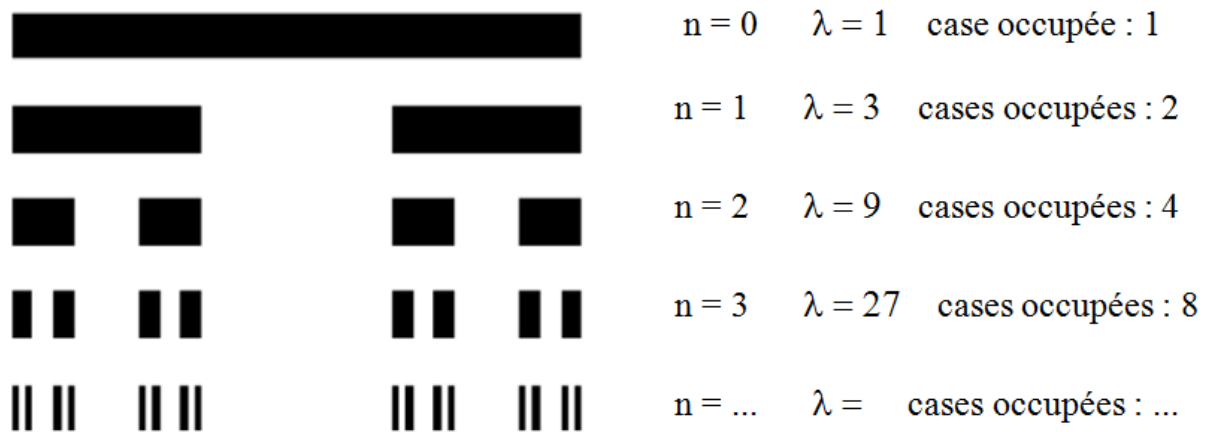


Figure 2.2 Calcul de la dimension fractal sur l'exemple de la poussière de cantor.

D'un point de vue pratique, pour déterminer la dimension fractale  $D_f$  on utilise généralement la méthode dite de comptage de boîtes (box counting) dans laquelle l'objet est maillé par une grille dont la longueur d'une arête vaut  $l$  :

$$D_f = \lim_{l \rightarrow 0^+} \frac{-\ln(N(l))}{\ln(l)}$$

$N(l)$  désigne le nombre de boîtes de taille  $l$  nécessaires pour recouvrir l'objet. L'idée de la méthode est d'approximer la relation entre  $l$  et  $N$  par une loi puissance de la forme :  $N(l) = l^{-D_f}$

Dans le cas de la poussière de Cantor, le nombre de segments noirs à l'itération  $n$  est  $N = 2^n$  et la taille de la grille vaut  $l = 1/3^n$ . Il en résulte que  $D_f = \frac{\log(2)}{\log(3)} = 0.6309$  et possède donc une valeur très inférieure à l'unité.

Notons que la co-dimension fractale  $C_f$ , est définie par  $C_f = D - D_f$  (Mandelbrot, 1967; 1983; Falconer, 1990) où  $D$  est la dimension géométrique de l'espace contenant la fractale. La co-dimension est liée à la probabilité qu'un pixel de l'espace géométrique (de dimension  $D$ ) appartienne à l'objet fractal :

$$\Pr(\text{Pixel} \in \text{Objet}) = \frac{\text{Nbre Pixels} \in \text{Objet}}{\text{Nbre Pixels Total}} = \frac{\lambda^{D_f}}{\lambda^D} = \lambda^{-C_f}$$

Nous généraliserons ce concept de co-dimension dans la partie suivante relative aux multifractals.

## 2.1.5 Les Multifractals

### 2.1.5.1 Notion de singularité

Les champs géophysiques (pression, humidité, température, ...)  $\varepsilon_\lambda$  (champ à la résolution  $\lambda$ ) sont en général des grandeurs valuées. La notion de fractal ne peut donc pas s'appliquer directement sur ces grandeurs puisqu'il n'est pas possible dans ce cas d'appliquer le concept d'appartenance ou non à un ensemble. Ce problème peut être contourné aisément en appliquant un seuil  $S_\lambda$  au champ considéré : toute valeur supérieure à ce seuil sera considérée comme appartenant à l'objet tandis qu'une valeur inférieure ne le sera pas. Il devient ainsi possible d'associer une dimension fractale (et une co-dimension) au champ. Cependant, la dimension fractale dépend en général de la valeur du seuil, lequel est lui-même dépendant de la résolution. On dit dans ce cas que le champ est multifractal. Afin de rendre ce seuil indépendant de la résolution, on préfère lui substituer la notion de singularité  $\gamma$  définie de la façon suivante (Schertzer et al., 2002) qui permet de définir non pas un seuil mais une famille de seuil  $S_\lambda = \lambda^\gamma$

Avec cette définition la probabilité qu'un pixel appartienne à l'objet multifractal est :

$$\Pr(\varepsilon_\lambda > \lambda^\gamma) \sim \lambda^{-c(\gamma)} \quad \text{Équation 2.1}$$

Ainsi la fonction de co-dimension  $c(\gamma)$  ne dépend plus de la résolution, elle dépend seulement du seuil de singularité. Cette fonction est croissante et convexe, ainsi lorsque le seuil de singularité augmente la dimension fractale a tendance à décroître.

### 2.1.5.2 Fonction d'échelle des moments

En considérant le champ géophysique  $\varepsilon_\lambda$  du point de vue d'une variable aléatoire de densité de probabilité  $p(x)$ , il est possible d'exprimer son moment d'ordre  $q$  :

$$E\{\varepsilon_\lambda^q\} = \int_{-\infty}^{+\infty} x^q p(x) dx \quad \text{Équation 2.1b}$$

Sachant que pour une variable aléatoire  $x$  à valeurs positives, la densité de probabilité  $p(x)$  et les moments statistiques sont reliés par une transformation de Mellin. Cette relation suggère que les propriétés multi-échelle duales vont apparaitre pour les moments (Parisi et Frisch, 1985 ; Verrier, 2011 ainsi que Schertzer et al. 2002, section 4.2).

A partir des deux équations 2.1 et 2.1b, on montre qu'une bonne approximation du moment d'ordre  $q$  (Schertzer et al., 2002) est donnée par :

$$E\{\varepsilon_\lambda^q\} \approx \lambda^{K(q)} \quad \text{Équation 2.2}$$

$K(q)$  est dénommée fonction d'échelle des moments. Tout comme la fonction de co-dimension elle caractérise les propriétés statistiques du champ.

Les fonctions  $K(q)$  et  $c(\gamma)$  peuvent être liées mutuellement par une transformée de Legendre comme suit (Parisi et Frisch, 1985) :

$$c(\gamma) = \max_q \{q\gamma - K(q)\} \quad \text{et} \quad K(q) = \max_\gamma \{q\gamma - c(\gamma)\}$$

$K(q)$  est une fonction convexe qui possède deux points invariants :  $K(0) = K(1) = 0$ .

### **2.1.6 Les cascades multiplicatives**

Les cascades multiplicatives permettent de créer de façon simple une dépendance inter échelles et d'obtenir ainsi des distributions vérifiant l'équation 2.1. L'idée originelle provient de Richardson (1922) qui dans le domaine de la turbulence a imaginé des structures de plus en plus petites imbriquées les unes dans les autres. Cette idée a ensuite été reprise par Kolmogorov en 1941 qui a montré que dans le cas de la turbulence il fallait considérer la variable flux d'énergie. En 1966, Yaglom a introduit le modèle de cascade multiplicative, ce modèle a ensuite été repris et développé par plusieurs chercheurs (Schertzer et al., 1993).

#### **2.1.6.1 Cascade continue et discrète**

Une cascade multiplicative discrète  $\varepsilon_{\lambda_n}$  est composée d'une série de variables aléatoires indépendantes et identiquement distribuées à valeurs positives ou nulles notées  $\mu_\varepsilon$  dénommées « incréments multiplicatifs ». Une telle cascade est construite en passant d'une échelle donnée à une échelle plus fine par un processus itératif. Ce passage d'une résolution  $\lambda_l$  à une résolution  $\lambda_n$  est représenté par la formule suivante :

$$\varepsilon_{\lambda_n} = \varepsilon_{\lambda_l} \cdot \prod_{i=l+1}^n \mu_{\varepsilon_i}$$

Dans un tel processus itératif, les valeurs de  $\lambda_n$  constituent une suite discrète, soit  $\lambda_n = \lambda_l^n$ . En pratique nous prendrons  $\lambda_l = 2$ . Le processus de génération d'une cascade discrète est représenté à la Figure 2.3 en dimension 1 et 2.

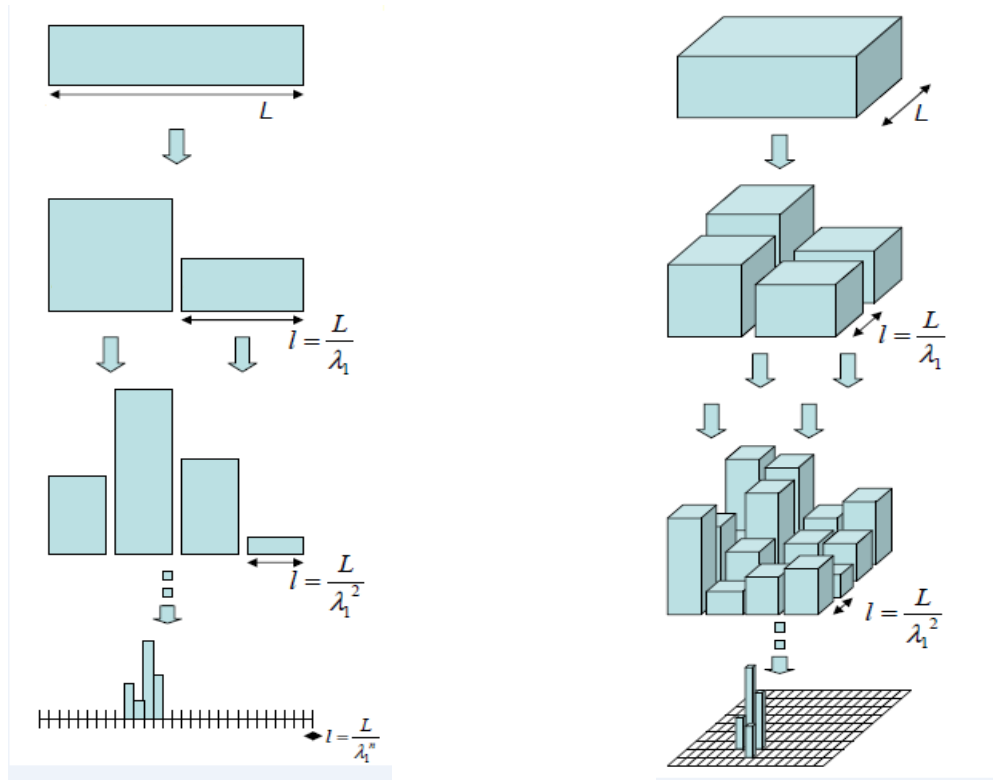


Figure 2.3 Schéma de principe des cascades multiplicatives discrètes 1D et 2D

La construction précédente peut se généraliser à celles de cascades multiplicatives continues en échelle. L'idée sous-jacente est à partir d'une cascade discrète de densifier le nombre de cascades entre les échelles  $\lambda_n$  et  $\lambda_{n+1}$  par un facteur  $N$  puis en faisant tendre  $N$  vers l'infini (ce qui revient à faire tendre le rapport  $\lambda_n / \lambda_{n+1}$  vers 1). En poursuivant ce raisonnement, on arrive à des singularités infinies (Bernadara et al. 2007). Il est donc nécessaire d'imposer une condition de non-divergence ainsi que de conservation de la moyenne. Lorsque la résolution augmente, lors de la simulation, une solution classiquement retenue consiste à imposer la conservation exacte de la moyenne sur toutes les réalisations. Cette solution est baptisée « conservation micro-canonique ». Elle présente toutefois l'inconvénient d'imposer une dépendance entre les incréments multiplicatifs générés pour passer d'une étape de la cascade à une autre. Pour cette raison Mandelbrot (1974) a proposé une condition moins restrictive dite « conservation canonique » qui consiste à assurer uniquement la conservation de l'espérance mathématique, c'est-à-dire :

$$\forall \lambda, E\{\varepsilon_\lambda\} = \text{Constante}$$

Équation 2.3



### 2.1.7 Modèle multifractal universel :

Dans le cas général, les modèles multifractals sont déterminés par leur fonction d'échelle des moments  $K(q)$ . Comme nous l'avons vu précédemment, cette fonction n'impose que peu de contraintes (convexité,  $K(0) = K(1) = 0$ ). Il en résulte qu'un grand nombre de paramètres peut être nécessaire pour la définir induisant ainsi une certaine complexité pour la modéliser. En pratique on souhaite plutôt un nombre réduit de paramètres ayant chacun une signification propre. L'alternative proposée par (Schertzer et Lovejoy, 1987) explore l'existence de générateurs attractifs stable qui mène à une classe de générateurs multifractals dénommée par leurs auteurs Multifractals Universels (MU) et nécessitant seulement 2 paramètres. Les auteurs ont montré que le processus de cascade multiplicative converge sous certaines conditions vers la classe des multifractals universels (Schertzer et Lovejoy, 1987 ; Schertzer et Lovejoy, 1997). Cette notion d'universalité des processus multifractal a également été abordée dans plusieurs études notamment par (Gupta et Waymire, 1993).

Dans le cadre des multifractals universels, la fonction d'échelle des moments est définie ainsi :

$$\begin{cases} K(q) = \frac{C_1}{\alpha-1} (q^\alpha - q) & \alpha \neq 1 \\ K(q) = C_1 q \cdot \ln(q) & \alpha = 1 \end{cases} \quad \text{Équation 2.4}$$

Et la fonction de co-dimension est définie par :

$$\begin{cases} c(\gamma) = C_1 \left( \frac{\gamma}{C_1 \alpha'} + \frac{1}{\alpha} \right)^{\alpha'} & \alpha \neq 1 \\ c(\gamma) = C_1 \exp\left(\frac{\gamma}{C_1} - 1\right) & \alpha = 1 \end{cases} \quad \text{Équation 2.5}$$

Avec

$$\frac{1}{\alpha} + \frac{1}{\alpha'} = 1$$

Le cas  $\alpha = 0$  correspond au cas monofractal, et  $\alpha = 2$  correspond au cas log-normal.

Les deux paramètres  $\alpha$  et  $C_1$  possèdent une signification propre :

$C_1$  : Il correspond à l'intermittence moyenne. Il mesure la rareté moyenne du champ. Plus précisément, c'est la co-dimension correspondant à la singularité moyenne du champ conservatif, sa valeur est comprise entre 0 et  $D$ .

$\alpha$  : Index de multifractalité. Il mesure la vitesse à laquelle évolue l'intermittence quand la singularité diffère légèrement de la singularité moyenne du champ.

### **2.1.8 Champs conservatifs / champs non conservatifs**

Comme on l'a dit dans l'introduction, selon la gamme d'échelle considérée, on caractérise plus spécifiquement les effets (de la plus grande échelle à la plus petite) de la variabilité climatique, du forçage barocline, de la turbulence, du mélange des gouttes, de la coalescence, du break-up, ... Selon les échelles considérées, on fera appel à un modèle multifractal conservatif ou non. Pour la pluie, aux grandes échelles, le comportement est conservatif.

Le modèle multifractal décrit précédemment permet de représenter des champs dits conservatifs, c'est à dire que leur espérance statistique est indépendante de l'échelle  $\lambda$  (Equation 2.3).

Cependant, des grandeurs atmosphériques associées à la turbulence, notamment la mesure du vent ou la concentration en polluant, sont non conservatives. Comme cela a été montré dans les thèses de Verrier et Montera, une prise en compte adéquat du support de pluie permet de faire ressortir les propriétés non conservatives de la pluie aux fines échelles. Quand, comme on le fait ici, on s'intéresse aux petites échelles (associées à la turbulence ainsi qu'à la coalescence et au break-up des gouttes), il est donc nécessaire d'utiliser un modèle non conservatif permettant de prendre en compte la variation de la moyenne du phénomène en fonction de l'échelle. Schertzer et Lovejoy (1987) ont proposé une extension du modèle MU permettant d'assurer une certaine non conservativité du champ. Cette extension consiste à intégrer fractionnairement à l'ordre  $H$  le champ (Schertzer et Lovejoy, 1991). Le champ résultant  $R_\lambda$  est un champ multifractal non conservatif appelé par ses auteurs modèle FIF (Fractionnally Integrated Flux).

La loi d'invariance d'échelle de sa fonction de structure est donnée par :

$$\langle |\Delta R_\lambda(\Delta t)|^q \rangle \sim \lambda^{-\zeta(q)} \quad \text{Équation 2.6}$$

Où  $\Delta R_\lambda(\Delta t)$  correspond à un incrément du champs « non conservatif » sur un intervalle de durée  $\Delta t$ .  $H$  est le degré de non stationnarité et  $\zeta(q)$  est l'exposant de la fonction de structure d'ordre  $q$  donné par:

$$\zeta(q) = qH - K(q)$$

On montre que plus H est fort, plus le processus est lisse.

Le modèle FIF complet fait donc intervenir trois paramètres ( $\alpha$ ,  $c_1$  et H).

## 2.2 Simulation et analyse multifractale : modèle FIF

Afin de réaliser un simulateur basé sur une approche multifractale qui reproduise les propriétés observées des précipitations. La première étape de mon travail a consisté à investiguer les problèmes liés à l'analyse et l'estimation des paramètres du modèle FIF sur des séries de taux précipitants observées (cf. section 3.6). Une seconde étape a consisté à générer des séries et des champs multifractals ayant les propriétés souhaitées. Des études antérieures réalisées au sein de l'équipe ont conclu que le modèle FIF était bien adapté à la modélisation des précipitations, nous avons donc volontairement limité notre étude à ce modèle. Les codes Matlab utilisés pour la génération et l'analyse du modèle FIF sont ceux développés par Sébastien Verrier durant sa thèse (Verrier, 2011), eux-mêmes provenant des codes Mathematica fournis par S. Lovejoy.

Les paragraphes ci-après ont pour objectifs de mettre en lumière certains aspects relatifs au modèle FIF et éventuellement d'apporter quelques améliorations.

### 2.2.1 Caractéristiques du générateur multifractal universel

Le principe du générateur multifractal utilisé est succinctement décrit ci-après. Ce générateur permet de simuler (ou de synthétiser) des champs en dimension D quelconque connaissant les trois paramètres multifractals décrits plus haut et la taille L désirée du champ. La première étape consiste à générer un bruit blanc de Lévy  $l_\alpha(\vec{x})$  de paramètre  $\alpha$  obtenu à partir d'un bruit de loi uniforme et d'un bruit de loi exponentielle. Ce bruit est ensuite multiplié par un terme  $\sigma$  afin d'obtenir la valeur de  $C_1$  désirée (Pecknold, 1993 ;Lovejoy et Schertzer, 2010) :

$$\sigma = \left( \frac{c_1}{|\alpha-1|} \right)^{\frac{1}{\alpha}} \quad \text{Équation 2.7}$$

Le bruit de Levy est ensuite convolué par le noyau de convolution  $g_\alpha(x)$  suivant :

$$g_\alpha(\vec{x}) = \begin{cases} \frac{1}{|\vec{x}|^{\frac{D}{\alpha}}} & (|\vec{x}| \leq \lambda) \\ 0 & (|\vec{x}| > \lambda) \end{cases} \quad \text{Équation 2.8}$$

permettant d'obtenir le générateur  $\Gamma_\lambda$  :

$$\Gamma_\lambda(\vec{x}) = \sigma \times g_\alpha(\vec{x}) * l_\alpha(\vec{x}) = \sigma \int_{1 \leq |\vec{x}'| \leq \lambda} \frac{l_\alpha(\vec{x}')}{|\vec{x} - \vec{x}'|^\alpha} D^d \vec{x}' \quad \text{Équation 2.9}$$

Après exponentiation du générateur, le champ conservatif  $\varepsilon_\lambda$  est obtenu et correspond donc au modèle MU.

$$\varepsilon_\lambda(\vec{x}) = e^{\Gamma_\lambda(\vec{x})} \quad \text{Équation 2.10}$$

L'obtention d'un champ fractionnairement intégré  $R_\lambda$  à l'ordre H nécessite l'ajout d'une seconde convolution avec le noyau  $g_H(x)$  :

$$g_H(\vec{x}) = \frac{1}{|\vec{x}|^{D-H}} \quad \text{Équation 2.11}$$

$$R_\lambda(\vec{x}) = \int d^D \vec{x}' \cdot \frac{\varepsilon_\lambda(\vec{x}')}{|\vec{x} - \vec{x}'|^{D-H}} \quad \text{Équation 2.12}$$

$R_\lambda(\vec{x})$  correspond au modèle FIF, la Figure 2.4 montre le schéma général du modèle MU et ou du modèle FIF.

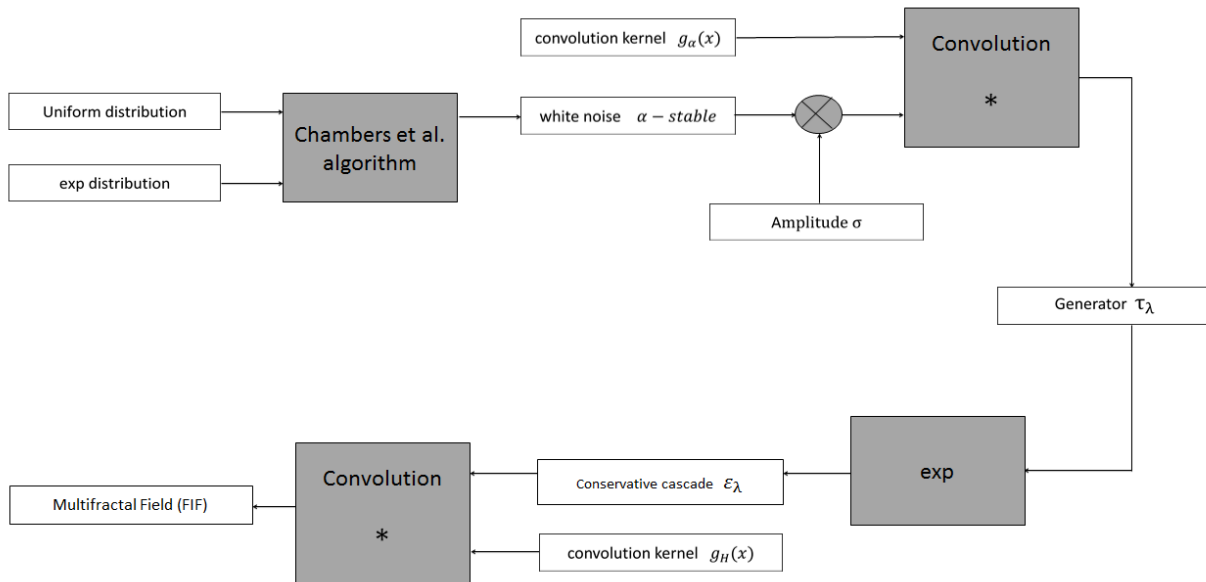


Figure 2.4 Schéma de principe du générateur multifractal (D'après Verrier, 2011)

L'extension à des champs anisotropes nécessite une modification adéquate des noyaux de convolution mais ne pose pas de problèmes particuliers (Schertzer et Lovejoy, 1985) ; elle n'est pas mise en œuvre dans le cadre de mes travaux.

Les paragraphes qui suivent abordent quelques aspects relatifs à la mise en œuvre de ce générateur.

### **2.2.1.1 Simulation des zéros (intermittence de la pluie)**

Bien que le modèle FIF permette de représenter correctement la variabilité de nombreux champs géophysiques, le problème se corse en ce qui concerne la modélisation des précipitations. En effet, à la différence des grandeurs géophysiques telles que la pression ou la température, la pluie est un phénomène intermittent dont la durée des épisodes est comprise entre quelques secondes et plusieurs heures. Il en découle que la modélisation de champs de pluie doit prendre en compte cette intermittence. En effet, il ne faut pas oublier qu'une série de taux précipitants mesurée par exemple avec un pluviomètre sur un laps de temps relativement long comportera sous nos latitudes environ 95% de zéros. (Cet ordre de grandeur est le résultat de l'intermittence de la pluie perçue à la résolution d'un pluviomètre qui est de l'ordre de 5 minutes.) Le modèle FIF dans sa version originale permet de représenter des champs multifractals de dimension quelconque mais sans valeur nulle. Il n'est donc pas directement adapté ni à l'analyse ni à la génération de champs intermittents.

Dans le cadre multifractal diverses approches ont été étudiées pour simuler des valeurs nulles (Olsson, 1998; Schmitt et al., 1998 ; de Montera et al., 2009 ; Gires et al., 2013). Certains chercheurs ont introduit le modèle «bêta», c'est à dire un modèle de cascade acceptant des incréments multiplicatifs nuls (Schertzer et al., 2002). Le modèle bêta-log-stable généralise la plupart des modèles multifractals proposés pour la simulation des précipitations. Ceux-ci sont aussi appelés modèles bêta-log-normale ou bêta-log-Levy, selon que le paramètre de stabilité  $\alpha$  est égal à 2 ou strictement inférieur (Gupta et al., 1996; Serinaldi, 2010; Veneziano et al., 2006; Veneziano et Lepore, 2012). Bien que ce modèle permette de générer des valeurs nulles, il ne permet pas d'obtenir des durées de pluie et de sécheresse réalistes (Schmitt et al., 1998). La généralisation de ce processus pour les cascades continues, tel que proposé par (Schmitt, 2014), fournit une modélisation plus réaliste, mais il n'est toujours pas en mesure de corriger l'incapacité intrinsèque du modèle à simuler les périodes sèches et humides de façon réaliste. Une autre approche utilisée par (De Montera et al., 2010) consiste à appliquer à la série simulée un seuil adéquat permettant ainsi d'obtenir le bon pourcentage de zéro. Bien que cette approche soit

simple elle ne permet pas non plus d'obtenir des durées de pluie et de sécheresse réalistes. Nous en concluons qu'il n'existe pas de modèle suffisamment réaliste concernant la génération des zéros dans le cadre des modèles multifractals et que des mesures spécifiques devront être mises en œuvre dans la suite de notre étude.

### 2.2.1.2 Produits de convolution

Ce paragraphe met en lumière un problème numérique du générateur multifractal tel qu'il est proposé dans sa version de base. Nous proposons ci-après une modification simple du mode de calcul afin de pallier ce problème.

Comme on l'a vu au début de la section (cf. figure 2.4), la génération de champs multifractal implique le calcul d'un ou de deux produits de convolution (selon que l'on génère un champs multifractal conservatif ou FIF). En pratique, le calcul est réalisé dans le domaine de Fourier permettant ainsi de transformer les produits de convolution en produits simples. Toutefois du fait de la taille finie de la série il apparaît un phénomène de « fenêtrage », notamment aux fréquences élevées. Ce phénomène est facilement identifiable sur le spectre des séries simulées sur lesquels un brusque changement de pente apparaît au-delà d'une certaine fréquence. Pour illustrer ce phénomène, j'ai simulé dans un premier temps des séries avec les paramètres suivants :  $\alpha=1.6$ ,  $C_1=0.1$  et  $H=0$  et une résolution  $\lambda = 14$  (la taille des champs multifractals simulés est donc  $n = 2^{14}$ ). Dans un second temps, j'ai fait une analyse spectrale de ces séries simulées. La Figure 2.5 montre le spectre pour une série (Figure 2.5 a) et le spectre moyen pour 100 séries (Figure 2.5 c). Il apparaît clairement qu'au-delà d'une certaine fréquence ( $\omega > 1$  le spectre n'est plus en loi puissance et par conséquent la série ne possède plus les propriétés d'invariance d'échelle requises. Bien que ce phénomène soit un artefact dû à un effet de bord il convient de le corriger au mieux.

Dans la version originelle, pour réaliser une série temporelle de taille  $M$ , on tire  $M$  réalisations d'une variable aléatoire de Levy que l'on convolue par l'intermédiaire d'une fonction noyau (Équation 2.8). La solution, alternative, adoptée ici consiste à :

- Tirer un nombre de variables de Lévy supérieur d'un facteur N à la taille de la série souhaitée
- Appliquer le noyau de convolution sur cette série.
- Appliquer une moyenne glissante de largeur N
- Décimer d'un facteur N

Une moyenne glissante s'apparente à un filtrage linéaire, elle n'est donc pas de nature à changer la probabilité suivie par les éléments de la série. Il ne reste plus qu'à utiliser l'Équation 2.13 caractérisant les variables  $\alpha$ -stables. On peut ainsi déduire la constante de normalisation supplémentaire à introduire, pour conserver les paramètres de la loi ayant servi à simuler les données :

$$\sum_{i=1}^n X_i = n^{\frac{1}{\alpha}} X \quad \rightarrow \quad X = \frac{\sum_{i=1}^n X_i}{n^{\frac{1}{\alpha}}} \quad \text{Équation 2.13}$$

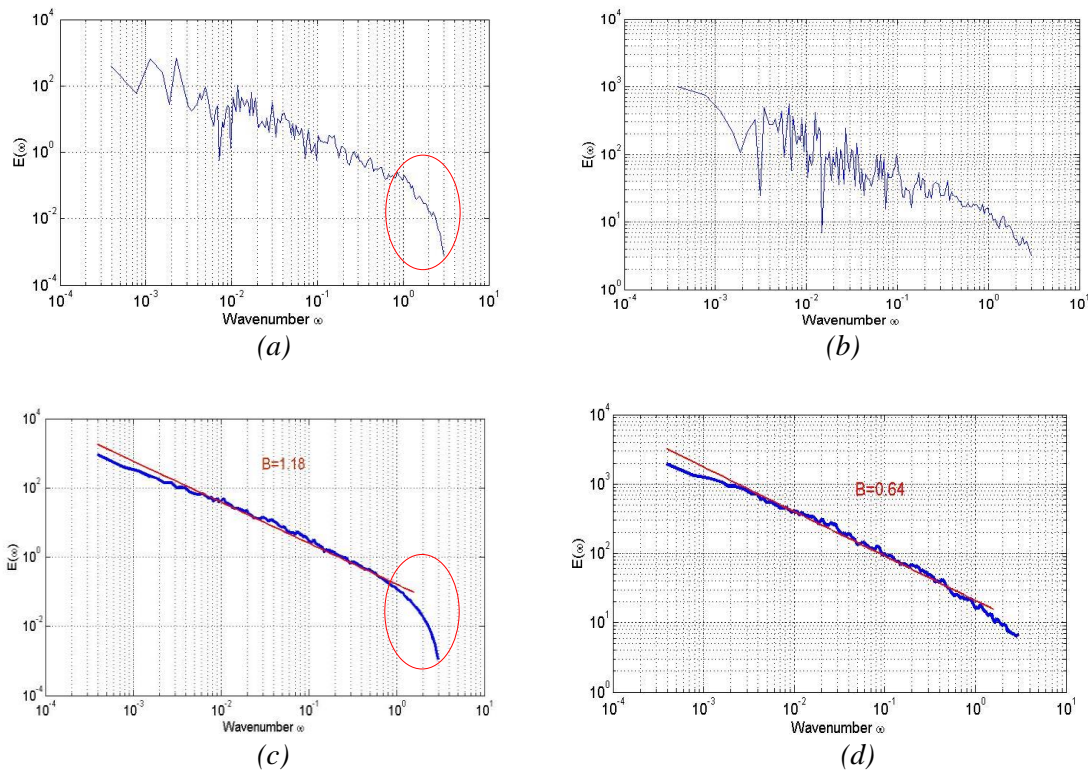
Pour illustrer l'amélioration obtenue nous avons généré des séries temporelles en utilisant les mêmes paramètres que précédemment mais en prenant  $N=8$ , c'est-à-dire que nous avons généré des séries de variables de Lévy 8 fois plus longues pour réaliser le produit de convolution. Les Figure 2.5 (a, b, c, d) montrent sur les spectres obtenus sur 1 réalisation avec le générateur originel et le générateur modifié (figures du haut) et les spectres moyens obtenus sur 100 réalisations (figures du bas). Il est clairement visible que le caractère en loi de puissance est beaucoup mieux préservé avec le générateur modifié ( $N=8$ ) y compris aux hautes fréquences.

Un autre effet bénéfique de cette correction est d'obtenir une pente plus proche de sa valeur théorique. En effet, compte tenu des paramètres utilisés pour  $\alpha$  et  $C_1$ , la pente théorique du spectre vaut :

$$\beta = 1 - K(2) + 2H = 1 - 0.17 + 0 = 0.83 \quad \text{Equation 2.13 b}$$

Cette valeur est à comparer avec les pentes obtenues à l'aide d'un ajustement sur les spectres moyens d'une part à partir du générateur d'origine (1.18) et la pente obtenue avec le générateur modifié (0.64) (Figure 2.5 c et Figure 2.5 d). Dans le premier cas, la pente est supérieure à l'unité ce qui impliquerait une valeur de H strictement positive, c'est-à-dire un processus non conservatif alors que nos simulations ont été réalisées dans le cas conservatif ( $H = 0$ ). Le

générateur modifié avec une valeur de pente beaucoup plus proche de la valeur théorique et inférieure à l'unité conduit bien à un champ conservatif des séries générées



**Figure 2.5** *a et b* Spectre obtenu sur une seule série simulée, *c et d* spectre moyen obtenu sur 100 séries simulées. Les deux figures de gauche sont obtenues avec le modèle original et les deux figures de droite sont simulées avec le modèle modifié ( $N=8$ )

### 2.2.2 L'Analyse multifractale

La méthodologie concernant l'analyse multifractale de séries ou de cartes est largement exposée dans la littérature. Il convient toutefois de préciser que l'intensité de pluie est non conservative ( $H > 0$ ). Cette non conservativité se traduit par une pente théorique du spectre  $\beta$  (cf. Equation 2.13 b) supérieure à 1 (voir le spectre de la figure 3.24). La modélisation requière donc une intégration fractionnaire d'ordre  $H$  en sortie de la cascade multiplicative continue  $\varepsilon_\lambda$  (modèle MU). Par conséquent, il n'est pas possible de réaliser directement l'analyse multifractale de champs non conservatifs pour déterminer les paramètres  $\alpha$  et  $C_1$  du modèle MU. Il faudrait donc en théorie dériver fractionnairement le champ non conservatif avant de faire l'analyse multifractale. En pratique cette dérivation est délicate d'autant plus que le paramètre  $H$  n'est pas forcément connu avec précision. Lavallée et al., (1993) ont montré qu'une bonne approximation



du champ conservatif à la résolution la plus fine lorsque  $H$  est proche de 0.5 (ce qui on le verra est notre cas) est la valeur absolue de la dérivée du champ non conservatif :

$$\varepsilon_{\lambda_{max}} \approx \frac{|\Delta R_{\lambda_{max}}(\Delta t)|}{\Delta t} \quad \text{Équation 2.14}$$

Ainsi chaque fois que nous aurons affaire à un champ non conservatif, nous appliquerons systématiquement la formule ci-dessus afin d'estimer le champ conservatif sous-jacent  $\varepsilon_\lambda$ .

L'analyse multifractale du champ conservatif, c'est-à-dire l'estimation des paramètres  $\alpha$  et  $C_1$  fait appel à la détermination préalable de la fonction d'échelle des moments  $K(q)$  (Équation 2.4), elle-même estimée à partir de l'estimation des moments du champ conservatif (Équation 2.2).

Un problème fondamental dans l'analyse multifractale de champs observés est l'estimation de la fonction  $K(q)$  qui comme nous l'avons vu nécessite l'estimation des moments (tout au moins jusqu'à un certain ordre  $q_{max}$ ). Il est donc primordial d'avoir une idée de la valeur de  $q_{max}$  utilisable en pratique. Les deux paragraphes ci-après permettent de mettre en lumière quelques aspects importants concernant  $q_{max}$ .

### 2.2.2.1 Limitation de la taille de l'échantillon et observation des extrêmes

Nous avons vu que l'on peut caractériser un processus multifractal par sa fonction d'échelle des moments  $K(q)$  (Équation 2.4) qui est directement liée au régime de scaling et qui permet d'estimer les paramètres multifractals (Veneziano et Furcolo, 2009 ; Villarini et al., 2007). Nous rappelons cette équation ci-après :

$$E\{(\varepsilon_\lambda)^q\} \propto \lambda^{K(q)} \quad \text{Équation 2.15}$$

Il est bien évident que la taille de l'échantillon peut limiter l'analyse statistique des caractéristiques des séries observées ou simulées. Pour les processus multifractals, cela se traduit par une singularité maximum  $\gamma_s$  et de façon duale par un ordre de moment  $q_s$  au-delà duquel les moments empiriques ne correspondent plus aux moments statistiques. En effet les moments d'ordres supérieurs sont associés à des singularités beaucoup trop rares et donc sous représentées dans l'échantillon. Par conséquent, on ne doit pas utiliser les moments empiriques d'ordre  $q > q_s$  pour estimer la fonction d'échelle des moments  $K(q)$  (Lovejoy et Schertzer, 2007).

Plus précisément, on considère  $N_s$  réalisations indépendantes à la résolution  $\lambda$ . Dans un espace de dimension  $D$ , nous avons  $\lambda^D$  valeurs par réalisation. La singularité maximum  $\gamma_s$  que l'on peut s'attendre à observer est définie par :

$$Pr(\varepsilon_\lambda \geq \lambda^{\gamma_s}) \approx \frac{1}{N_s \lambda^D} \quad \text{Équation 2.16}$$

Par introduction de la notion de dimension d'échantillon  $D_s$  :  $N_s = \lambda^{D_s}$ , il vient :

$$Pr(\varepsilon_\lambda \geq \lambda^{\gamma_s}) \approx \lambda^{-(D+D_s)} = \lambda^{-c(\gamma_s)} \quad \text{Équation 2.17}$$

Avec :  $c(\gamma_s) = D + D_s$

La fonction de co-dimension est donc bornée à la valeur  $c(\gamma_s)$  alors qu'on s'attend à ce qu'elle tende vers l'infini lorsque la singularité  $\gamma$  tend vers l'infini. On montre d'après la transformée de Legendre que dans ces conditions la fonction d'échelle des moments devient linéaire pour  $q > q_s$  :

$$K(q) = \gamma_s(q - q_s) + K(q_s) \quad \text{Équation 2.18}$$

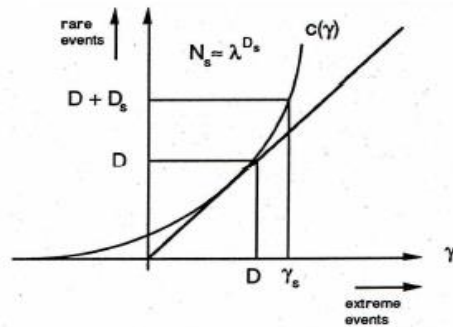


Figure 2.6 Ordre maximum de la singularité (Schertzer et Lovejoy, 1993)

Dans le cadre des multifractals universels, il a été montré que :

$$\gamma_s = \alpha' C_1 \left( \frac{D+D_s}{C_1} \right)^{\frac{1}{\alpha'}} - \frac{C_1}{\alpha-1} \quad \text{Équation 2.19}$$

Et

$$q_s = \left( \frac{D+D_s}{c_1} \right)^{\frac{1}{\alpha}} \quad \text{Équation 2.20}$$

Le choix adéquat de la dimension d'échantillon  $D_s$  est discuté dans Schertzer et Lovejoy (1993). Un exemple d'utilisation de cette notion est proposé dans Royer et al. (2008) pour étudier l'impact du changement climatique sur les précipitations extrêmes. Les auteurs ont analysé la série journalière de 12 pixels situés sur le territoire français et simulés par le modèle de climat CNRM-CM3 de Météo France sur la période 1860-2100. Ils ont constaté que l'intermittence moyenne ( $C_1$ ) a tendance à augmenter, tandis que la multifractalité ( $\alpha$ ) diminue. Étant donné que ces tendances ont des effets contradictoires sur les extrêmes, ils ont étudié la singularité maximale observable  $\gamma_s$  et ont montré que celle-ci semble augmenter pour la plupart des points de la grille, ce qui signifie que l'intermittence moyenne a des effets dominants, ce qui les a amenés à dire qu'on peut s'attendre à des précipitations plus fortes dans les 100 prochaines années. D'autres auteurs utilisent cette notion pour estimer ou prédire les précipitations maximales probables (Douglas et Barros, 2003; Hubert et al., 1993).

La taille de l'échantillon pour l'estimation des moments est donc un paramètre important. Meneveau et Sreenivasan, (1991) proposent un critère statistique. Ils ont en effet constaté qu'une taille de  $10^q$  peut-être suffisante pour l'estimation du moment d'ordre  $q$ . Cependant, ce résultat n'est pas toujours valable surtout dans le cas de données corrélées (Lombardo et al., 2014). Dans notre cas, nous analyserons des séries d'une durée minimale de 32 minutes ce qui correspond à 128 observations. L'application de la formule de Meneveau et Sreenivasan fournit une valeur  $q_s = 2.1$ .

### 2.2.2.2 Divergence des moments

De façon analogue au paragraphe précédent, la divergence des moments peut induire une mauvaise estimation de ces derniers et donc de la fonction  $K(q)$ . En effet, certaines distributions qui ont la propriété d'être à queue lourde voient leurs moments diverger au-delà d'un ordre  $q_d$ . Dans le cas des variables  $\alpha$ -stables, les moments divergent lorsque l'ordre est supérieur à la valeur  $\alpha$ . Schertzer et Lovejoy (1987) ont montré que les moments d'une cascade multifractale divergent à partir d'un ordre critique  $q_d$ , solution de l'Équation 2.21 suivante, sous réserve de son existence :

$$\begin{cases} K(q_d) = D(q_d - 1) \\ q_d > 1 \end{cases} \quad \text{Équation 2.21}$$

En effet, l'existence de ce moment  $q_d$  est directement liée aux queues de distributions lourdes, autrement dit aux extrêmes observés.

Dans le cas des précipitations, certains auteurs ( Hubert, 2005 ; Macor et al. 2007) suggèrent que la valeur  $q_d$  serait aux alentours de 3. En pratique il est difficile de mettre en évidence la divergence des moments car ces derniers sont estimés de façon empirique par la formule ci-dessous, ils sont donc toujours finis :

$$M_q(\lambda) = \frac{1}{n} \sum_{j=1}^n ((\varepsilon_j)_\lambda)^q \quad \text{Équation 2.22}$$

Lorsque les moments ne sont pas définis, c'est-à-dire pour des ordres supérieurs à  $q_d$ , les moments empiriques obtenus par la formule ci-dessus dépendent de la taille de l'échantillon. Ainsi, plus la taille de la série est grande, plus les moments seront élevés. (Lombardo et al., 2014) ont étudié la convergence des moments, en utilisant différentes distributions ayant des queues plus ou moins lourdes et en se plaçant dans un cadre multifractal. Ils ont montré que pour les moments d'ordre  $q \geq 3$ , il existait des écarts importants entre les moments théoriques et les moments empiriques. De même les incertitudes relatives aux estimations sont de plus en plus grandes et seuls les moments d'ordre  $\leq 2$  sont bien estimés.

### 2.2.2.3 Estimation des moments

Nous concluons des deux paragraphes précédents que bien qu'il serait intéressant de disposer de valeurs de moments d'ordre élevé pour l'estimation de  $K(q)$ , l'ordre maximum des moments empiriques que nous calculerons doit être suffisamment faible afin d'être inférieur non seulement à  $q_d$  mais également à  $q_s$ . Pour cette raison, nous nous limiterons dans la suite à l'estimation des moments d'ordres inférieurs à  $q_{max} = \min(q_s, q_d) = 2.1 \cong 2$ .

La Figure 2.1 représente les moments empiriques obtenus sur une série de tailles  $N=2^{14}$  pour diverses valeurs d'ordre  $q$  compris entre 0.01 et 1.5 en fonction de la résolution  $\lambda$ . Pour chaque valeur de l'ordre  $q$ , la pente de la droite fournit la valeur de  $K(q)$  correspondante. Cependant comme on peut le constater sur la Figure 2.7 de gauche, les courbes ne sont pas rectilignes, notamment aux résolutions grossières pour lesquelles le nombre de valeurs à moyenner devient faible. Afin d'améliorer l'estimation des moments (notamment aux résolutions grossières), les moments seront estimés sur un « grand » nombre de réalisations  $N_R$  (typiquement une centaine) chacune de taille  $N$ . De plus seuls ceux dont la qualité de l'ajustement sera supérieure à :  $R^2 > 0.80$  seront pris en compte. La Figure 2.7 de droite montre une très nette amélioration ( $N_R=100$ ), on observe toutefois encore une légère courbure pour les ordres élevés aux résolutions les plus fines ( $\lambda > 10$ ) c'est-à-dire celles qui incluent des composantes fréquentielles élevées.

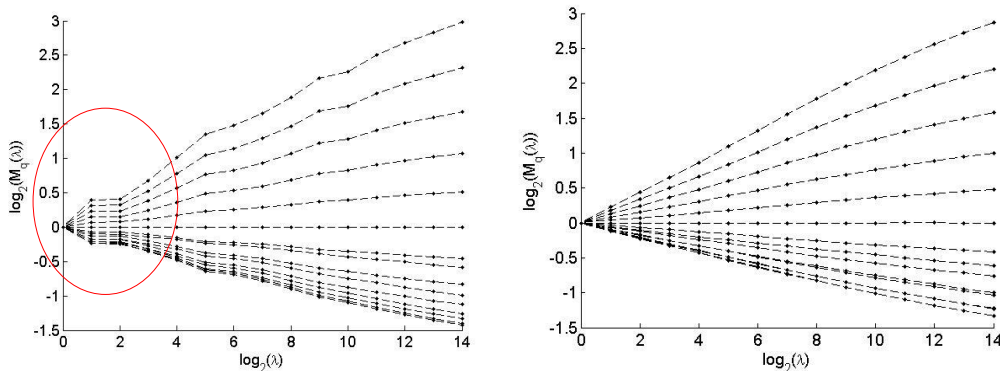


Figure 2.7 Moments estimés sur une seule réalisation (à gauche) et sur 100 réalisations (à droite)

## 2.3 Restitution des paramètres multifractals en pratique

Dans cette section, nous illustrons les résultats d'analyses multifractales sur des séries temporelles simulées avec les modèles UM et FIF en dimension 1. Dans un premier temps, nous simulons 100 séries avec une résolution  $\lambda = 14$  et avec les paramètres multifractals estimés sur des données réelles obtenues par un spectropluviomètre (cf. Chapitre 3), soit  $\alpha = 1.6$ ,  $C_1 = 0.1$  et cela dans les deux cas UM et FIF (c.-à-d.  $H = 0$  et  $H = 0.4$ ). Nous avons également réalisé d'autres simulations en utilisant des paramètres issus de la littérature (eg. Verrier et al., 2011) soit:  $\alpha = 0.6$ ,  $C_1 = 0.3$ . Au total 8 jeux de paramètres ont été utilisés. Pour chacun de ces jeux,

nous avons simulé 100 séries à l'aide du modèle multifractal original ainsi qu'à partir du modèle modifié tel qu'il a été décrit dans la section 2.2.1 avec  $N = 8$ . Pour chacun des cas nous avons estimé les paramètres multifractals à l'aide de la méthode des moments traces (cf. annexe 1). Le Tableau 2.1 présente les paramètres estimés pour les différentes configurations :

Paramètres du modèle FIF			Paramètres restitués à partir des séries simulées lorsqu'elles sont générées avec le modèle original			Paramètres restitués à partir des séries simulées lorsqu'elles sont générées avec le modèle modifié (N=8)		
H	$\alpha$	C1	H	$\alpha$	C1	H	$\alpha$	C1
0	1.6	0.1	0.11	1.61	0.10	0.06	1.64	0.11
		0.3	0.05	1.40	0.27	0.02	1.50	0.28
	0.6	0.1	0.09	0.64	0.10	0.04	0.65	0.13
		0.3	0.05	0.75	0.35	0.01	0.76	0.39
0.4	1.6	0.1	0.52	1.67	0.09	0.46	1.70	0.1
		0.3	0.50	1.62	0.27	0.43	1.65	0.27
	0.6	0.1	0.50	0.90	0.07	0.46	0.98	0.08
		0.3	0.48	1.18	0.26	0.42	0.92	0.36

**Tableau 2.1** paramètres multifractals estimés pour 8 jeux de paramètres appliqués au générateur multifractal original et au générateur multifractal modifié. Dans chaque cas de figure 100 séries à la résolution  $\lambda=14$  ont été générées.

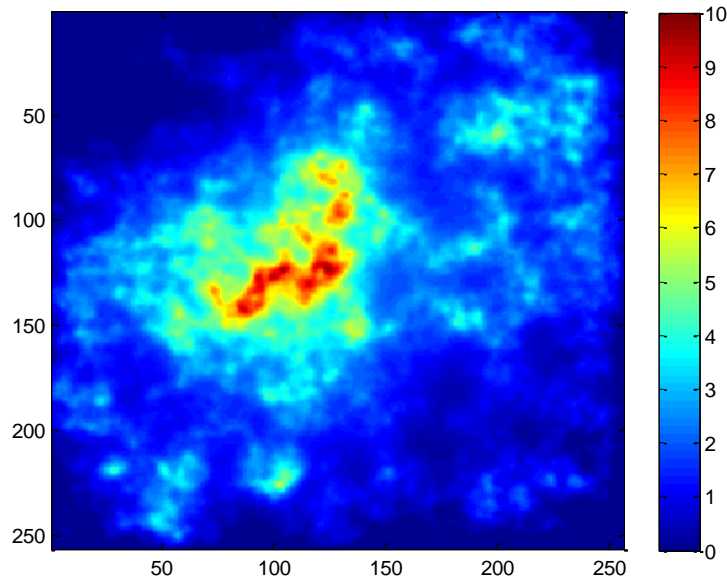
Comme nous pouvons le remarquer, le paramètre  $H$  est plus ou moins surestimé dans tous les cas de figure. Cependant, comme on pouvait l'attendre,  $H$  est mieux estimé sur les séries générées par le modèle modifié. En moyenne une valeur de  $H$  égale à 0.075 est obtenue à partir du simulateur original lorsque les données ont été générées avec  $H = 0$  et une valeur de 0.5 est obtenue lorsque les données sont générées avec une valeur de  $H = 0.4$ , soit un biais maximum de 0.1. En qui concerne le modèle modifié on obtient dans les mêmes conditions  $H = 0.03$  à la place de  $H = 0$  et 0.44 à la place de 0.4, soit un biais maximum de 0.04.

Le paramètre C1 est globalement bien restitué quel que soit le modèle utilisé même si le modèle original semble fournir un résultat légèrement meilleur avec un biais maximum de 0.05 au lieu de 0.09 pour le modèle modifié. Concernant le paramètre  $\alpha$ , les choses sont plus nuancées. On constate en effet que globalement ce paramètre est bien estimé lorsqu'il est supérieur à 1. Par contre pour  $\alpha$  inférieur à 1 celui-ci est systématiquement surestimé avec une erreur pouvant aller jusqu'à 100% dans le cas du modèle original. Il est difficile d'expliquer la raison d'une telle erreur : erreur dans la génération des données ou erreur dans l'estimation des paramètres ? Il

convient toutefois de remarquer que dans notre cas nous nous intéressons à la valeur  $\alpha = 1.6$  et que nous ne sommes par conséquent pas dans ce cas de figure.

## 2.4 Généralisation de cartes en dimension 2

Le passage en dimension 2 ne pose pas de problème particulier hormis la capacité mémoire du ordinateur qui limite compte tenu de notre ordinateur la taille des cartes à 8192x8192 pixels. La Figure 2.8 ci-dessous montre un exemple de champ généré avec les paramètres  $\alpha = 1.6$ ,  $C_1 = 0.1$  et  $H = 0.4$ .



*Figure 2.8 Exemple de carte 2D 256x256 pixels générée à l'aide du modèle FIF avec les paramètres  $\alpha=1.6$ ,  $C_1=0.1$  et  $H = 0.4$ .*

Il est à noter que les champs générés (en 1D ou 2D) sont sans unité et de valeur moyenne égale à l'unité.

## 2.5 Conclusion

Après avoir introduit les notions de fractals et multifractals, nous nous sommes focalisés sur le modèle multifractal universel (MU) et son extension le modèle multifractal universel fractionnellement intégré (FIF). Ce modèle est en effet reconnu pour bien représenter la

variabilité spatiale et/ou temporelle des précipitations et notamment la non conservativité du champs de pluie aux fines échelles. Il y a toutefois un bémol : son impossibilité de représenter la valeur zéro. Le modèle FIF ne peut donc pas à lui seul représenter correctement les précipitations. Nous avons proposé de modifier légèrement le générateur initial afin d'améliorer ses propriétés spectrales, notamment aux hautes fréquences. Il s'ensuit que non seulement le spectre en loi puissance est mieux représenté mais aussi l'estimation du paramètre  $H$  relative à l'intégration fractionnaire. Concernant l'estimation des trois paramètres multifractals nous avons vu que dans la gamme de valeurs des paramètres qui nous intéresse ( $\alpha = 1.6$ ,  $C_1 = 0.1$  et  $H = 0.4$ ) les techniques d'estimation multifractales habituelles fournissent de bons résultats moyennant quelques précautions concernant la qualité des ajustements et un choix pas trop élevé de l'ordre maximal des moments empiriques.





# Chapitre 3

## GENERATEUR DE SERIES TEMPORELLES DE PRECIPITATIONS

### Sommaire

---

3	GENERATEUR DE SERIES TEMPORELLES DE PRECIPITATIONS .....	59
3.1	Introduction.....	59
3.2	Séries temporelles de taux précipitants observées par le DBS.....	63
3.3	Périodes sèches et périodes pluvieuses .....	66
3.4	Hypothèses retenues .....	67
3.5	Analyse et modélisation du support de pluie des données observées par le DBS .....	68
3.5.1	Analyse des propriétés du support de pluie observé .....	68
3.5.2	Modélisation du support à partir d'une distribution GEV .....	72
3.5.3	Modélisation fine du support de pluie .....	76
3.6	Analyse et modélisation des séries de taux précipitants observés.....	80
3.6.1	Analyse des évènements observés et estimation des paramètres multifractals .....	80
3.6.2	Génération des séries temporelles de pluie .....	83
3.7	Générateur 1.....	83
3.7.1	Introduction.....	83
3.7.2	Calibration quantile à quantile .....	84
3.7.3	Performances.....	87
3.7.3.1	Analyse spectrale.....	88
3.7.3.2	Analyse multifractale .....	90
3.7.3.3	Autocorrélation.....	91
3.7.4	Analyse critique .....	92
3.8	Générateur 2.....	92
3.8.1	Introduction.....	92
3.8.2	Relation durée/intensité.....	94
3.8.3	Performances.....	97
3.8.3.1	Analyse spectrale.....	97
3.8.3.2	Analyse multifractale .....	98
3.8.3.3	Distribution de probabilité.....	99
3.9	Conclusion .....	101

---



## 3 GENERATEUR DE SERIES TEMPORELLES DE PRECIPITATIONS

### 3.1 Introduction

L'objectif de ce chapitre est de réaliser un simulateur capable de générer des séries chronologiques de taux précipitants. Plus précisément, il s'agit de développer un simulateur présentant des caractéristiques statistiques les plus proches possibles de celles observées, et ce, dans une large gamme d'échelle en partant de l'échelle la plus fine possible (en adéquation avec les caractéristiques de nos instruments de mesure). En effet, une étude du spectre de taux précipitants réalisée par de (De Montera et al. 2009) a clairement mis en évidence que les propriétés d'invariance d'échelle sont vérifiées jusqu'à des échelles de quelques dizaines de secondes (typiquement 15 s) voire moins<sup>2</sup>. **Nous nous fixons donc comme objectif de générer des séries chronologiques exhibant des propriétés statistiques identiques à celles observées pour des échelles allant de 15 secondes jusqu'à une échelle la plus large possible compte tenu des observations dont nous disposons.**

Afin de modéliser la pluie, il est nécessaire, dans un premier temps, de décider si les précipitations sont considérées comme un processus continu ou discret ; dans ce contexte, plusieurs études existent dans la littérature, par exemple (De Michele et Ignaccolo, 2013). Une fois ce choix fait, différentes approches peuvent être utilisées. Ainsi, (Cooley et al., 2007) proposent un modèle hiérarchique Bayésien caractérisant les événements de précipitations extrêmes produits par un modèle climatique régional. (Wang et al., 2012) choisissent de modéliser les périodes sèches et humides en utilisant des modèles à base de chaînes de Markov. Un grand nombre de générateurs est basé sur une hypothèse de processus de Poisson, on peut citer parmi d'autres (Onof et al., 2000; Burton et al., 2008; Evin et Favre, 2013; Bernardara et al., 2007). Les techniques géostatistiques comme le krigeage sont également largement utilisées pour modéliser la pluie (Schleiss et al., 2009, 2014; Leblois et Creutin, 2013). D'autres types de

---

<sup>2</sup> Une surface de capture des gouttes de pluie plus importante serait nécessaire pour étudier des durées plus faibles.

modèles existent comme le modèle de Hurst-Kolmogorov (Tyrallis et Koutsoyiannis, 2011). Parmi tous ces nombreux modèles, les modèles multifractals décrits au chapitre précédent ont reçu beaucoup d'attention dans la littérature géophysique (Deidda et al., 1999 ; Deidda, 2000 ; Gires et al., 2013 ; Over et Gupta, 1996 ; Rupp et al., 2009 ; Schertzer et Lovejoy, 1987 ; Schmitt et al., 1998 ; Serinaldi, 2010 ; Tessier et al., 1996 ; Veneziano, et al., 1996 ; Veneziano et Iacobellis, 2002 ; Veneziano et Lepore, 2012). Basés sur les propriétés d'invariance d'échelle ces modèles utilisent le concept de cascades multiplicatives discrètes ou continues (e.g., Schertzer et Lovejoy, 1987; Lovejoy et Schertzer, 1995; Lovejoy et al., 2008; Lovejoy et Schertzer, 2010; Serinaldi, 2010; Gaume et al., 2007). La différence entre ces différents modèles réside principalement dans la manière de prendre en compte (ou non) certaines caractéristiques des précipitations:

- **L'intermittence et les régimes de scaling:** Comme nous l'avons déjà précisé, le générateur proposé par (Schertzer et Lovejoy, 1987) et décrit au chapitre précédent ne permet pas de générer de valeurs nulles. De plus, il ne permet de générer qu'un seul régime de scaling, c'est-à-dire des paramètres multifractals identiques pour toutes les échelles considérées. L'analyse spectrale des séries de taux précipitants met en évidence un comportement en loi de puissance pour certaines gammes d'échelle. L'existence de plusieurs régimes de scaling a été montrés par (Fraedrich et Iardner, 1993 ; Olsson 1995 ; Fabry, 1996 ; Molini et al. 2009; Rodríguez et al., 2013). Ces régimes de scaling classiquement observés sur les spectres des séries chronologiques de précipitations traduisent les propriétés d'invariance d'échelle pour certaines gammes d'échelles. L'analyse multifractale de séries chronologiques observées à fine échelle montre une cassure à l'échelle de la demi-heure. Les travaux de (De Monterra et al., 2009) ont mis en évidence le lien entre l'intermittence et les cassures observées entre les différents régimes de scaling. Verrier et al., (2011) ont mis en évidence un premier régime de scaling à l'échelle sub-horaire compris entre 30 secondes et 30 minutes environ. Ce régime est interprété par les auteurs comme un régime dit intra événement, qui relie les propriétés de relation d'échelle de la pluie au sein d'un même événement pluvieux. Aux petites échelles, la pente du spectre ( $\beta$ ) estimée par les auteurs se situe aux alentours de 1,5 (voir figure 3.18) signifiant ainsi que le processus est clairement non conservatif. (Comme vu à la fin de la section 2.2.1.2 avec l'équation 2.13 b, un  $\beta > 1$  implique que le paramètre de

non conservativité  $H$  est différent de 0.) Au-delà de trois jours la pente du spectre s'aplatit très nettement indiquant par-là une décorrélation importante au-delà de cette période. Un régime de transition est visible entre 30 minutes et trois jours avec une pente s'aplatissant vers les larges échelles. Ce régime de transition est scindé en 2 par ces auteurs. Au final ceux-ci définissent trois cassures dans le spectre définissant ainsi 3 périodes de 'scaling' : 30 secondes – 30 minutes, 30 minutes – 3 heures, 3 heures – 3 jours. Ces travaux ont souligné la nécessité de distinguer les propriétés de l'intermittence de la variabilité des taux précipitants non nuls. Il est important de noter qu'au-delà de 30 minutes les cassures sont liées aux caractéristiques du support de la pluie. Ainsi, comme décrit par (De Montera et al., 2009 ; Veneziano et Lepore, 2012 ; Gires et al., 2012), les paramètres multifractals estimés sur la série temporelle complète dépendent fortement des caractéristiques du support. Différents auteurs (Verrier et al., 2010; De Montera et al., 2009, 2010) ont étudié les paramètres multifractals intra-événement pour 3 régions climatiques différentes à partir de données des disdromètres et des radars. Ils montrent que les paramètres multifractals restent relativement constants pour les 3 zones considérées et que ces paramètres sont identiques en temps et en espace. Verrier et al., (2011) ont quantifié l'effet de l'inclusion des zéros sur l'estimation des paramètres multifractals. Plus récemment une autre étude a confirmé ce point (Gires et al., 2013). De ces travaux antérieurs relatifs à l'analyse multifractale de taux précipitants à haute résolution on retiendra pour notre étude, que les paramètres MU du régime sub-horaire compris entre 30 secondes et 30 minutes environ sont universels (i.e. associées aux équations partielles régissant la physique régissant les petites échelles de la pluie et de la turbulence) et utilisables quel que soit le lieu géographique considéré, qu'au-delà de 30 minutes les cassures et les propriétés d'invariance d'échelle sont liées à l'intermittence des précipitations (et donc aux conditions spécifiques à chaque région, par exemple, les forçages et les conditions géographiques). Il faut toutefois rappeler que les paramètres multifractals permettent de modéliser la variabilité des précipitations et non leur intensité, qui elle, peut être très différente d'une région à une autre.

- **Les échelles de représentation** : la plupart des modèles existants ont été développés pour des résolutions temporelles typiquement de 5 mn, une heure voire 24 heures. L'utilisation de ces modèles pour générer des données à des résolutions plus fines nécessiterait une

étape supplémentaire de type downscaling pour passer à des résolutions plus fines (e.g. Gaume et al., 2007 ; Rupp et al., 2009). Notre objectif ici n'est pas de développer une méthode de downscaling mais plutôt de développer notre propre modèle directement à l'échelle fine (15 s).

- **Distribution des taux de pluie** : Certains modèles proposés (Hubert, 2001 ; Serinaldi, 2010) ont tendance à restituer les propriétés d'échelle des précipitations. Ils ne s'intéressent qu'assez peu à la distribution des taux de pluie restitués à la fine échelle. Enfin, on peut noter l'étude de (Pathirana et al., 2003), s'intéressent plus spécifiquement aux taux de pluie, qui ont proposé un modèle qui permet d'avoir de bons résultats en terme de périodes de retour des durées et des volumes d'eau mais pas en terme d'autocorrélation des taux précipitants au-delà de quelques heures.

L'objectif principal de ce chapitre étant de développer un générateur de précipitations 1D réaliste, c'est-à-dire capable de reproduire les séries observées à différentes échelles, il convient de définir précisément les différentes caractéristiques que devra posséder ce dernier :

- **Les propriétés d'invariance d'échelle** : les différents régimes d'invariance d'échelle et les coefficients multifractals qui les caractérisent, le spectre et la fonction d'autocorrélation de la série.
- **L'intermittence de la pluie** : la probabilité d'occurrence de la pluie, la dimension fractale du support de pluie, les périodes de retour des durées de pluie, les fonctions de survie empiriques (ESF) des durées de pluie à plusieurs échelles de temps.
- **La distribution des taux précipitants**: les fonctions de survie empiriques (ESF) des taux précipitant et des hauteurs d'eau à plusieurs échelles de temps.

Nous nous appliquerons donc à valider le simulateur à partir de sa capacité à restituer convenablement l'ensemble des caractéristiques ci-dessus. Certaines caractéristiques telles que la fonction de survie empirique (ESF) des taux de pluie doivent être validées à plusieurs échelles de temps (typiquement 5 minutes, 1 heure, 1 jour). Cela nous semble en effet absolument indispensable : une série temporelle simulée peut en effet présenter une « bonne » distribution statistique à une échelle de temps donnée mais avoir une distribution totalement erronée à une autre échelle si l'intermittence de la pluie n'est pas correctement modélisée.

## 3.2 Séries temporelles de taux précipitants observées par le DBS

Cette étude repose sur l'utilisation de séries temporelles de taux de pluie (ci-après RR) obtenues à partir d'un disdromètre développé au LATMOS (Dual-Beam Spectropluviometer, DBS) (Figure 3.1 Schéma de principe du Dual Beam Spectropluviometer) décrit en détail dans (Delahaye et al., 2006). Cet instrument permet d'enregistrer le temps d'arrivée des gouttes de pluie ainsi que leur diamètre  $D$  et leur vitesse de chute verticale  $V$ . Chaque goutte est horodatée avec une précision de 1 milliseconde, assurant ainsi une bonne résolution temporelle pour la récupération du support de pluie. La surface de captation correspond à un rectangle de surface  $S=100 \text{ cm}^2$ . De la même manière qu'un pluviomètre, l'observation des précipitations est très localisée. En intégrant sur un laps de temps  $T= 15$  secondes, les gouttes recueillies permettent d'estimer le taux de pluie correspondant :

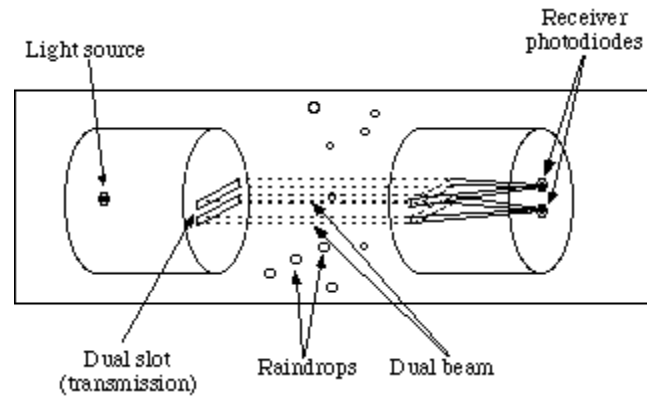
$$RR = \frac{600\pi}{TS} \sum_{i=1}^N D_i^3$$

Nous avons utilisé deux jeux de données couvrant chacun une période de deux ans. Le premier a été observé entre Juillet 2008 et Juillet 2010 et a été utilisé pour le choix des modèles et l'apprentissage tandis que le second a été observé entre octobre 2010 et juin 2013 et a été utilisé pour la validation du modèle. Durant toute cette période, le DBS était basé sur le Site d'Instrumentation et de Recherche sur la Télédétection Atmosphérique (SIRTA<sup>3</sup>) situé sur le campus de l'Ecole Polytechnique à Palaiseau en région parisienne. Il est à noter que ces données incluent une correction des épisodes neigeux qui sont aisément distinguables par le spectropluviomètre puisque les flocons de neige ont un diamètre large et une vitesse plus lente que les gouttes de pluie.

---

<sup>3</sup> <http://sirta-dev.ipsl.jussieu.fr/joomla/index.php/85-article-sans-categorie/71-sirta-home-page>





*Figure 3.1 Schéma de principe du Dual Beam Spectropluviometer*

Les propriétés des séries temporelles de précipitation dépendent, d'un côté, de la résolution temporelle du capteur mais aussi de son seuil de détection. Un seuil de détection élevé peut négliger les gouttes de faible diamètre, et donc surestimer les durées des périodes sèches. En revanche, un seuil trop bas peut augmenter les risques de fausses détections (poussière, insectes, turbulence de l'air en périodes très ensoleillées), conduisant à une sous-estimation des durées des périodes sèches. Après différents essais nous avons sélectionné un seuil de détection sur le taux précipitant de 0.05 mm/h. Cela nous a paru en effet être un bon compromis entre sensibilité aux petites gouttes et un seuil de fausses détection suffisamment bas pour l'étude du support de pluie. La série temporelle des taux de pluie utilisée pour l'apprentissage est composée de  $5.2 \times 10^6$  observations enregistrées avec une résolution de 15 secondes (taux de pluie intégré sur 15 secondes). La valeur maximale enregistrée à cette résolution est 185mm/h. Dans cette série 4.6% des observations sont non nuls (pluie) ce qui correspond à environ 1000 heures de pluie distribuées sur 12195 périodes pluvieuses (ayant des durées comprises entre 15 secondes et 12 heures). Les caractéristiques statistiques des séries de pluie utilisées sont résumées dans le Tableau 3.1 Certains événements de pluie extraits de ce même jeu de données (mais avec un prétraitement légèrement différent) ont été utilisés dans des travaux antérieurs (Verrier et al., 2011) pour analyser les propriétés d'invariance d'échelle à fine résolution temporelle.

	Série apprentissage		Série validation	
Durée totale de la série (mois)	24		20	
Temps d'intégration	15 s	1h (dégradée)	15 s	1h (dégradée)
Durée de la période de pluie (heure)	1000	2370	1230	3511
Occurrence de pluie (%)	4,6	1.29	5.31	1.49
Nombre de périodes pluvieuses = Nombre de périodes sèches (+ ou - 1)	12 195	759	15775	1135
Nombre d'événements de durées supérieures à 8 min	1049		1614	
Durée maximale des périodes pluvieuses (heure)	12	21	12	24
Durée maximale des périodes sèches (heure)	589	689	380	380
Valeur minimale de RR (mm/h)	0,05(seuil fixe)	0.002	0,05 (seuil fixe)	0.003
Valeur maximale de RR (mm/h)	185	22.38	148	16.47
Valeur moyenne de RR sur la série (mm/h)	0.05	0.05	0.06	0.06
Valeur moyenne de RR sur la période de pluie (mm/h)	1.09	0.37	1.12	0.39
Ecart type de RR sur la série (mm/h)	0.7	0.38	0.7	0.35
Ecart type de RR sur la période de pluie (mm/h)	3.15	0.99	2.85	0.83
Quartiles Q1/Q2/Q3 sur la période de pluie (mm/h)	0.14/0.40/1.06	0.007/0.06/0.36	0.22/0.5/1.16	0.01/0.09/0.40

**Tableau 3.1** Principales caractéristiques des séries chronologiques de taux précipitants à la résolution initiale de 15s utilisées pour mettre au point et valider le simulateur. Caractéristiques obtenues pour les séries agrégées avec un pas de temps horaire.

L'analyse des caractéristiques des données du DBS doit nous permettre de développer un simulateur capable de reproduire les propriétés de cette série à fine échelle mais aussi à des échelles plus larges (horaire, journalière) par simple agrégation. Le simulateur ainsi réalisé doit nous permettre de simuler des séries synthétiques de taux précipitants « observables » à

Palaiseau, c'est-à-dire comme étant des réalisations possibles. Bien entendu, la simulation de séries synthétiques « observables » dans d'autres régions climatiques nécessitera l'ajustement de certains paramètres du modèle, ce point sera discuté en fin de chapitre.

### 3.3 Périodes sèches et périodes pluvieuses

Selon la définition météorologique classique, un événement de pluie peut contenir de courtes périodes sèches (des zéros). Cette notion est inévitablement conditionnée par les caractéristiques du capteur utilisé. En effet, les pluviomètres classiques utilisent des augets basculants délivrant une impulsion lorsqu'une certaine hauteur d'eau a été collectée (0.1 ou 0.2 mm typiquement). En présence de pluie faible l'auget peut ne pas se remplir complètement retardant ainsi la fourniture de l'impulsion au prochain événement de pluie. De même, si le temps de remplissage de l'auget est supérieur au temps d'intégration, une valeur nulle sera estimée alors qu'en réalité il pleut. Les pluviomètres ne peuvent par conséquent pas rendre compte, de façon fidèle, des transitions pluie / non pluie. Dans le cas du DBS, celui-ci mesure avec précision le temps d'arrivée des gouttes, ce qui permet de distinguer clairement pour chaque tranche de temps de 15 secondes s'il pleut ou pas.

Dans l'approche que nous proposons, **nous n'avons pas utilisé la définition météorologique des événements de pluie**, mais plutôt **la notion de «continuité» des périodes pluvieuses et sèches**. Dans ce contexte, **une période sèche (drought duration sur certaines figures) est définie comme une série des taux de pluie (intégrés sur 15 secondes) consécutifs de valeur nulle**. Les durées sèches sont noté  $d_{nr}$ . De même, **une période pluvieuse (rain duration sur certaines figures) est une série de taux de pluie consécutifs et non nuls à la résolution de 15s**. Les durées pluvieuses sont noté  $d_r$ . Il est important de noter que, au sens météorologique, quel que soit la nature de l'évènement stratiforme ou convective, un événement de pluie est composé de plusieurs périodes pluvieuses qui sont susceptibles d'être régis par la même dynamique.

### 3.4 Hypothèses retenues

Notre simulateur est conçu sur la base de 3 hypothèses principales. La première est liée au support de pluie, les deux autres sont liées des propriétés internes aux périodes pluvieuses. La validité de ces hypothèses par rapport à nos données est discutée dans les sections suivantes.

**Hypothèse 1 :** En se basant sur les conclusions de (Lavergnat et Golé, 2006) et (Schmitt et al. 1998), nous considérons que le support de pluie peut être modélisé comme une alternance de périodes pluvieuses  $d_r$  et des périodes sèches  $d_{nr}$ . On suppose également que les durées des périodes pluvieuses et sèches sont mutuellement indépendantes. Nous verrons plus bas que cette hypothèse forte est raisonnable sur notre jeu de données à la résolution de 15 secondes.

**Hypothèse 2 :** Le taux précipitant RR à l'intérieur d'une période pluvieuse a des propriétés d'invariance d'échelle et peut être représenté par une modélisation multifractale.

Le modèle multifractal universel (MU) intégré « Fractional Integrated Flux » (FIF) est adapté pour modéliser la variabilité du taux précipitant de chaque période pluvieuse (e.g. Schertzer et Lovejoy 1987, Montera et al. 2009 ; Verrier et al. 2010 ; 2011). Conformément aux résultats des études antérieures présentées en introduction de ce chapitre, nous utilisons un seul jeu de paramètres multifractals que nous considérerons comme « universels ». Ce jeu de paramètres caractérise les relations d'échelle à l'intérieur des périodes pluvieuses qui est donc supposé indépendant de la période pluvieuse considérée. Pour chaque période pluvieuse, le modèle FIF à 3 paramètres permet de générer une série de taux de pluie normalisés notés  $RR_N$ .

**Hypothèse 3 :** Une étape de dé-normalisation est nécessaire afin de simuler des séries de taux précipitants RR réalistes à partir des taux de pluie normalisés  $RR_N$ . Le taux de pluie moyen  $\langle RR \rangle$  varie d'une période de pluie à l'autre. Deux approches sont envisagées pour cette étape de dé-normalisation. Une approche globale appliquée sur l'ensemble de la série normalisée et une approche de dé-normalisation appliquée sur chaque période de pluie en fonction de sa durée. Ces deux approches ont fait l'objet de deux études et sont implémentées dans deux versions de générateurs:

- Générateur 1 (section 3.7) : les séries chronologiques simulées de longueur identique à celle observée ont la même distribution de probabilité cumulée. Pour cela on déterminera une « fonction de calibration » du type « lookup table » permettant le passage des valeurs

normalisées aux valeurs dé-normalisées. Dans ce cas la distribution de probabilité des taux précipitants de la série chronologique simulée est par construction identique à celle observée.

- Générateur 2 (section 3.8): pour chaque période pluvieuse, on déterminera le taux précipitant moyen de la période pluvieuse considérée  $\langle RR \rangle_{d_r}$  à partir de sa durée  $d_r$  (résultant de la simulation du support). Le taux précipitant dé-normalisé d'une période pluvieuse  $RR$  sera obtenu par multiplication du taux précipitant normalisé  $RR_N$  par  $\langle RR \rangle_{d_r}$ . Dans ce cas la distribution de probabilité des taux précipitants de la série chronologique simulée est liée à la simulation du support et à la simulation des taux précipitants moyens à partir des durées  $d_r$  des évènements.

### 3.5 Analyse et modélisation du support de pluie des données observées par le DBS

Dans un premier temps le support de la série DBS est analysé sous deux aspects :

- ✓ On étudie le caractère fractal (au sens statistique) du support de la pluie qui assure la propriété dite d'invariance d'échelle dans une certaine gamme d'échelle (que nous préciserons plus loin). Ce caractère fractal du support devra être préservé sur les séries simulées, autrement dit les probabilités des durées des périodes pluvieuses devront suivre une loi de puissance relative à l'échelle temporelle considérée.
- ✓ Conformément à l'hypothèse 1, on cherche à modéliser de façon indépendante les distributions des durées des périodes de pluie et de sécheresse.

#### 3.5.1 Analyse des propriétés du support de pluie observé

Comme nous l'avons mentionné précédemment l'adéquation entre les propriétés des séries synthétiques et celles des séries observées à différentes échelles repose sur la répartition dans le temps des taux précipitants simulés. Il s'agit donc en premier lieu de simuler de manière réaliste l'intermittence des périodes de pluie/non pluie. Pour cela, on définit le support de pluie  $I_\lambda(n)$  :

$$I_\lambda(n) = \begin{cases} 0 & \text{si } RR_\lambda(n) < 0.05 \text{ mm/h} \\ 1 & \text{si } RR_\lambda(n) \geq 0.05 \text{ mm/h} \end{cases} \quad \text{Équation 3.1}$$

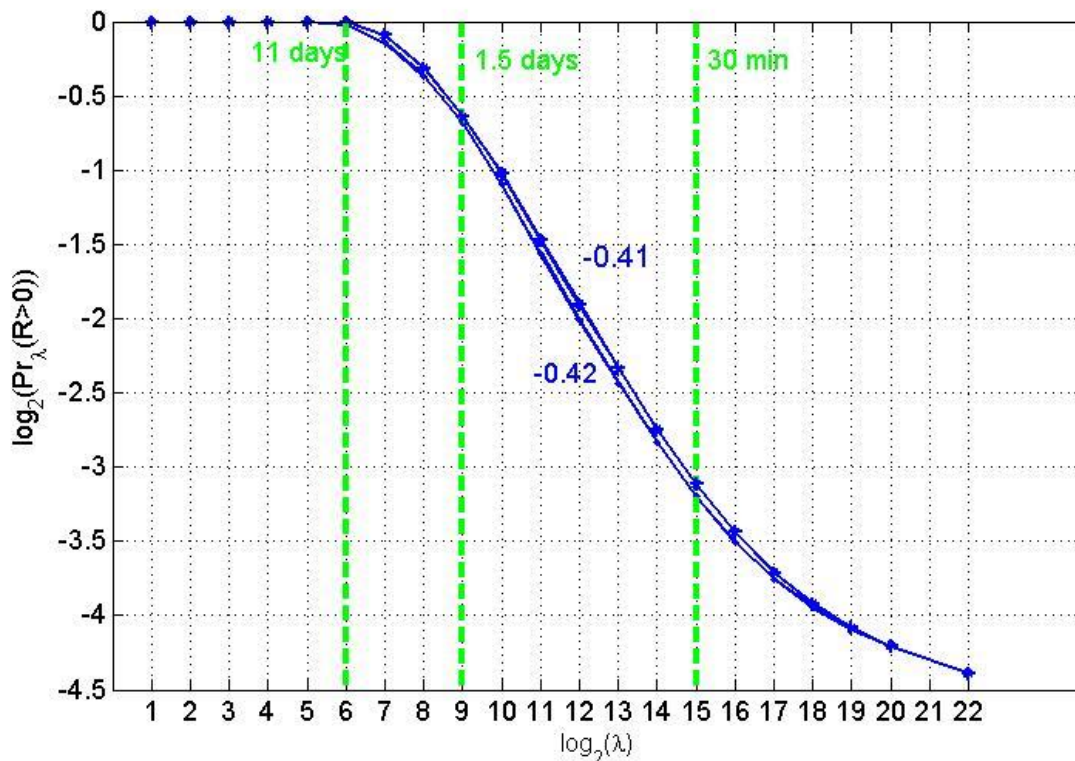
L'indice  $n$  correspondant à la position de l'échantillon et  $RR_\lambda(n)$  au  $n^{\text{ième}}$  taux précipitant de la série à la résolution  $\lambda$ .

Le support possède certaines propriétés statistiques notamment celle de pouvoir être vu comme un objet fractal dans une certaine gamme d'échelle (Olsson et al. 1993 ; Hubert et Carbonnel, 1993 ; Schmitt et al. 1998 ; Lavergnat et Golé 1998; Ghanmi et al. 2013). Si tel est le cas, la probabilité qu'une valeur  $RR_\lambda(n)$  soit non nulle ou de façon équivalente que  $I_\lambda(n)$  soit égale à 1 suit une loi de la forme :

$$Pr(RR_\lambda > 0) = \frac{Card(I_\lambda=1)}{Card(I_\lambda)} \approx \lambda^{-C_f} \quad \text{Équation 3.2}$$

La méthode dite de « comptage de boîte » (box-counting) qui consiste à estimer la pente de la droite lorsque cette équation est considérée dans un repère bi-logarithmique permet d'obtenir une estimation de la co-dimension  $C_f$ . La Figure 3.2 fournit une illustration de cette équation obtenue à partir du support de notre série de données de 2 ans. Nous remarquons que sur la gamme d'échelle 30 minutes à 1,5 jours il existe une relation linéaire de pente  $C_f=0.42$ . Pour les échelles supérieures à 12 jours la probabilité qu'il pleuve tend vers 1 ( $C_f=0$ ), il n'existe pas de période de sécheresse supérieure à 28 jours dans notre jeu de données. Pour les petites échelles inférieures à 30 minutes, la probabilité qu'il pleuve tend vers  $2^{-4.4}$  soit 4.6%, valeur attendue sous nos latitudes. Olsson et al. (1993) ont obtenu par une méthode similaire une pente  $C_f= 0.63$  entre 45 mn et une semaine à partir de données collectées en Suède. Schmitt et al. (1998) ont estimé une valeur de  $C_f= 0.45$  entre 10mn et 3 jours pour une série observée en Belgique. Veneziano et Lacobellis (2002) une pente  $C_f= 0.5$  entre 20 mn et 3 jours. Ghanmi et al. (2013) obtiennent une pente de  $C_f= 0.56$  entre 5 mn et 2 jours pour des séries observées en milieu semi-aride. Hubert et Carbonnel, (1993) ont obtenu des valeurs de la pente proche de 0.2 entre 1h et 32 h mais en utilisant des séries horaires relativement courtes. Les différences obtenues par ces auteurs à partir de jeux de données différents concernent aussi bien les gammes d'échelles que la valeur de la pente. Ces différences sont dues à la physique des événements, elles peuvent aussi s'expliquer en grande partie par les propriétés du capteur utilisé, le support observé étant très lié à son seuil de détection et à sa résolution temporelle. Il est à ce jour impossible de déterminer précisément quelle est la part respective entre le système d'observation et la situation météorologique

considérée dans la variabilité de la dimension fractale du support estimé. Dans le cadre de cette étude on s'attachera donc à reproduire les propriétés du support telles qu'elles sont observées par le DBS puisque les simulations sont réalisées avec un seuil de mesure et une résolution identique.



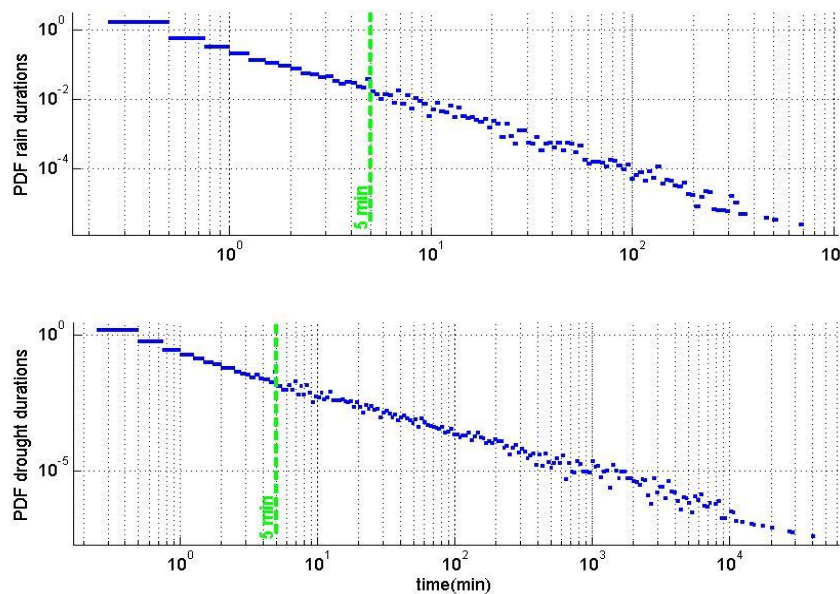
**Figure 3.2** Fréquence des durées des périodes pluvieuse en fonction de la résolution des séries d'apprentissage et de validation accompagnée de leur co-dimension.

Nous avons fait comme hypothèse l'indépendance des durées des périodes sèches  $d_{nr}$  avec les périodes pluvieuses  $d_r$  (hypothèse 1). Il s'agit de vérifier cette hypothèse conformément à plusieurs études, notamment celle de (Lavergnat et Golé, 1998 ; 2006) qui ont montré que les durées entre deux arrivées de gouttes de pluie sont décorrélées et que le coefficient de corrélation entre deux arrivées consécutives de gouttes est de 0.08. Cette propriété a été confirmée sur notre jeu de données (qui concerne non pas des gouttes mais des taux précipitants sur 15s) avec un coefficient de corrélation très faible (0.003). Par ailleurs, la distribution des couples  $(d_{nr}, d_r)$  ne fait apparaître aucune relation entre les deux variables. Nous en concluons, que bien que cette validation soit partielle. Il est possible de considérer les durées des périodes sèches et des périodes pluvieuses comme deux variables aléatoires indépendantes.

Ce résultat permet de simuler de façon simple le support de pluie comme étant une succession alternée de tirages de durées de périodes sèches  $d_{nr}$  et pluvieuses  $d_r$ .

La Figure 3.3 présente les densités de probabilités empiriques des périodes sèches  $d_{nr}$  et des périodes pluvieuses  $d_r$  observées sur la série d'apprentissage. Il est cependant impossible de conclure quant aux valeurs réelles de troncature des durées maximales. Nous ne disposons en effet que de réalisations sur deux ans. Il est par conséquent impossible de dire si les distributions sont réellement tronquées pour des raisons dues à la physique du milieu ou en raison de la taille limitée de nos séries de données. Comme nous pouvons l'observer, la Figure 3.3 montre un comportement en loi puissance à la fois sur les durées des périodes de pluie  $d_r$  et celles de sécheresse  $d_{nr}$ .

Une observation plus détaillée montre cependant que deux comportements peuvent être distingués : avant et après 5 minutes. Cette remarque s'applique aux périodes pluvieuses mais aussi aux périodes sèches : nous y reviendrons plus tard (cf. section 3.5.3) .



**Figure 3.3. Densités de probabilité empiriques des durées des périodes de pluie (rain duration) et des périodes de sécheresse (drought duration) de la série d'apprentissage (cf. section 3.3 pour les définitions de ces deux types de périodes.)**



### 3.5.2 Modélisation du support à partir d'une distribution GEV

Dans un premier temps, nous avons choisi une modélisation simplifiée malgré l'observation de deux comportements sensiblement différents avant et après 5 minutes. Les meilleurs ajustements des deux distributions empiriques ont été obtenus à l'aide d'une loi GEV non tronquée, ce qui nous a amené à choisir cette loi pour modéliser les durées des périodes sèches et pluvieuses. Un comportement global en loi de puissance est donc considéré.

La fonction de distribution de probabilité de la GEV a la forme générale suivante :

$$f(d | k, \mu, \sigma) = \left(\frac{1}{\sigma}\right) \exp\left(-\left(1 + k \frac{(d-\mu)}{\sigma}\right)^{-\frac{1}{k}}\right) \left(1 + k \frac{(d-\mu)}{\sigma}\right)^{-1-\frac{1}{k}} \quad \text{Équation 3.3}$$

La variable  $d$  représente la durée des périodes (sèche/pluvieuse). Cette loi requiert trois paramètres,  $k$  est le paramètre de forme, il contrôle la 'lourdeur' de la queue de distribution,  $\mu$  est le paramètre de localisation et  $\sigma$  le paramètre d'échelle.

- $k = 0$  correspond à une décroissance de type exponentiel (loi de Gumbel), les queues légères et non bornées.
- $k > 0$  correspond aux distributions à queue lourdes et bornées à gauche (Fréchet)
- $k < 0$  Correspond aux distributions bornées à droite (Weibull)

Les paramètres de la GEV ( $k$ ,  $\sigma$  et  $\mu$ ) ont été estimés sur les durées  $d_r$  et  $d_{nr}$  observées à l'aide d'un estimateur basé sur la méthode de maximum de vraisemblance. Nous avons respectivement obtenu les valeurs suivantes :

- $k=4.38$ ,  $\sigma=0.34$ ,  $\mu=0.08$  pour les durées des périodes sèches  $d_{nr}$
- $k=3.42$ ,  $\sigma=0.02$ ,  $\mu=0.05$  pour les durées des périodes pluvieuses  $d_r$ .

Nous avons estimé  $k > 0$  pour les deux séries de durées, ce qui correspond à la loi de Fréchet. Les paramètres  $k$  et  $\sigma$  sont plus grand pour les durées  $d_{nr}$  comparées à ceux estimés sur les durées  $d_r$ . Cela s'explique simplement par le fait que les durées des périodes sèches sont globalement plus longues que les durées des périodes pluvieuses.

La Figure 3.4 montre la pdf empirique des durées observées sur la série d'apprentissage et une pdf simulée obtenue en simulant une série de deux ans avec les paramètres estimés ci-dessus. Pour les périodes sèches, nous avons une bonne correspondance globale entre pdf observée et simulée. Pour les durées longues on note cependant un sur-échantillonnage des durées simulées au-delà de 12 heures. Pour les pdf des durées des périodes pluvieuses, nous observons un sur-échantillonnage des longues durées simulées (supérieures à 3 heures) et un sous-échantillonnage des courtes durées simulées (entre 30 secondes à 30 minutes).

Les paramètres des lois GEV obtenus sont utilisés pour générer 100 réalisations d'une durée de deux ans et demi à la résolution de 15 secondes. La Figure 3.5 reprend les résultats de l'analyse fractale du support (cf. Figure 3.2 probabilité qu'il pleuve en fonction de la résolution  $\lambda$ ) obtenue sur les séries observées et sur les 100 séries simulées. Bien que l'on puisse noter, sur la Figure 3.5, un décalage vertical entre les 100 réalisations et les 2 réalisations observées, il apparaît néanmoins un même comportement fractal durant des périodes comprises entre 30 minutes et 1,5 jours environ. Dans cette gamme d'échelle, Les co-dimensions fractales estimées sur les supports synthétiques sont proches de celle calculées sur les supports observés (Tableau 3.4). Le décalage signalé précédemment s'explique facilement : à la résolution  $\lambda = 6$  correspondant à 11 jours, nous remarquons que, pour les deux supports observés, les courbes atteignent la valeur zéro (probabilité de pluie à 1), tandis que pour les supports synthétiques, elles ne l'ont pas encore atteint. Comme nous l'avons signalé au paragraphe précédent, une explication plausible est que notre modèle a tendance à surestimer les durées des périodes sèches  $d_{nr}$  (voir Figure 3.4) : autrement dit le modèle peut générer des périodes sèches (ici de l'ordre de 30 jours pour les séries simulées) plus longues que ce qui a été observé (24 jours maximum pour les observations ayant servi à apprendre le modèle). Cela reste, tout de même, proche de la valeur observée sur la série de validation (28 jours). A la vue de la longueur de la série modélisée (de l'ordre de deux ans), il convient cependant de remarquer qu'il n'est pas forcément judicieux de vouloir faire absolument coller le modèle aux observations sur ce point. Rappelant que l'estimation des propriétés fractales du support est très sensible à la qualité et aux caractéristiques des observations utilisées.

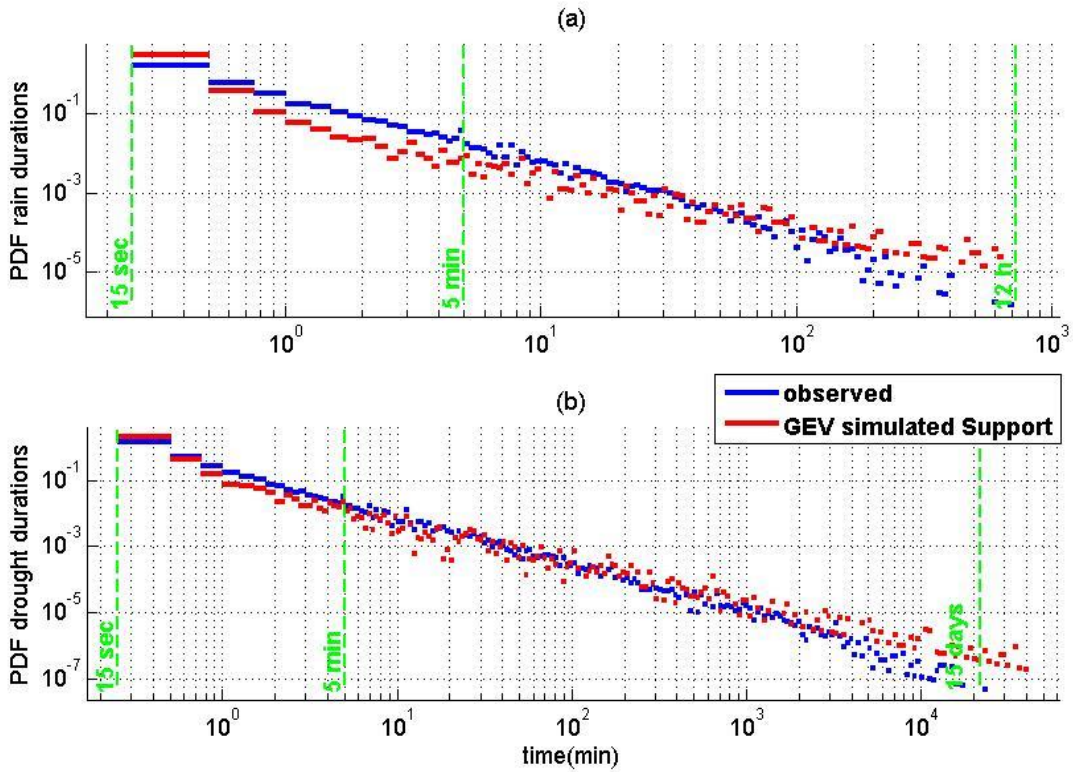


Figure 3.4 Densité de probabilité des périodes de pluie et des périodes de sécheresse de la série d'apprentissage et d'une série simulée

L'utilisation de deux distributions GEV permet de reproduire relativement bien les pdfs des durées des périodes sèches et humides. La co-dimension estimée sur 100 réalisations est très proche de celle observée. La probabilité d'occurrence de pluie à la résolution de 15 secondes (en moyenne 4.29%) reste très proche de celle estimée sur les deux supports observés (4.6% et 5.3%) (voir Tableau 3.4). En revanche la restitution de la durée des événements en termes de périodes de retour ne correspond pas à celle observée.

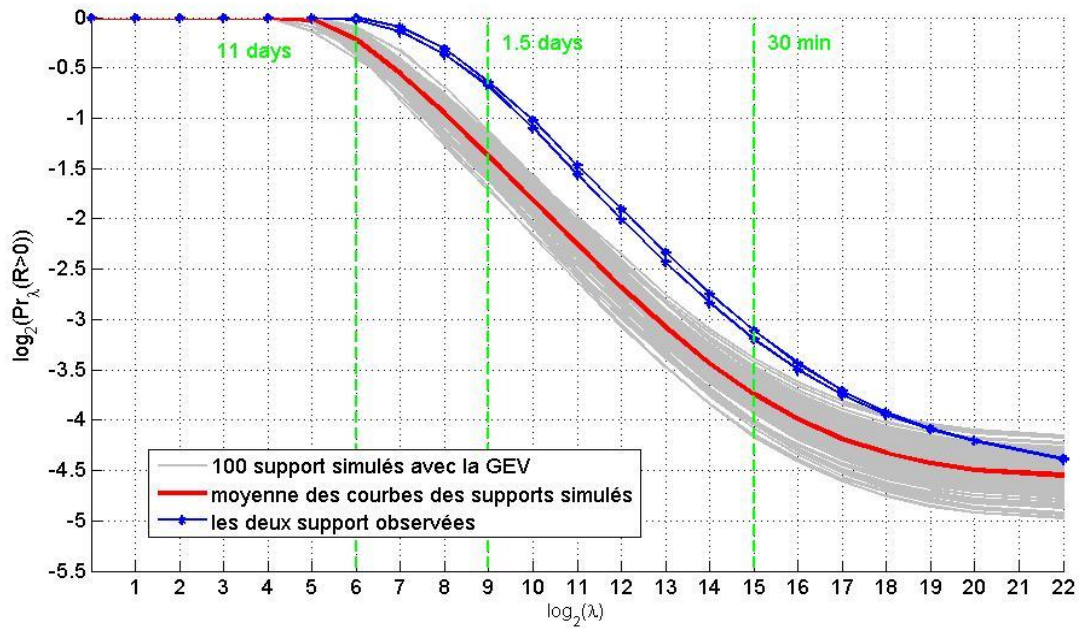


Figure 3.5. Fréquence de la pluie du support de la série apprentissage et du support simulé en fonction de la résolution

La Figure 3.6 montre la durée des périodes pluvieuses en fonction du temps de retour.

Cette figure permet d'apprécier l'incidence de la surestimation des durées pluvieuses courtes en termes de période de retour. Il apparait clairement que le modèle proposé ne permet pas de répondre à ce critère.

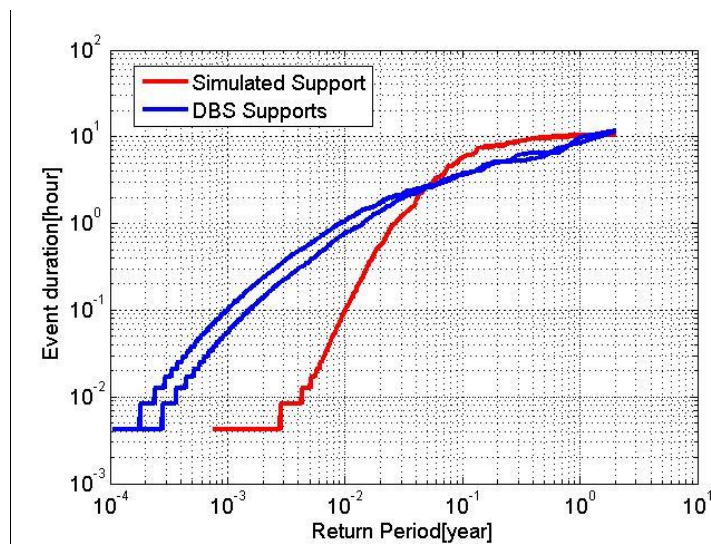


Figure 3.6. Période de retour des durées des événements des séries apprentissages et validation et de la série simulée (même exemple de support utilisé pour tracer la Figure 3.4).

### 3.5.3 Modélisation fine du support de pluie

Afin de remédier au problème identifié précédemment nous proposons une modélisation plus fine du support de pluie. En effet, lors de notre étude de la Figure 3.4, nous avons constaté une difficulté de notre modèle à ajuster les distributions observées des durées des périodes (sèche et humide) avant et après 5 minutes (Figure 3.3). Inspirés par les travaux de Lavergnat et Golé 1998, nous avons analysé séparément les distributions des durées observées inférieures et supérieures à 5 minutes. Nous dénommerons par la suite durées courtes les durées inférieures à 5 minutes et par durées longues les durées supérieures à 5 minutes.

Chaque durée  $d_r$  et  $d_{nr}$  peut être une durée courte ou une durée longue. Le choix entre les deux est fait par tirage aléatoire en considérant les probabilités a priori des durées courtes et des durées longues. Ces probabilités déterminées empiriquement sur notre jeu de données ont pour valeur :  $P_{dr<5\text{ mn}} = 0.87$  et  $P_{dnr<5\text{ mn}} = 0.78$ .

Probabilités P	Durées courtes	Durées longues
Périodes pluvieuses	0.87	0.13
Périodes sèches	0.78	0.22

Tableau 3.2 Probabilités empiriques des durées des périodes courtes et longues pour les périodes pluvieuses et sèches

Comme signalé précédemment, un comportement en loi de puissance est observé. L'étude des distributions associées aux quatre types de durée montre que les distributions des durées  $d$  peuvent être modélisées par une distribution de Pareto généralisée :

$$f(d | k, \sigma, \theta) = \left(\frac{1}{\sigma}\right) \left(1 + k \frac{(d-\theta)}{\sigma}\right)^{-1-\frac{1}{k}} \quad \text{Équation 3.4}$$

Habituellement utilisée pour modéliser la distribution de valeurs au-delà d'un seuil, cette distribution fait appel à 3 paramètres :  $k$  le paramètre de forme,  $\sigma$  est paramètre d'échelle et  $\theta$  le paramètre de position. Comme pour la GEV, on distingue trois types de décroissance de la queue de distribution :

- $k=0$  : Décroissance exponentielle.
- $k>0$  : Décroissance en loi puissance.
- $k<0$  : queues finies.

Les paramètres des quatre lois de Pareto ont été estimés par maximum de vraisemblance sur notre jeu de données. Le Tableau 3.3 présente les 4 jeux de paramètres. Comme dans le cas de la modélisation par une loi GEV, le paramètre  $k$  est positif, signifiant une décroissance en loi puissance.

	Paramètre de la Pareto estimés sur les durés courtes			Paramètre de la Pareto estimés sur les durés longues		
	$k$	$\sigma$	$\theta$	$k$	$\sigma$	$\theta$
Durées sèches $d_{nr}$	1.56	0.32	0.25	1.88	14.35	4.75
Durées pluvieuses $d_r$	1.79	0.31	0.25	0.74	7.77	4.75

**Tableau 3.3 Paramètres de la loi de Pareto estimés sur les durées courtes et longues du support de la série apprentissage**

Finalement la génération du support de pluie est effectuée par l’algorithme suivant :

*Tant que la taille de la série n’est pas atteinte :*

*Si période précédente = pluvieuse alors*

*période courante = sèche*

*Tirer aléatoirement une variable  $X$  suivant une loi de Bernoulli de probabilité*

*$P_{d_{nr} < 5 \text{ mm}}$*

*- Si  $X=1$  : Tirer aléatoirement la durée avec les paramètres de la loi de Pareto correspondante (période sèche, durée courte).*

*- Si  $X=0$  : Tirer aléatoirement la durée avec les paramètres de la loi de Pareto correspondante (période sèche, durée longue).*

*Sinon*

*période courante = pluvieuse*

*Tirer aléatoirement une variable  $X$  suivant une loi de Bernoulli de probabilité*

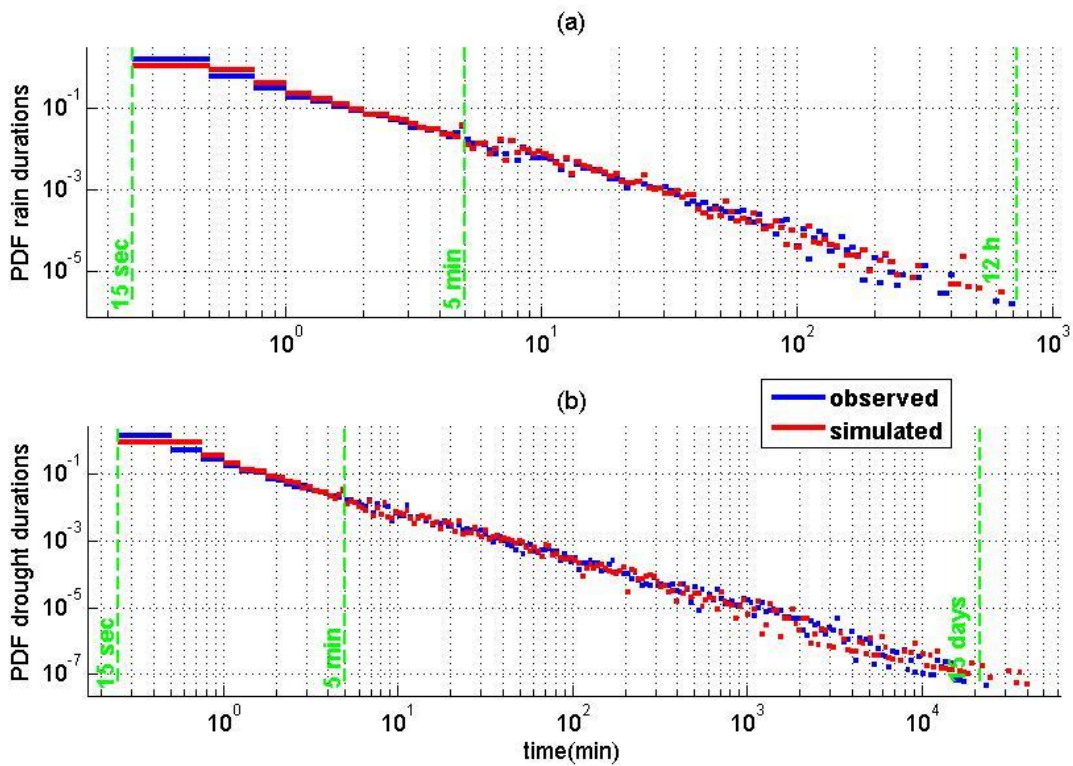
*$P_{d_r < 5 \text{ mm}}$*

*- Si  $X=1$  : Tirer aléatoirement la durée avec les paramètres de la loi de Pareto correspondant (période pluvieuse, durée courte).*

*- Si  $X=0$  : Tirer aléatoirement la durée avec les paramètres de la loi de Pareto correspondant (période pluvieuse, durée longue).*

*Fin Tant que*

La Figure 3.7 ci-dessous montre un exemple de densités de probabilités obtenues sur un support généré à l'aide de l'algorithme ci-dessus et pour une durée de deux ans (courbes en rouge). La figure du haut concerne les périodes pluvieuses et celle du bas les périodes sèches. Il apparaît une très bonne correspondance entre les pdfs simulées et observées (courbes en bleu). Une petite sous-estimation des durées de 15 secondes (pluvieuses et sèches) et une petite surestimation des durées longues sèches supérieures à 15 jours sont également observées. Les mêmes résultats ont été observés sur 100 simulations différentes.



**Figure 3.7 Pdfs normalisées des durées pluvieuses et des durées sèches de la série d'apprentissage et de la série simulée**

Afin de valider le modèle, il convient également de vérifier que le caractère fractal du support simulé est préservé. La Figure 3.8 montre pour 100 réalisations la probabilité empirique qu'il pleuve (courbes grises) en fonction de la résolution. La probabilité moyenne est représentée par la courbe rouge tandis que les deux courbes bleues ont été obtenues sur les séries mesurées. L'allure des courbes est proche et ne diffère pas fondamentalement de celles obtenues avec le modèle précédent (Figure 3.5). On remarque cependant une légère dégradation de la valeur de la co-dimension par rapport à la modélisation précédente. Nous remarquons que le pourcentage de

pluie, pour une résolution donnée, est plus important sur la série observée que sur les séries simulées. Pour la résolution la plus fine (à 15 secondes), nous obtenons des cas de figure différents (pourcentage de pluie dans les supports simulés est identique, inférieur ou légèrement supérieurs à ceux des supports observés). La synthèse des 100 simulations est présentée Tableau 3.4. Il convient de rapprocher la Figure 3.8 avec la Figure 3.5. Il apparaît qu'il n'y a pas de différences fondamentales entre les deux modèles concernant les fréquences de pluie aux différentes résolutions.

	DBS		Support lois GEV			Modélisation fine du support lois Pareto		
			moyenne	min	max	moyenne	min	max
Cf	0.42	0.41	0.40	0.37	0.43	0.38	0.34	0.42
% de pluie à 15s de résolution	4,6%	5.31%	4.29%	3.07%	5.05%	3.85%	3.4%	4.98%

Tableau 3.4 Caractéristiques du support (Co-dimension et probabilité d'occurrence) des séries deux observées (apprentissage et validation) et des 100 supports simulés avec les deux méthodes.

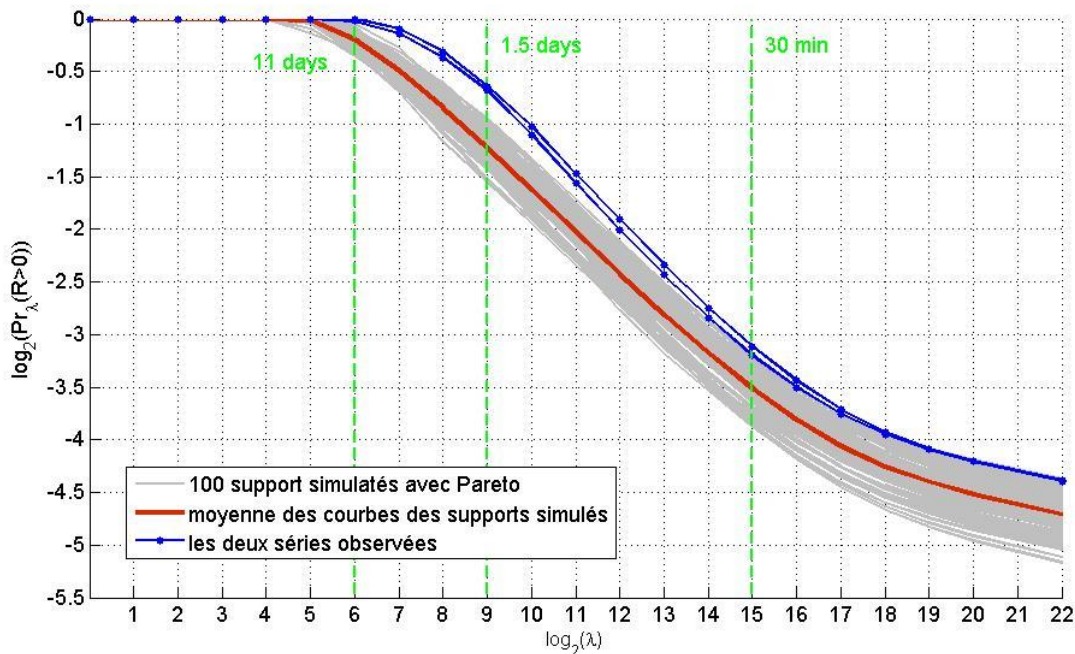


Figure 3.8. Probabilité qu'il pleuve en fonction de la résolution pour les séries d'apprentissage et de validation (bleu) et pour 100 simulations (courbes en gris). La courbe en rouge représente la moyenne des 100 séries simulées.



La Figure 3.9 montre les performances du modèle en termes de périodes de retour, elle est à rapprocher de la Figure 3.6. Contrairement au modèle précédent qui ne fournissait pas de résultats corrects en dessous d'une douzaine de jour, ce modèle permet l'obtention de périodes de retour identiques à celles observées sur l'ensemble de la gamme (15 secondes – deux ans). Par conséquent nous conserverons ce deuxième modèle pour construire notre générateur.

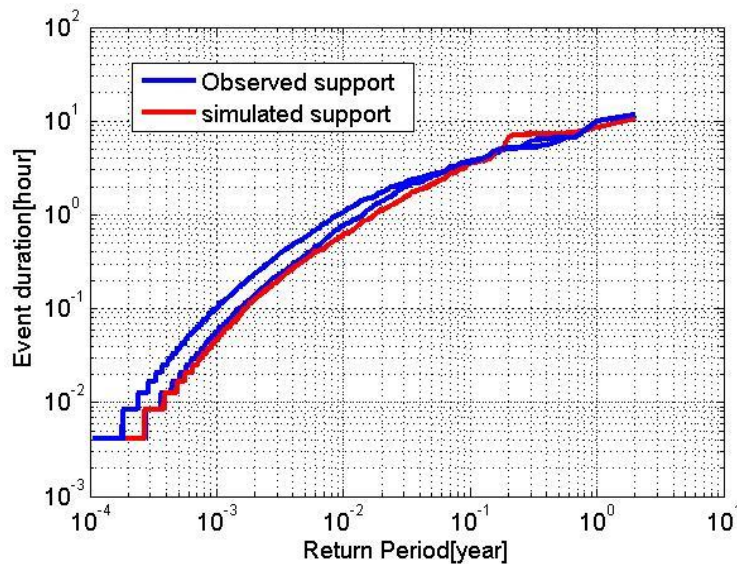


Figure 3.9. Période de retour des durées des événements observés et synthétiques (même exemple de support synthétique utilisé pour tracer la Figure 3.7)

## 3.6 Analyse et modélisation des séries de taux précipitants observés

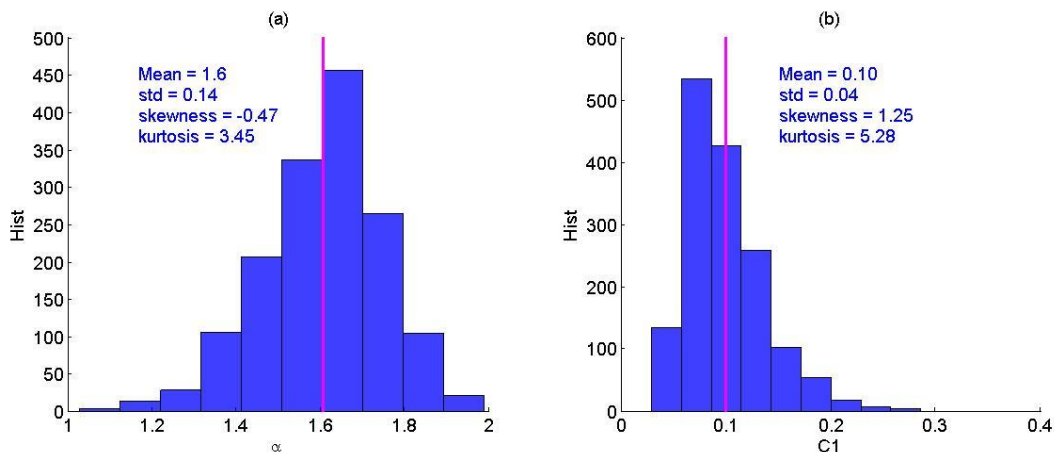
### 3.6.1 Analyse des événements observés et estimation des paramètres multifractals

Conformément à l'hypothèse 2, le taux précipitant peut être modélisé par le modèle FIF.

Dans cette section, nous présentons les résultats de l'analyse multifractale de notre série temporelle issue du DBS. D'après des études réalisées précédemment (de Montera et al., 2009 ; Veneziano et Lepore, 2012; Verrier et al., 2011) un biais important des paramètres multifractals universels existe. Il est dû à la proportion importante de zéros dans les séries analysées. Par conséquent, les paramètres multifractals seront estimés sur les périodes de pluie continues extraites de la série chronologique originale. Nous avons extrait plus de 12000 périodes de pluie

ayant des durées comprises entre 15 secondes et 12 heures. Cependant, seules les périodes ayant une durée de plus de 15 minutes (au minimum 60 observations) sont considérés. En effet, pour des durées plus courtes la gamme d'échelle est trop petite pour estimer correctement les paramètres multifractals.

La procédure classique d'analyse multifractale est utilisée, voir (Schertzer et Lovejoy, 2011 ; Tessier et al., 1996 ainsi que les thèses de S. Verrier et L. de Montera) à chacun des événements constituant la série complète. La Figure 3.10 représente l'histogramme des valeurs des paramètres  $\alpha$  et  $C_1$  estimés sur les événements extraits de la série DBS, les valeurs de la moyenne, l'écart type, le skewness, le kurtosis sont également précisés. Le paramètre  $\alpha$  présente une distribution relativement symétrique (skewness = -0.47) et un kurtosis de 3.45 proche de celui d'une distribution gaussienne (3) et une valeur moyenne de 1.6. La distribution de  $C_1$  est fortement dissymétrique et possède un kurtosis élevé (5.28) et une valeur moyenne de 0.1. Rappelons que l'estimation des paramètres multifractals nécessite d'estimer au préalable les moments empiriques (cf. chapitre 2 et en particulier la section 2.2.2.3) puis l'estimation de la régression d'ordre 1 de ces moments en fonction de la résolution. Seul les paramètres issus de séries de régressions ayant un coefficient de détermination moyen  $R^2 > 0.8$  sont conservés.



**Figure 3.10** Histogramme ainsi que les moments associés pour des valeurs des paramètres  $\alpha$  (a) et  $C_1$ (b) estimés sur les événements de la série apprentissage

Comme nous le verrons plus tard, la pente du spectre est supérieure à 1 dans la gamme 1 minute - 30 minutes. Autrement dit, dans cette gamme le champ est non conservatif et la valeur de  $H > 0$ . La Figure 3.11 présente la fonction de structure d'ordre 1 (cf Equation 2.6 pour  $q=1$ ) calculée sur

166 évènements. Chacun de ces évènements, extrait de la série observée, a une durée aux alentours de 32 minutes.

L'étude de la fonction de structure sur des évènements plus longs (d'une heure) montre un changement rapide de la pente aux alentours de 32 minutes. Pour des durées inférieures à cette valeur le paramètre  $H$  estimé par l'ajustement d'une droite de régression vaut 0.4. Au-delà, le paramètre  $H$  semble décroître. Toutefois il est difficile d'avoir une idée claire de l'évolution de  $H$  pour des durées supérieures à 64 minutes compte tenu du nombre relativement faible d'évènements disponibles conduisant à une mauvaise estimation de  $H$  en terme de variance.

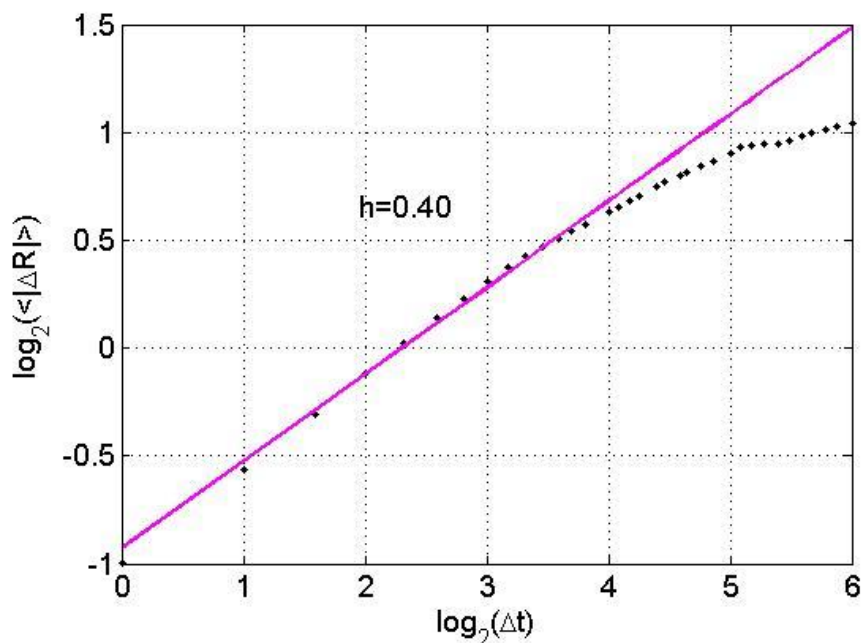


Figure 3.11. Estimation du paramètre  $H$  sur les évènements de pluie

Les paramètres multifractals retenus sont les valeurs moyennes estimées précédemment sur la série observée, soit  $\alpha=1.6$ ,  $C_1=0.10$  et  $H=0.40$ . Il convient cependant de remarquer sur la Figure 3.10 une certaine variabilité des paramètres  $\alpha$  et  $C_1$ , notamment ce dernier dont le kurtosis vaut 5.28 laissant présager l'existence de données « aberrantes ». Cette variabilité marquée est due très probablement au nombre d'observations insuffisant sur certains évènements. Il ressort des travaux antérieurs (thèses de S. Verrier et de L de Montera) relatifs à l'analyse multifractale de taux précipitants à haute résolution que les paramètres multifractals (MU), du régime compris entre 30 secondes et 30 minutes, peuvent être considérés comme constants (thèses de S. Verrier et de L de Montera). Il convient de remarquer en effet que des analyses multifractales faites dans les mêmes conditions (évènements sans zéros) ont fournis un jeu de paramètres très proche du notre

(Montera et al. 2010 ; Verrier et al., 2011) alors que les études qui portent sur l'analyse de séries complètes aboutissent à des jeux de paramètres influencés par les propriétés du support (Verrier et al. 2010). Il est à ce jour impossible de déterminer précisément quelle est la part respective de la méthodologie de l'analyse et de la situation météorologique considérée dans la variabilité des paramètres MU.

### ***3.6.2 Génération des séries temporelles de pluie***

Le modèle FIF permet de prendre en compte les relations d'échelles mais il fournit des séries de valeurs normalisées, de moyenne égale à 1. De ce fait, la distribution des données simulées ne peut pas correspondre à la distribution réelle de la variable observée. Pour prendre en compte cet aspect, nous avons introduit une étape de « dé-normalisation », hypothèse 3. Deux méthodes de calibration ont été développées et testées. Le générateur 1 est basé sur une méthode de calibration globale quantile à quantile entre les distributions de probabilités cumulées (CDF) de la série simulée et de la série observée. La calibration du générateur 2 consiste à modéliser la distribution conditionnelle de l'intensité moyenne d'un événement sachant sa durée. Ensuite, il s'agit de calibrer chaque événement par son intensité moyenne tirée aléatoirement suivant cette loi.

## **3.7 Générateur 1**

### ***3.7.1 Introduction***

Dans cette partie, je présente la première version du générateur proposé, la méthodologie et les résultats obtenus. En ce qui concerne le support, ce générateur est basé sur la première modélisation du support à l'aide de lois GEV (cf. section 3.5.2) bien que comme nous l'avons vu cette modélisation n'est pas totalement satisfaisante (elle correspond cependant à une évolution chronologique de l'étude). Concernant la calibration nous avons utilisé une méthode de type quantile à quantile.

Le schéma suivant résume la méthodologie suivie pour simuler des séries temporelles de longueur quelconque  $L$  :

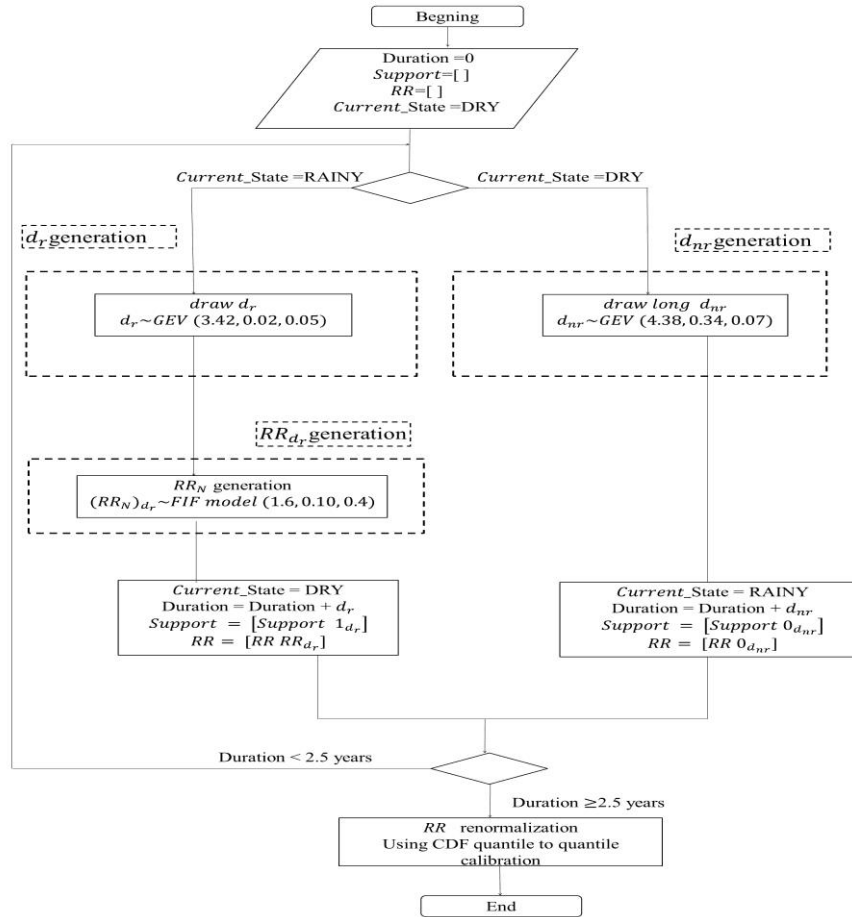


Figure 3.12 Schéma de principe du générateur 1

### 3.7.2 Calibration quantile à quantile

La calibration quantile à quantile permet d’assurer une correspondance exacte entre la CDF estimée sur les données empiriques et la CDF estimée sur les données simulées pour une résolution donnée (15s).

Notre objectif est de générer des séries temporelles réalistes (par rapport aux critères fixés à la fin de la section 3.1) en reproduisant certaines propriétés statistiques des séries observées. Nous allons faire en sorte que la distribution cumulée de nos données simulées soit bien cohérente avec celle obtenue avec la série simulée à la résolution la plus fine. La Figure 3.14 présente la fonction de survie « 1-CDF » des données simulées ainsi que celles des données réelles. Une calibration quantile à quantile est proposée. Pour chaque valeur  $x$  de la série simulée, la calibration vise à

trouver une correspondance entre la valeur  $x$  et le taux de pluie mesuré  $RR$ . Cela est équivalent à trouver une fonction  $\widetilde{RR}(x)$  donnant une valeur réelle  $RR$  pour chaque valeur simulée  $x$ :

$$\widetilde{RR}(x) = g^{-1}(f(x))$$

avec  $g$  la distribution cumulée (CDF) des taux de pluie mesuré  $RR$ , et  $f$  celle des valeurs simulées par le modèle FIF.

La Figure 3.13 montre la fonction de calibration  $\widetilde{RR}(x)$  nous remarquons trois régimes :

- le plateau à gauche (les valeurs faibles) est lié à la limitation de mesure de l'appareil car celui-ci possède un seuil de détection, au-dessous duquel les gouttes ne sont pas détectées. Les valeurs simulées inférieures à ce seuil de détection sont recalées avec cette valeur minimale.
- Le plateau à droite correspond aux valeurs fortes, il est lié à la valeur maximale mesurée qui est dépendante de la durée d'observation de la série temporelle. Il est à noter que les événements extrêmes étant rares, l'utilisation d'une série de deux ans est une limitation. En effet, les valeurs simulées supérieures à cette valeur maximale observée sont recalées avec cette valeur maximale. (Cette méthode de dénormalisation/calibration induit une « sturation ». En effet, quelle que soit la dynamique du modèle multifractal, les taux de pluie simulés ne pourront pas excéder la plus grande valeur observée qui est de l'ordre de 150mm/h.)
- Le régime linéaire entre les deux plateaux permet de recaler les valeurs théoriques entre le minimum et le maximum observés.

L'utilisation de séries plus longues serait donc préférable afin de pouvoir reproduire les événements extrêmes. Comme attendu, après calibration la fonction de survie calibrée de la série simulée est identique à la fonction de survie empirique (Figure 3.14)

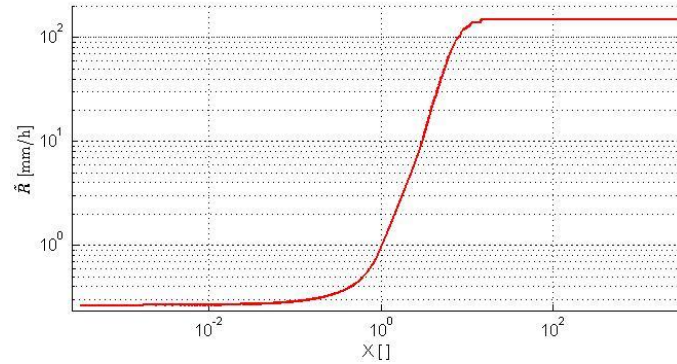


Figure 3.13. Fonction de calibration  $\widetilde{RR}(x)$  :  $\widetilde{R}$  (RR observé) versus  $X$  (x simulé)

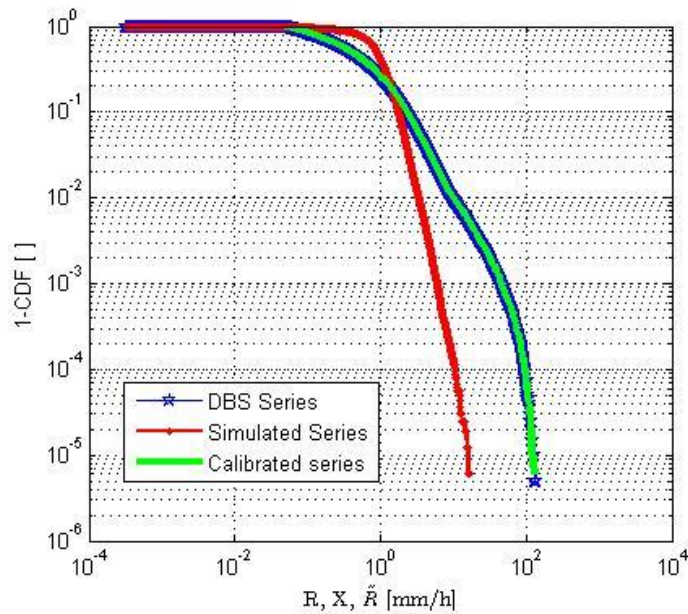
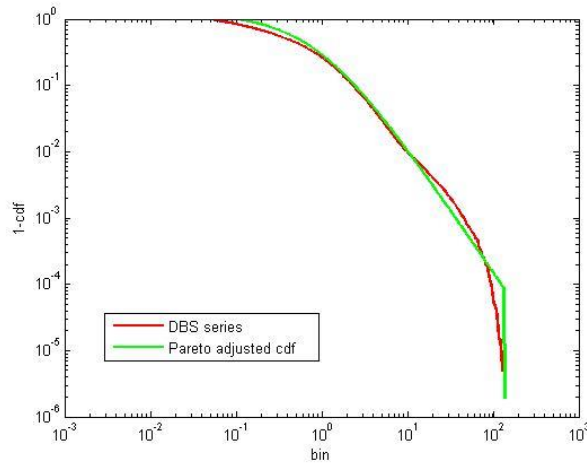
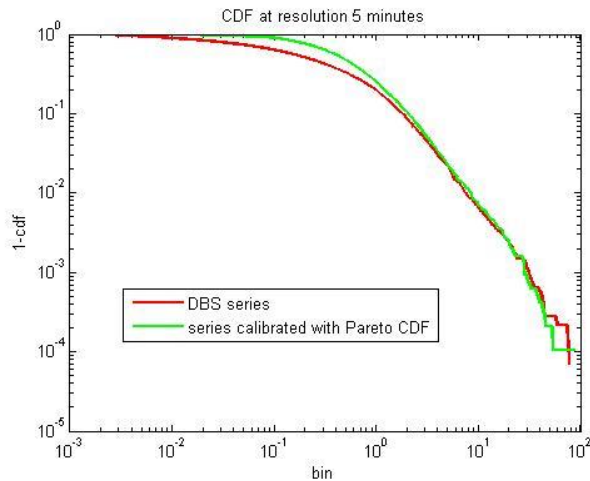


Figure 3.14 Fonction de survie de la série temporelle issue du DBS, et celle de la série simulée avant et après calibration

La calibration réalisée à partir d'une CDF empirique nécessite de disposer d'une telle série pour son utilisation. On peut cependant faire l'hypothèse que la CDF observée suit une loi de Pareto tronquée. (Cette troncature est due au phénomène de « saturation » mentionné précédemment.) La Figure 3.15 montre en effet qu'il est possible d'obtenir un bon ajustement de la loi de Pareto tronquée au maximum observé dans la série DBS (de paramètre de forme  $k=0.23$  et de paramètre d'échelle  $\sigma=0.67$ ) à la distribution des taux de pluie de la série DBS pour une résolution de 15 secondes. En dégradant la résolution des séries calibrées avec la CDF Pareto tronquée à 5 minutes, nous retrouvons la bonne fonction de survie (sans troncature apparente), un exemple est présenté sur la Figure 3.16.



*Figure 3.15. Fonction de survie de la série DBS et ajustement de loi de Pareto tronquée à la résolution de 15 secondes (les bins sont en mm/h).*



*Figure 3.16. Fonction de survie de la série DBS et exemple d'une réalisation calibrée à l'aide de la loi de Pareto tronquée à une résolution dégradée de 5 minutes)*

L'étape de calibration du générateur 1 peut donc être réalisée soit à partir d'une CDF empirique soit à partir d'une loi de Pareto (retenue pour la qualité de son ajustement) tronquée à 4 paramètres (3 paramètres de la Pareto et la valeur à laquelle la Pareto est tronquée).

### 3.7.3 Performances

Le simulateur étant calibré nous analysons les séries simulées. Pour cela, nous avons généré, avec cette première version du générateur, 100 séries chacune d'une durée de 2.5 ans à la résolution de 15s.



Par ailleurs, il convient de signaler que la calibration par la distribution empirique ou par la loi de Pareto tronquée donne des résultats identiques. La distribution empirique est utilisée ci-dessous.

### 3.7.3.1 Analyse spectrale

L'analyse spectrale permet de vérifier si les différents régimes d'invariance d'échelle, mentionnés en introduction de ce chapitre, sont bien présents dans les séries simulées et de mettre en évidence l'effet éventuel de la calibration sur les régimes de scaling. Les spectres des séries simulées avant et après calibration sont donc comparés à ceux de la série DBS ( Figure 3.17 ). Nous remarquons que la série DBS possède trois régimes de scaling. Il convient donc de vérifier les propriétés spectrales pour chaque régime. Aux petites échelles le régime représente les propriétés intra évènements. Dans cette gamme, la pente du spectre  $\beta$  est supérieure à 1, ce qui implique que nous sommes dans le cas non conservatif ( $\beta$  supérieur à 1 implique H différent de 0 cf. section 2.2.2). Aux grandes échelles (échelles au-delà de 30 minutes), le régime représente les propriétés inter évènements.

La Figure 3.17 montre une différence remarquable entre le spectre de la série observée DBS et celui de la série simulée avant la phase de calibration. La calibration améliore notablement le spectre en corrigeant les régimes de scaling comme le confirme le Tableau 3.5 qui compare les pentes des trois régimes de la série observée DBS et celles de la série synthétique simulée avant et après calibration. Nous pouvons donc conclure que le simulateur permet de reproduire les cassures du spectre aux mêmes endroits que ceux observés à partir des données réelles. Néanmoins, les pentes sur le 1<sup>er</sup> régime ne sont pas bien restituées.

La Figure 3.18 et le Tableau 3.5 présentent les spectres obtenus sur 100 simulations avec ce générateur ainsi que ceux des séries d'apprentissage et de validation. Comme nous pouvons l'observer sur cette figure, les régimes de scaling pour des durées supérieures à 30 minutes sont bien reproduits. Pour le régime compris entre une et trente minutes la pente moyenne s'avère être plus faible que celles observées, un décrochage apparaît pour des durées inférieures à la minute. Ce décrochage implique que les hautes fréquences sont mal restituées, et plus précisément que les variations rapides du taux précipitant sont sous représentées. Il est difficile d'avancer une explication, on peut cependant penser que la méthode de calibration utilisée détruit partiellement les propriétés d'invariance d'échelle. En effet, la Figure 3.13 montre clairement que la

transformation  $\widetilde{RR}(x)$  obéit à une loi puissance dans sa partie centrale, préservant ainsi les propriétés d'invariance, alors qu'elle possède ensuite deux coudes sur les côtés détruisant les propriétés d'échelle.

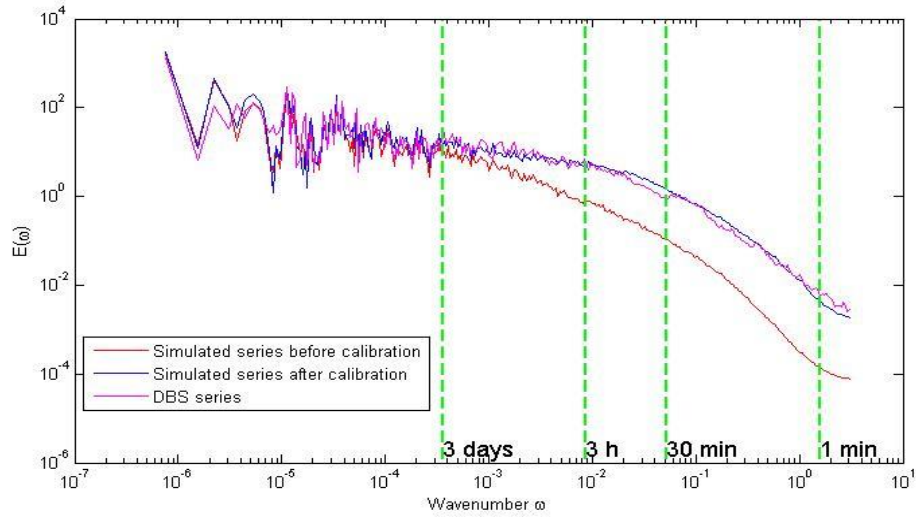


Figure 3.17. Spectre de puissance de la série DBS (bleu) ayant servi pour apprendre les paramètres du modèle, d'une série simulée avant (rouge) et après calibration (rose)

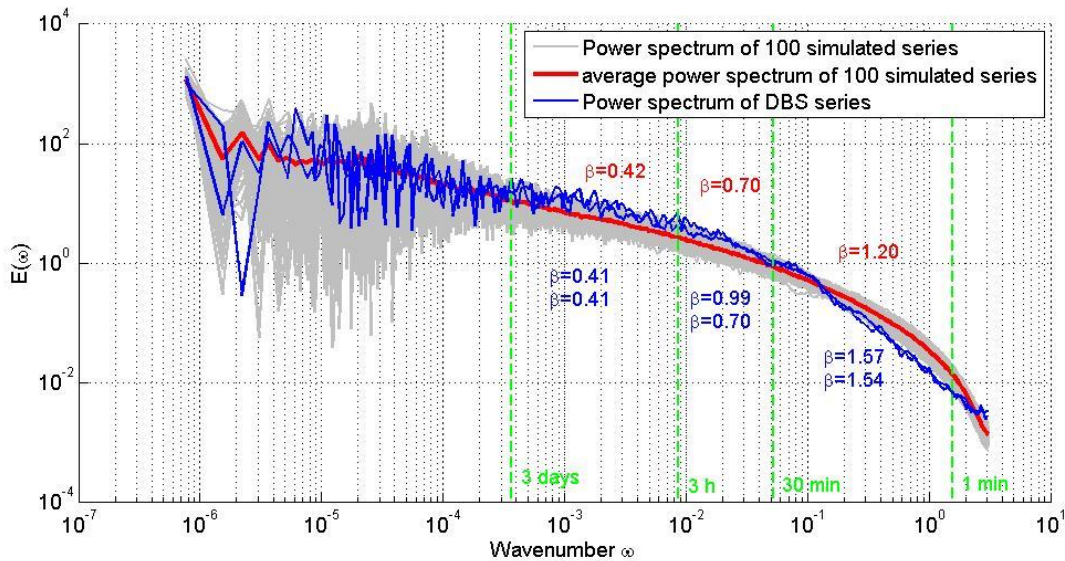


Figure 3.18. Spectre en loi de puissance de 100 séries simulées avec le générateur 1 et des séries apprentissage et validation

	$\beta$ in regime scaling 1min $\rightarrow$ 30 min				$\beta$ in regime scaling 30 min $\rightarrow$ 3 h				$\beta$ in regime scaling 3 h $\rightarrow$ 3 days			
DBS	1.57				0.99				0.41			
	mean	std	min	max	mean	std	min	max	mean	std	min	max
100 séries Simulées avant calibration	1.41	0.05	1.12	1.60	1.40	0.08	1.2	1.45	1.20	0.05	0.80	1.34
100 séries Simulées après calibration	1.20	0.09	1.10	1.40	0.70	0.09	0.5	0.98	0.42	0.05	0.30	0.54

Tableau 3.5 Pentas des spectres de la série apprentissage et des séries simulées par le générateur 1 avant et après calibration

### 3.7.3.2 Analyse multifractale

Il convient de vérifier si l'étape de calibration, qui comme nous l'avons entrevu, a des conséquences sur la forme du spectre, n'altère pas les paramètres multifractals. L'analyse multifractale des évènements de pluie des 100 séries simulées a été réalisée. Comme pour les données observées, seules les périodes ayant une durée de plus de 15 minutes sont considérées. Environ 1500 périodes continues ont été analysées pour chaque série simulée permettant ainsi de calculer pour chacune d'elle une valeur moyenne des  $\alpha$  et des  $C_1$  estimés. La Figure 3.19 montre les distributions des paramètres  $\alpha$  et  $C_1$  estimés sur ces séries d'une part et sur la série apprentissage d'autre part. Pour les deux paramètres, il existe une différence assez importante entre les distributions que ce soit en termes de forme ou de valeurs moyennes. Concernant  $\alpha$ , sa valeur moyenne sur les 100 réalisations vaut  $\alpha = 1.38 \mp 0.05$ . Cette valeur est nettement plus faible que celle utilisée dans le modèle MU ( $\alpha = 1.6$ ). La distribution issue de la série DBS est légèrement asymétrique (Skewness= -0.47) mais possède un kurtosis proche de la normalité (3.45) tandis que celle issue des données simulées est symétrique avec un kurtosis un peu inférieur à 3 (2.6). Concernant le paramètre  $C_1$ , sa valeur moyenne vaut  $C_1 = 0.20 \mp 0.02$  et est nettement plus importante que celle utilisée dans le modèle ( $C_1 = 0.10$ ). Le kurtosis sur les données observées est relativement élevé (5.28) signifiant une distribution assez ramassée sur elle-même alors qu'il a une valeur proche de 3 sur les données simulées. Ces remarques

impliquent que les propriétés intra-événement ne sont pas bien restituées et que les événements synthétiques sont plus lisses que les événements réels. Il convient de remarquer que ce résultat est en accord avec la remarque faite au paragraphe précédent concernant le décrochage du spectre aux hautes fréquences. Concernant le paramètre H, une valeur moyenne de  $H = 0.33 \mp 0.02$ , inférieure à celle estimée sur les observations, est obtenue. Cela confirme également la plus faible pente obtenue sur le premier régime de scaling (1-30 minutes) du spectre. La calibration dégrade donc notablement la restitution des paramètres multifractales.

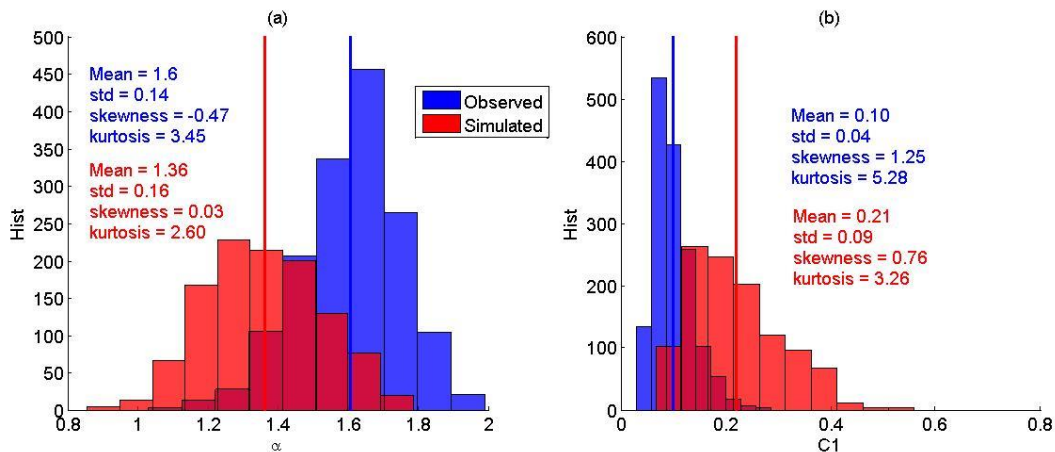


Figure 3.19. Distribution des paramètres multifractals obtenue sur 100 séries simulées.

### 3.7.3.3 Autocorrélation

L'autocorrélation, sur une échelle de 50 heures, des séries observées et un exemple d'une série simulée est présentée Figure 3.20. Une bonne correspondance est observée entre l'autocorrélation des données observées et les données synthétiques sur tous les retards.

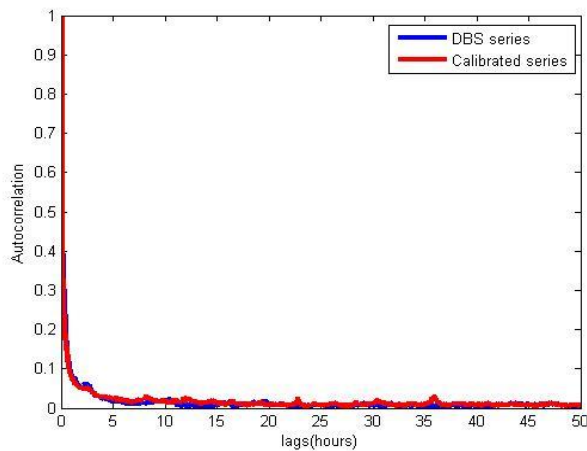


Figure 3.20. Comparaison de l'autocorrélation entre les données DBS observées et les données synthétiques

### **3.7.4 Analyse critique**

Le générateur proposé possède au total 12 paramètres (6 pour les 2 lois GEV concernant les durées et 3 pour la génération des événements et 3 paramètres de la Pareto permettant d'estimer la distribution). Si comme le supposent les études mentionnées en introduction de ce chapitre, les 3 paramètres multifractals caractéristiques des lois d'échelles intra-événements sont indépendants de la région climatique et peuvent être considérés comme universels, les autres paramètres qui caractérisent l'intermittence du processus et la distribution des taux précipitants dépendent quant à eux de la région considérée. En termes de paramètres, notre modèle nécessite donc un nombre relativement restreint de paramètres à ajuster en fonction de la localisation (9 paramètres). Cependant comme nous l'avons constaté plus haut ce générateur ne permet pas de restituer de façon satisfaisante certaines propriétés des séries observées que ce soit en termes de distributions des durées des périodes pluvieuses et sèches, de spectre ou de paramètres multifractals. Pour cette raison, nous proposons dans les pages suivantes une autre modélisation du générateur que nous dénommerons «générateur 2» (Akrou et al., 2015).

## **3.8 Générateur 2**

### **3.8.1 Introduction**

Afin de résoudre les problèmes rencontrés précédemment dans le générateur 1, il convient d'améliorer deux aspects du modèle :

- La génération du support
- La calibration

Concernant la génération du support nous allons utiliser le modèle présenté au paragraphe 3.5.3 intitulé « Modélisation fine du support de pluie » dans laquelle une distinction est faite entre les durées dite courtes ( $< 5$  mn) et les durées dites longues. Concernant la calibration, nous avons dé-normalisé les taux précipitants simulés  $RR_N$  à l'aide d'une modélisation durée - intensité moyenne des événements observés présentée dans le paragraphe suivant.

Le schéma de la Figure 3.21 résume la méthodologie suivie.

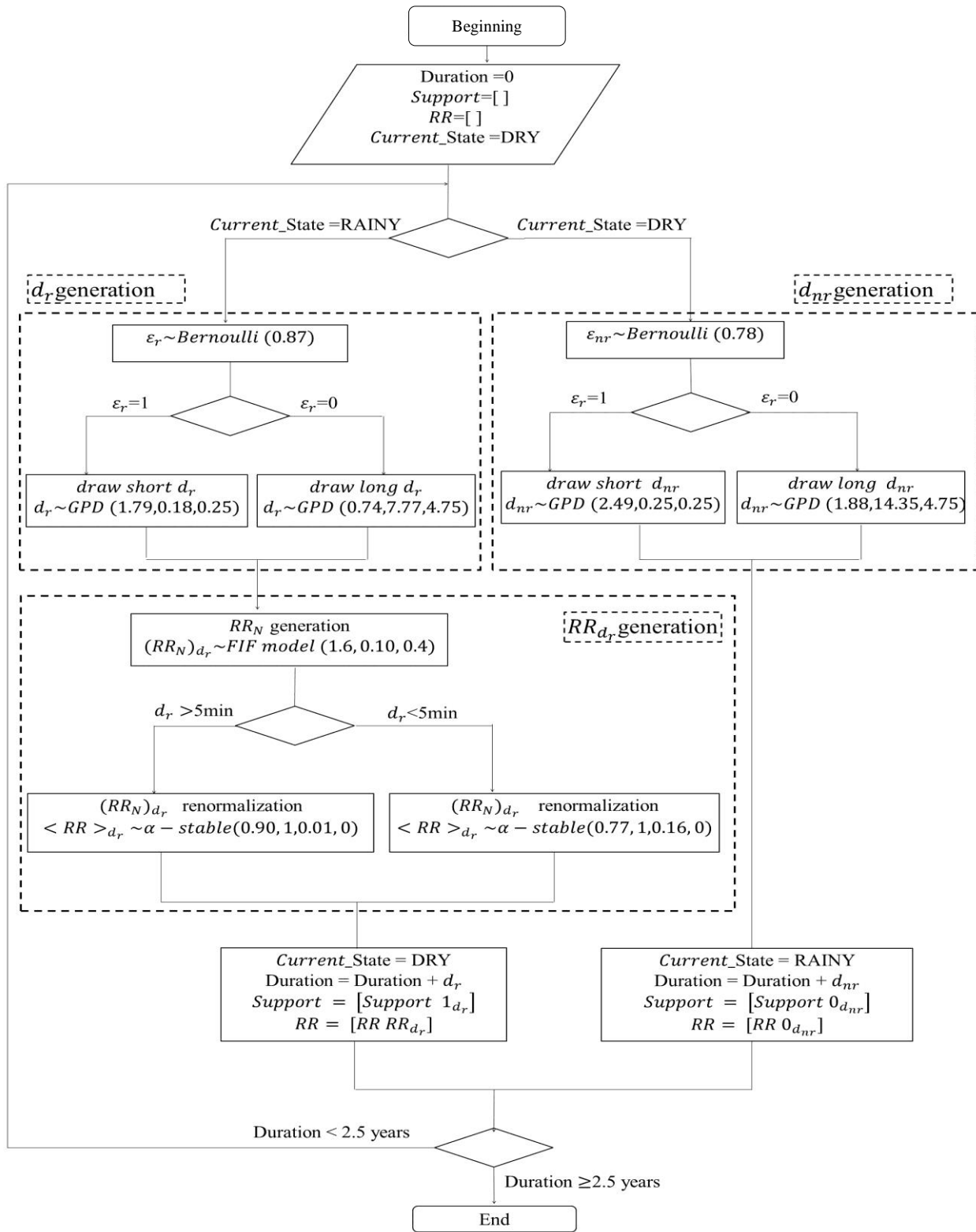


Figure 3.21 Schéma de principe du générateur 2

### 3.8.2 Relation durée/intensité

Chaque période pluvieuse est caractérisée par sa durée  $d_r$  et le taux précipitant moyen correspondant  $\langle RR \rangle_{d_r}$  (moyenne du taux de pluie sur la période pluvieuse). La relation entre les durées  $d_r$  et  $\langle RR \rangle_{d_r}$  a fait l'objet de nombreuses études. Certaines d'entre elles concluent qu'il n'y a pas de relation entre les deux, et par conséquent les deux variables peuvent être modélisées indépendamment l'une de l'autre. C'est le cas de l'étude de Menabde et Sivapalan, (2000) qui ont étudié des séries temporelles de données ayant une résolution de 6 minutes. Ils ont démontré que la distribution Levy-stable permet de mieux reproduire les distributions des durées et des intensités des événements par rapport à une distribution de type gamma. Au contraire, De Michele et Salvadori (2003) ont montré que l'indépendance s'avère être une hypothèse non réaliste. Ils ont décrit la dépendance entre les deux variables (durée et intensité) au moyen d'un 2-copula (Pareto marginale à la fois pour la durée et l'intensité moyenne par événement). Ils ont montré que 2-copula permet de décrire non seulement la dépendance mais aussi de reproduire la variabilité marginale des durées des événements et de l'intensité moyenne par événement. Il est à noter cependant que les différentes études ont été réalisées dans des conditions différentes de celles considérées ici notamment en ce qui concerne la résolution des données, la taille des séries observées, la définition d'un événement.

Pour étudier la relation durée / intensité moyenne, nous avons extrait toutes les périodes de pluie selon la définition donnée au paragraphe 3.3 et en gardant toutes les périodes quelle que soit leur durée. La Figure 3.22 présente l'intensité moyenne  $\langle RR \rangle_{d_r}$  par période pluvieuse en fonction de sa durée  $d_r$  (points bleu). La moyenne pour chaque classe de durée est également reportée (points en rouge).

D'après la Figure 3.22, nous pouvons remarquer un changement net de comportement aux alentours de 5 minutes. Il semble en effet exister une dépendance entre durée et  $\langle RR \rangle$  pour des durées inférieures à 5 minutes et pour des valeurs de  $\langle RR \rangle$  comprises entre 0.1 et 0.3 mm/h. Au-delà d'une durée de 5 minutes, la distribution conjointe ne montre pas de dépendance entre les deux variables.

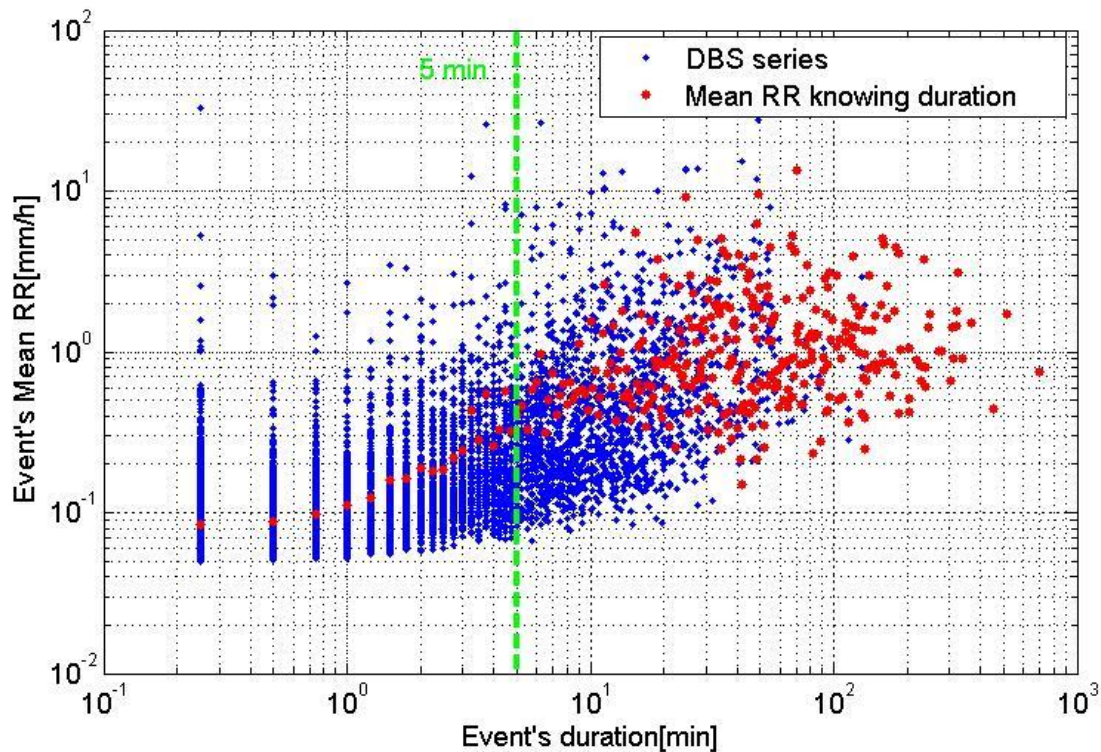


Figure 3.22 Intensité moyenne en fonction de la durée obtenue sur l'ensemble des périodes de pluie de la série apprentissage (en bleu). Intensité moyenne sur l'ensemble des périodes de même durée en rouge.

Les remarques précédentes nous conduisent à faire les trois hypothèses suivantes :

- Il existe deux comportements différents avant et après 5 minutes. (Ici, nous ne tranchons pas entre les raisons physiques et la discrétisation des données.)
- Concernant les durées longues ( $> 5\text{mn}$ ) il est clair que  $\langle RR \rangle$  peut être modélisé indépendamment de la durée.
- Bien qu'une dépendance semble exister entre les 2 variables pour des durées courtes, il s'est avéré que la prise en compte de cette dépendance n'améliorait pas de façon significative les performances du modèle par rapport à une hypothèse d'indépendance. La complexification du modèle n'ayant pas d'apport tangible à ce niveau, nous avons donc opté pour cette dernière hypothèse et donc pour le modèle le plus simple.

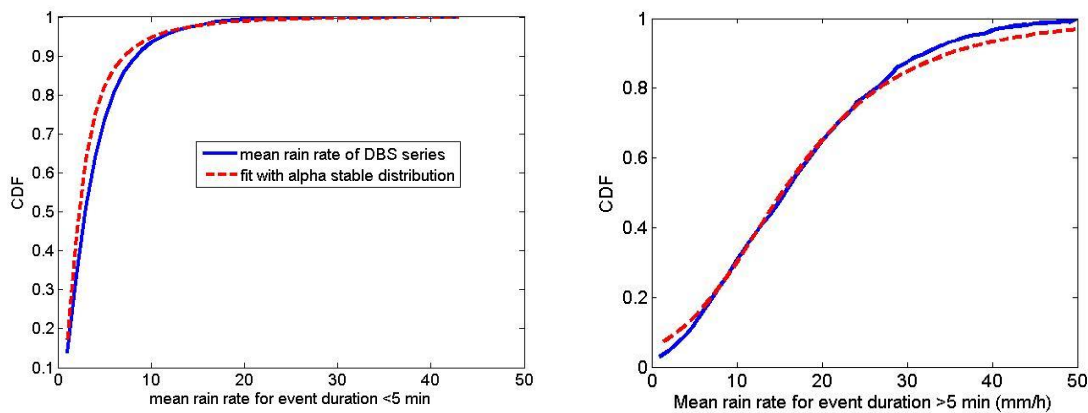
Il est à noter qu'un comportement similaire à celui observé sur les moyennes des taux de pluie est aussi observé sur leurs variances.



Ces résultats sont cohérents avec ceux de (Menabde et Sivapalan, 2000) qui ont montré que les intensités des évènements peuvent être modélisées indépendamment de la durée à partir de données à 6 minutes de résolution. Nous montrons que la distribution Alpha-stable avec les quatre paramètres ( $\alpha, \beta, \gamma, \delta$ ) permet de bien décrire la loi de distribution des taux précipitant moyen par période, aussi bien pour les périodes de durées supérieures à 5 minutes que pour celles de courtes durées . Le paramètre d'asymétrie (skewness)  $\beta$  a été fixé à 1, et le paramètre de localisation  $\delta$  a été fixé à 0, deux valeurs ayant un sens physique. Les deux autres paramètres sont estimés sur les périodes de pluie de la série apprentissage. Tous les paramètres obtenus sont résumés dans la table suivante. La Figure 3.23 illustre l'adéquation entre les CDF empiriques et les CDF alpha-stable obtenues.

Paramètres des intensités moyennes des durées courtes inférieures à 5 minutes				Paramètres des intensités moyennes des durées longues supérieures ou égales à 5 minutes			
$\alpha_r$	$\beta$	$\sigma$	$\delta$	$\alpha_r$	$\beta$	$\sigma$	$\delta$
0.90	1	0.01	0	0.77	1	0.16	0

**Tableau 3.6** Paramètres de la loi alpha stable estimés sur les intensités moyennes des évènements de durées courtes et longues.



**Figure 3.23.** CDF des intensités moyenne des périodes de pluie de la série apprentissage et de la loi Alpha-stable correspondant

### 3.8.3 Performances

Comme dans le cas du générateur 1, nous avons simulé 100 séries de deux ans et demie, à la résolution de 15 secondes, que nous avons analysées. Dans ce qui suit nous présentons les résultats obtenus.

#### 3.8.3.1 Analyse spectrale

La Figure 3.24 illustre, en échelle log-log, les spectres de puissance des deux séries observées, de 100 séries simulées de deux ans et demie. Le spectre moyen des 100 simulations est présenté en rouge. Nous remarquons une certaine variabilité des spectres (courbes en gris) des séries simulées. Malgré quelques disparités, le spectre moyen est cohérent avec les spectres des mesures. Le régime de scaling correspondant aux échelles de plus de 3 jours est principalement dû au support de pluie, et celui de 1 à 30 minutes, comme dit précédemment, caractérise l'intra des évènements. Sur le régime intra-évènement, la pente est égale à 1.57, la pente du spectre moyen est égale à 1.63 ; elles sont très proches ce qui confirme aussi une valeur de H bien restituée (comme nous le verrons après). La restitution de ce premier régime de scaling confirme que les évènements sont bien simulés. Les régimes de scaling dus au support de pluie (au-delà de 30 minutes) sont bien restitués sur les séries simulées et montrent que le support est bien simulé. Comparé à l'ensemble des simulations, nous remarquons que les spectres estimés sur les observations peuvent être une réalisation possible de notre modèle.

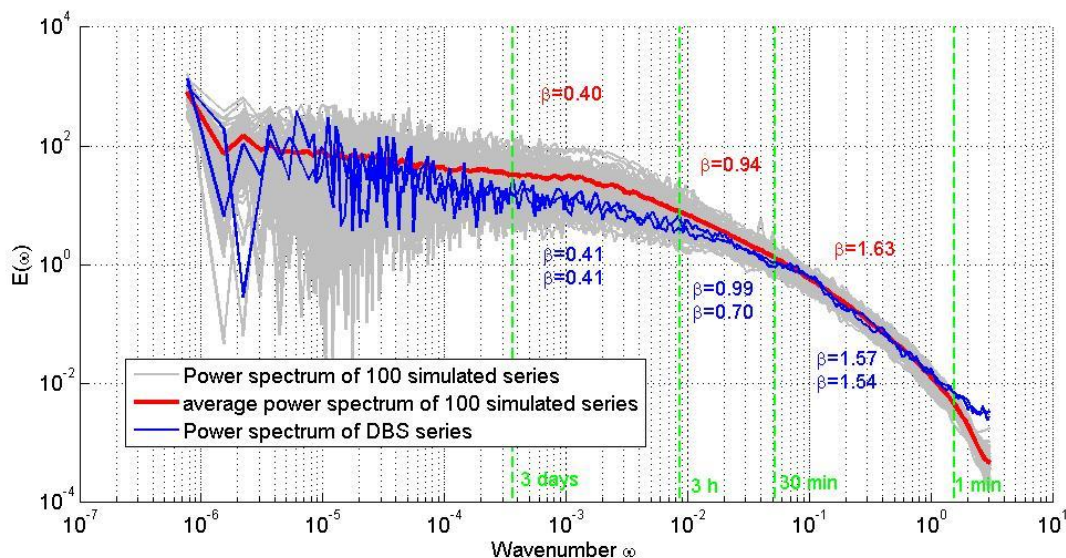


Figure 3.24. Spectre des séries apprentissage et validation, spectres des 100 séries simulées, spectre moyen des séries simulées

### 3.8.3.2 Analyse multifractale

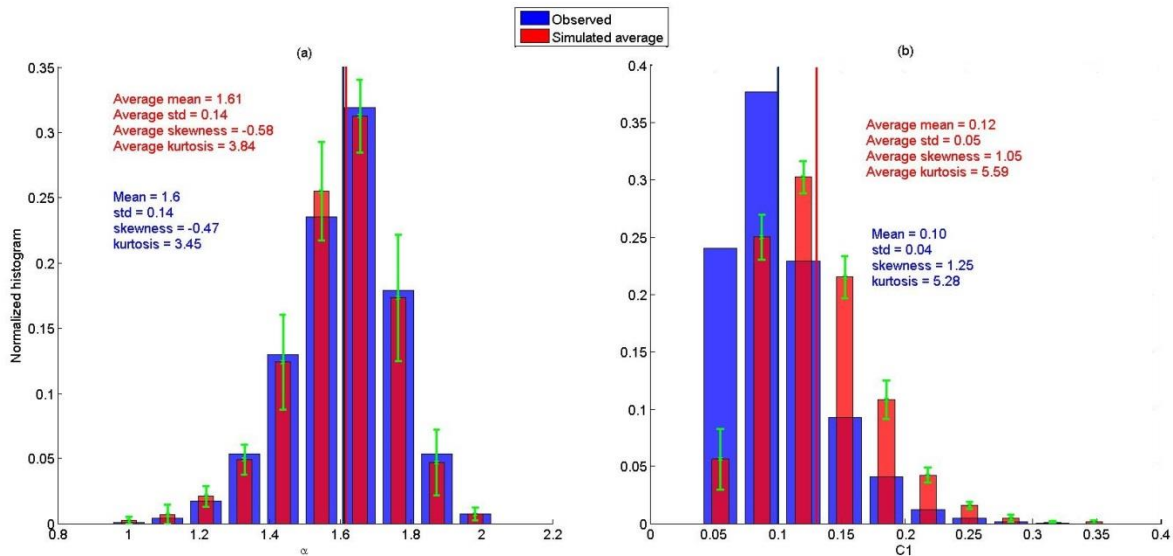
Nous avons effectué l'analyse multifractale des périodes continues de pluie. Comme pour les données observées, seules les périodes ayant une durée de plus de 15 minutes (60 observations) sont considérées conduisant en moyenne à 1500 évènements pour chaque série simulée et permettant ainsi de calculer pour chacune d'elle une valeur moyenne des paramètres  $\alpha$  et des  $C_1$  estimés. Au total, nous avons 100 couples de valeurs  $\alpha$  et  $C_1$ , chacune étant une moyenne de plusieurs valeurs (nombre d'évènements extrait de la série pris en compte dans le calcul).

Le Tableau 3.7 et la Figure 3.25 montrent la moyenne, l'écart type, le minimum et le maximum des 100 couples  $\alpha$  et  $C_1$  obtenue sur les 100 séries de 2 ans et demi simulées ainsi que leur distribution. La moyenne de  $\alpha$  (1.61) et la moyenne des  $C_1$  (0.11) sont cohérentes avec les valeurs estimées sur les observations et utilisées dans le modèle FIF, soit  $\alpha = 1.60$  et  $C_1 = 0.10$ .

100 séries simulées avec le générateur 2	$\alpha$ moyen estimé sur les évènements de chaque série				$C_1$ moyen estimé sur les évènements de chaque série			
	mean	std	min	max	mean	std	min	max
	1.61	0.005	1.59	1.63	0.11	0.002	0.09	0.12

**Tableau 3.7** Analyse des paramètres des 100 séries simulées avec le générateur 2. Pour chaque série simulée on considère la moyenne des paramètres estimés sur l'ensemble des périodes continues de pluie de durées supérieures à 15 mn.

Pour les deux paramètres l'écart type est le même que celui estimé sur les observations. Les distributions des deux paramètres estimés sur les séries simulées sont cohérentes avec les distributions des paramètres estimés sur les observations



**Figure 3.25.** Distribution des paramètres multifractals estimés sur les événements extraits de la série observée (histogramme en bleu) et la moyenne par bin des 100 séries simulées avec le générateur 2 (histogramme en rouge) avec l'écart type par bin (en vert)

Pour le paramètre  $H$ , nous avons estimé une valeur moyenne de 0.41 sur toutes les séries avec un écart type égal à 0.01, ce qui reste proche des valeurs observées (0.40).

### 3.8.3.3 Distribution de probabilité

Nous avons calculé les distributions (fonctions de survie empiriques noté ESF) de trois quantités: des taux précipitants globaux  $RR$  de toute la série, des durées des événements  $d_r$  et des hauteurs d'eau correspondantes des 100 séries simulées à 15 secondes. Il est important de vérifier que les propriétés d'échelle vérifiées à 15 secondes sont aussi valides à des résolutions plus grossières et que notre modèle peut simuler des séries chronologiques qui sont compatibles avec les mesures effectuées à des résolutions plus grossières. Pour cela, nous avons aussi calculé les fonctions de survie des trois quantités sur les 100 séries simulées dégradées à des résolutions de 5 minutes, 1 heure et une journée. Nous avons ensuite comparé les résultats aux fonctions de survie des séries observées dégradées aux mêmes résolutions. Les résultats sont présentés à la Figure 3.26. Pour chaque sous-figure de (a) à (l), les fonctions de survie correspondant à 100 séries simulées sont représentées (courbes en gris), la courbe en rouge présente la courbe moyenne des courbes grises, et les deux courbes en bleu présentent les observations. Les sous-figures sur la première ligne (a),

(b) et (c) montrent les résultats à 15 secondes. Chaque ligne suivante correspond aux fonctions de survie des séries dégradées à une résolution donnée (5 mm, 1 heure et 1 jour).

Un bon agrément est observé entre les ESFs calculées sur les observations et celles calculées sur les simulations, ce qui signifie que les relations à différentes échelles temporelles et taux précipitants sont bien représentées dans notre modèle et les observations peuvent être une réalisation possible du modèle. Le bon comportement des trois quantités à différentes résolutions valide la capacité de notre modèle à générer des séries temporelles de précipitations à différentes résolutions. Les ESFs illustrées dans la 2<sup>ème</sup> colonne (Figure 3.26 (b), (e), (h) et (k)) sur les durées des périodes pluvieuses confirment l'information révélée par la Figure 3.7. Malgré leur tendance à surestimer légèrement les courtes durées, les ESFs des durées pluvieuses observées sont très proches de celles simulées. Pour les grandes valeurs, une petite différence peut être observée entre les ESFs des observations et celles des simulations. Une légère courbure peut également être observée dans l'une des courbes mesurées dans la Figure 3.26 (e) - comme indiqué précédemment, ce type d'effet peut être expliqué par la variabilité naturelle de la simulation. Les ESFs, des hauteurs d'eau des événements, illustrées dans la 3<sup>ème</sup> colonne (Figure 3.26 (c), (f), (i) et (l)) sont proportionnelles au produit de la durée de la pluie et du taux de pluie moyen ( $\langle RR \rangle_{d_r} \times d_r$ ).

Étant donné que la consistance du support de pluie a été démontrée, cela peut être une méthode supplémentaire pour valider la relation durée/intensité moyenne.

La 1<sup>ère</sup> colonne (Figure 3.26 (a), (d), (g) et (j)) représente les ESFs des taux précipitants. Nous constatons une bonne correspondance entre les mesures et les simulations.

Cette Figure 3.26 illustre donc une bonne correspondance entre les mesures et les simulations à plusieurs résolutions et offre un critère supplémentaire de validation globale de notre modèle et des hypothèses sous-jacentes.

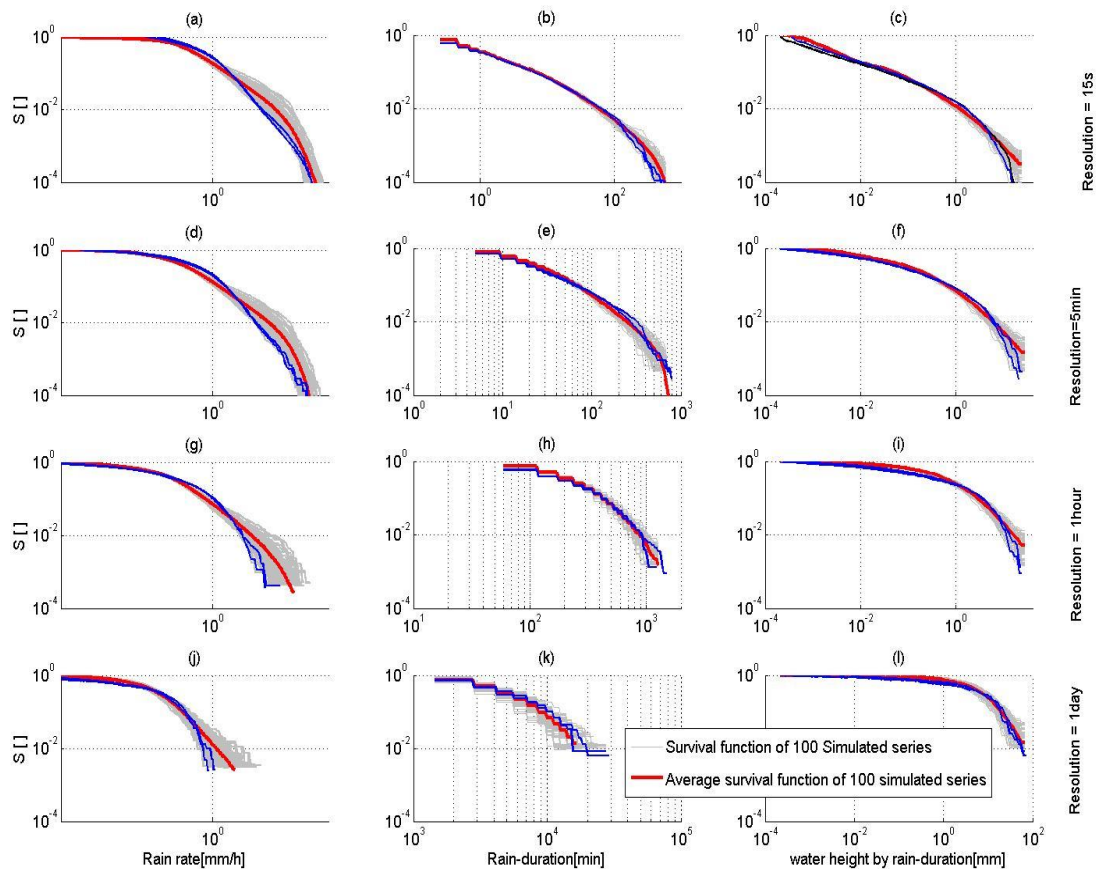


Figure 3.26. De gauche à droite: Fonctions de survie empiriques des taux de pluie RR, des durées des périodes pluvieuses et de la hauteur d'eau correspondante. De haut en bas: les fonctions de survie calculées à différentes résolutions 15sec, 5 min, 1 heure et 1 journée.

### 3.9 Conclusion

L'objectif principal de ce chapitre était de simuler des séries temporelles de précipitations avec des propriétés statistiques (distribution et invariance d'échelle) réalistes représentant le climat d'une région particulière. Notre méthodologie est basée sur deux principales étapes, à savoir la simulation de support de pluie et de taux précipitant réalistes. Ces étapes sont basées sur plusieurs hypothèses (dédiées de la littérature) qui ont été validées sur les données disponibles.

Deux générateurs de séries temporelles de précipitation, basés sur la notion des cascades multifractales ont été proposés. Le premier repose sur la simulation du support avec une distribution GEV, et une étape de calibration quantile à quantile. Cette version simule des séries

temporelles avec une très haute résolution (15 secondes) et a permis de valider certaines propriétés statistiques des précipitations vérifiées la section performance. Cependant, certaines limitations, qui concernent les propriétés spectrales et multifractales aux échelles inférieures à 30 minutes, ont été observées sur cette première version du générateur qui présente l'avantage toutefois de ne nécessiter que 9 paramètres.

Afin de remédier à ces limitations, nous avons proposé une deuxième version du générateur. Elle se base sur la distinction des durées inférieures ou supérieures à 5 mn aussi bien pour la simulation du support que pour la relation durée/intensité des événements. Ce deuxième générateur permet de valider toutes les propriétés statistiques sur toute la gamme d'échelle au prix toutefois d'un nombre plus élevé de paramètres.

Le générateur 2 permet d'obtenir un meilleur agrément entre les mesures et les simulations sur plusieurs critères et à différentes résolutions temporelles. Au total ce second générateur nécessite 25 paramètres, 14 pour le support, 3 pour la variabilité intra-événement et 8 pour le processus de re-normalisation. Cependant, certains paramètres sont fixés et d'autres ne devraient pas changer significativement d'un point à l'autre du globe. En effet, si l'on suppose que les processus en jeu aux courtes durées sont liés aux lois de la physique qui gouvernent la variabilité des précipitations à très petite échelle (inférieure 5 minutes) on peut supposer que les paramètres correspondant sont constants et donc universels. Au final, seuls les paramètres à plus large échelle nécessitent un ajustement d'une région à une autre sont ceux qui caractérisent les échelles supérieures à 5 minutes: 8 paramètres pour le support, 2 pour la re-normalisation. Le générateur 2 permet d'obtenir des performances nettement supérieures pour un nombre de paramètres à ajuster identique au générateur 1.

Plusieurs aspects restent à étudier pour définir les limites de ce simulateur. Les travaux réalisés ont permis de montrer qu'il reproduisait les propriétés observées à Palaiseau sur des séries de 2 ans à la résolution de 15s. L'utilisation de séries chronologiques longues pourrait permettre d'analyser et de prendre en compte l'effet saisonnier. L'utilisation de séries chronologiques observées à fine résolution en d'autres endroits du globe est indispensable pour confirmer le caractère général des hypothèses retenues.

# Chapitre 4

## GENERATION DE CARTES 2D DE PRECIPITATIONS

### Sommaire

---

4	GENERATION DE CARTES 2D DE PRECIPITATIONS .....	105
4.1	Introduction.....	105
4.2	Les données .....	107
4.3	Méthodologie.....	109
4.4	Analyse et modélisation du support de pluie .....	110
4.4.1	Quelques approches possibles pour la simulation du support .....	110
4.4.2	Méthode retenue (algorithme de simulation séquentielle) :Etape 1 .....	110
4.4.2.1	Principe de l'analyse variographique .....	112
4.4.3	Analyse et modélisation des variogrammes empiriques des supports de pluie observés .....	116
4.4.3.1	Choix du modèle paramétrique .....	118
4.4.3.2	Estimation des paramètres des variogrammes observés .....	119
4.4.4	Exemple de supports simulés .....	122
4.5	Simulation des lames d'eau dans les zones de pluie .....	123
4.5.1	Simulation des zones de pluie: étape 2.....	123
4.5.2	Re-normalisation des lames d'eau simulées: Etape 3.....	125
4.5.2.1	Relation entre la surface et la lame d'eau moyenne des évènements .....	125
4.5.3	Traitement des transitions pluie/non pluie .....	130
4.6	Validation des cartes simulées .....	134
4.6.1	Propriétés multifractales des simulations .....	134
4.6.2	Propriétés spectrales.....	135
4.6.2.1	Analyse des propriétés spectrales des observations.....	135
4.6.2.2	Analyse des propriétés spectrales des simulations obtenues avec support conditionnel (réel).....	139
4.6.2.3	Analyse des propriétés spectrales des simulations obtenues avec support inconditionnel (synthétique) .....	141
4.6.3	Distributions statistiques .....	142
4.6.3.1	Analyse statistiques des observations.....	142
4.6.3.2	Analyse statistique des simulations avec support conditionnel (réel).....	143
4.6.3.3	Analyse statistiques des simulations avec support inconditionnel (synthétique).....	144
4.7	Conclusion .....	145

---





## 4 GENERATION DE CARTES 2D DE PRECIPITATIONS

### 4.1 Introduction

La variabilité des précipitations concerne aussi bien l'espace que le temps, la variabilité de l'une est associée à la variabilité de l'autre via les relations liées à la dynamique météorologique, aux turbulences et aux effets du relief.

La variabilité spatiale des précipitations est considérée comme une importante source d'erreur pour les modèles hydrologiques, qui simulent les débits des rivières à partir des précipitations sur les bassins versants. Elle complique également la tâche des prévisionnistes ; les prévisions météorologiques étant basées sur des modélisations numériques effectuées à des résolutions spatiales au mieux de quelques kms.

C'est dans ce contexte que nous avons développé un modèle de simulation 2D, qui permet de simuler des champs précipitants à partir des analyses faites sur des mesures Radar à la résolution  $1 \times 1 \text{ km}^2$ .

Les modèles et les techniques pour la simulation des événements pluvieux synthétiques ont une longue histoire et couvrent une large gamme d'applications. Beaucoup de modèles de génération 2D existent dans la littérature. Les méthodes basées sur les processus de Poisson par exemple sont très utilisés pour simuler des séries chronologiques de précipitations (e.g., Onof et al., 2000; Burton et al., 2008; Evin et Favre, 2013). En revanche, leur implémentation devient rapidement complexe dans les applications en deux ou trois dimensions. D'autres types de modèles font appel aux techniques de la géostatistique pour simuler des cartes de précipitation (e.g. Schleiss et al., 2009, 2014; Leblois et Creutin, 2013). Ces techniques de simulation visent essentiellement à reproduire les structures spatiales (basées typiquement sur des variogrammes). Ces techniques n'exploitent pas toutes le potentiel de la géostatistique et peuvent encore être améliorées, en particulier par rapport à l'intermittence et d'autres sources de non-stationnarité (la variabilité, les extrêmes ...). (Schleiss et al. 2014) ont proposé un modèle basé sur la méthode dite « dry/drift » (qui fait varier le taux de pluie en un point en tenant compte de la distance qui le sépare des zones sèches l'entourant). Le modèle semble être réaliste bien qu'aucun critère de validation statistique en dehors de la structure spatiale n'ait été donné. Cependant en testant ce modèle sur nos

données, nous avons constaté que la distribution des taux de pluie ne correspond pas à celle observée. (Leblois et al. 2013) ont proposé un modèle géostatistique spatio-temporel, la simulation du support se fait en utilisant un modèle gaussien avec un seuil. D'autres approches basées sur les fractales (multi) et des considérations d'auto-similarité semblent plus prometteuses et plus populaire (e.g., Lovejoy et al., 1987; Gupta et Waymire, 1993; Menabde et al., 1997; Deidda, 2000; Pathirana et Herath, 2002). Comme nous l'avons précisé dans le chapitre précédent concernant la simulation 1D, les multifractales sont intéressantes car elles offrent un cadre très efficace pour résumer la variabilité spatiale et/ou temporelle des champs de précipitations à différentes échelles spatiales et/ou temporelles. Les hypothèses mathématiques (c.-à-d. stationnarité des incréments) derrière l'approche peut cependant être remise en question, en particulier en présence d'intermittence (e.g. Verrier et al., 2011; Gires et al., 2013, Serinaldi 2010, Véniziano 2012).

Notre simulateur, qui est une extension du modèle 1D de simulation de séries temporelles de précipitation présenté au chapitre précédant, se base sur la même méthodologie que ce dernier (simulation du support/ simulation multi fractale FIF des zones de pluie/ re-normalisation ).

Concernant la génération du support de pluie, nous avons mis en œuvre deux approches :

- **Simulation conditionnelle** : On utilise des supports de pluie observés via des observations radars. On dispose donc de la forme des cellules de pluie et seul les taux précipitants à l'intérieur de chaque cellule sont modélisés via une modélisation multifractale FIF.
- **Simulation inconditionnelle** : Les supports de pluie sont générés à partir d'une méthode géostatistique. Les taux précipitants sont alors générés de la même manière que précédemment.

De la même façon qu'en 1D, les cartes simulées avec le modèle FIF nécessitent une étape complémentaire de re-normalisation dont l'objectif est d'avoir des taux de pluie (ou plus exactement ici des lames d'eau) réalistes. D'après les résultats du chapitre précédent, une re-normalisation par évènement contrairement à une re-normalisation globale de la CDF, permet une meilleure restitution des propriétés spectrales et multifractales aux échelles fines. Cette étape de

re-normalisation est donc basée sur l'analyse de la relation longueur (ou surface) de la cellule de pluie / Intensité moyenne de la cellule observée.

## 4.2 Les données

Dans cette partie, nous présentons les données de Météo France utilisées pour faire notre analyse. Il s'agit des cartes mosaïques de lame d'eau radar : estimation en mm du cumul de précipitations durant 5 minutes sur l'ensemble du territoire français obtenu par composition des lames d'eau individuelles des radars du réseau de Météo-France (24 radars)(Figure 4.1). Les lames d'eau sont obtenues par combinaison des lames d'eau des différents radars et accompagnées en chaque pixel d'un code qualité.

Dans ce chapitre, j'ai utilisé une année de données mosaïques (1 janvier – 31 décembre 2012) couvrant toute la France soit environ 103633 mosaïques de lames d'eau radar. Chaque carte a une résolution spatiale de 1 km<sup>2</sup> et une résolution temporelle de 5 minutes ainsi qu'une surface de 1536x1536 km<sup>2</sup>.

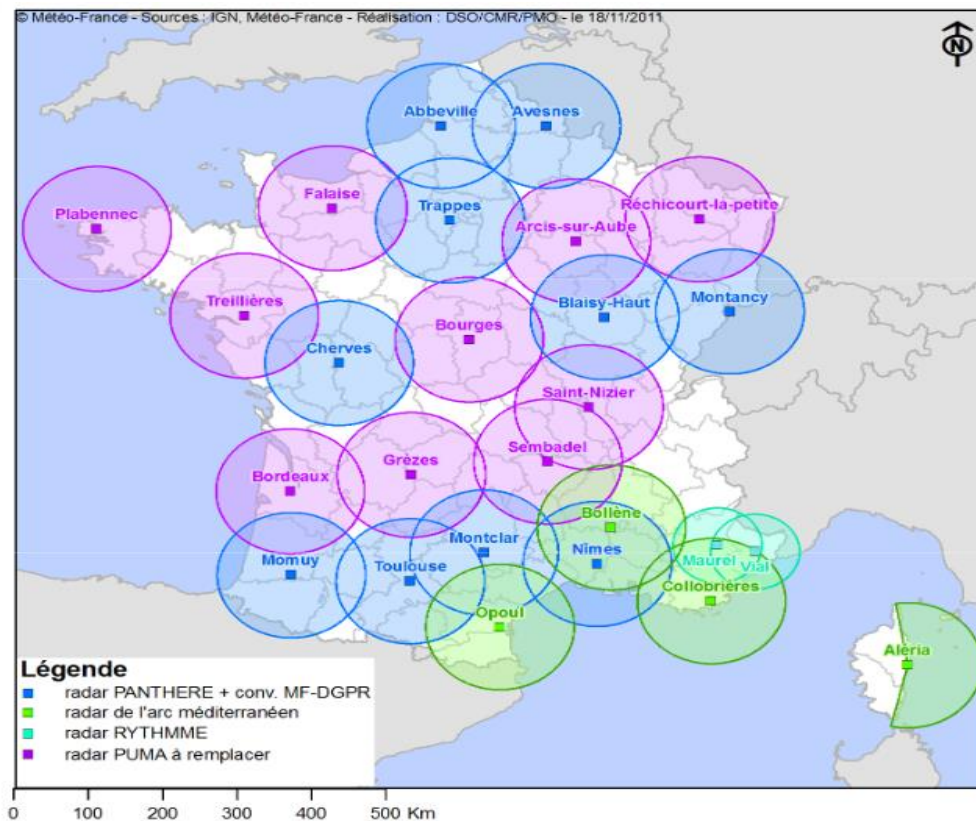


Figure 4.1. Réseau Radar Météo France.

Les cartes mosaïques permettent de renseigner sur les différents climats observés en France (6 climats très différents). Notre objectif étant, à ce stade, essentiellement méthodologique nous nous sommes focalisés sur une zone d'étude particulière. Le simulateur 1D présenté précédemment étant développé à partir des données observées au SIRTA en région parisienne (48.7145 Nord, 2.24575 Est) nous avons fait également le choix de développer le simulateur 2D pour cette même région. L'Ile-de-France se caractérise par la quasi absence de reliefs et par un climat océanique altéré (zone de transition entre le climat océanique et les climats de montagne). Un climat caractérisé par une certaine modération, que ce soit en termes de précipitations, de températures ou encore de vent. Une région aussi connue comme relativement sèche. On compare son taux de pluie moyen annuel (600mm) au taux de pluie moyen annuel de toute la France (environ 720 mm). En revanche, le nombre de jours où il pleut est très important (environ 170 jours) et se trouve au-dessus de la moyenne nationale. Il pleut donc souvent mais les pluies sont rarement intenses.

Le site du SIRTA étant proche du radar météorologique de Trappes, nous avons extrait des cartes de 130x130 km<sup>2</sup> centrées sur ce radar. La distance maximum au radar est de l'ordre de 90 kilomètres garantissant ainsi des données de bonne qualité car relativement proches du radar. En premier lieu, un prétraitement a été effectué sur ces données. En effet, après discussion avec des agents de Météo France il ressort que les deux premières classes de lames d'eau (0.12 et 0.24 mm) sont observées avec des barres d'erreur extrêmement importantes et notamment un taux important de fausses détections. Afin d'éviter tout problème, seules les valeurs appartenant à la troisième classe (0,48 mm) et au-delà ont été conservées.

Nous avons donc un total de 103200 cartes dont 60607 cartes de zéros (pas de pluie), ce qui donne un pourcentage de 41.52% de cartes de pluie sur l'ensemble de l'année 2012.

Dans les paragraphes suivants, ces données sont utilisées d'une part pour modéliser les propriétés du support et d'autre part pour déterminer une relation entre les lames d'eau moyennes des événements et leurs surfaces. Elles sont également utilisées pour étudier la capacité de notre modèle à restituer un certain nombre de propriétés statistiques des lames d'eau radar telles que les propriétés spectrales, multifractales et leurs distributions.

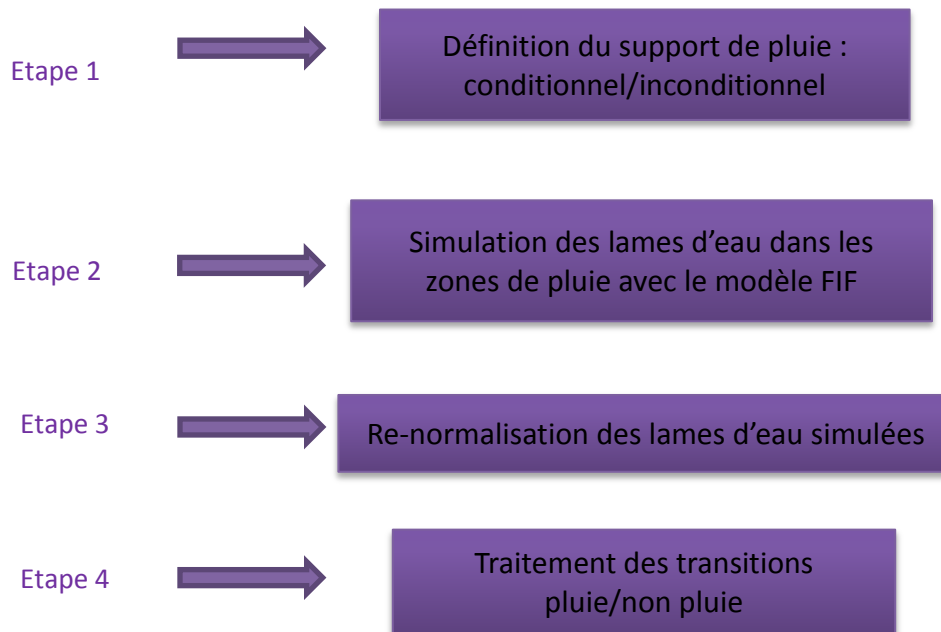
### 4.3 Méthodologie

Une carte de précipitation peut être vue comme un champ de variables aléatoires dans  $\mathbb{R}^2$ . En séparant le processus de lame d'eau  $R(mm)$  du processus d'occurrence, nous obtenons un masque représentatif du support de pluie  $I$  défini comme suit :

$$I(x) = \begin{cases} 0 & \text{si } R(x) = 0 \\ 1 & \text{si } R(x) > 0 \end{cases}$$

Où  $R(x)$  représentent la valeur de lame d'eau au point  $x \in \mathbb{R}^2$

La première étape de notre générateur de pluie est de définir le processus d'occurrence de pluie permettant de disposer d'un support de pluie  $I(x)$ . A partir de ce support de pluie, la deuxième étape est de simuler des évènements de pluie dans les zones pluvieuses, c'est-à-dire lorsque  $I(x) = 1$ . La troisième étape est de re-normaliser les valeurs générées pour se rapprocher de valeurs réalistes et avoir des distributions cohérentes avec les mesures. Comme nous le verrons plus loin, bien que le générateur simule des cartes globalement réalistes, il est nécessaire néanmoins de traiter les transitions pluie/non pluie. En effet, comme nous simulons la pluie indépendamment du support, les transitions peuvent dans certain cas être trop brutales. Afin de traiter ce problème une étape finale consistant en un pseudo filtrage sera appliquée. Le diagramme suivant résume les étapes que nous détaillons dans la suite.



## **4.4 Analyse et modélisation du support de pluie**

### ***4.4.1 Quelques approches possibles pour la simulation du support***

Le support de pluie a été relativement peu étudié dans la littérature. En effet des formes très simples sont utilisés pour modéliser les cellules de pluie comme les cercles (e.g. Morin et al. 2005, 2006) ou encore des ellipses, comme c'est le cas dans (Northop 1998 ; Von Hardenberg et al. 2003 ; Reborra et Ferraris 2006). D'autres ont proposé d'utiliser les modèle bêta ou en intégrant le composant bêta (une simple cascade on/off) ( Schertzer et al. 2002). D'autres chercheurs ont utilisé la technique de seuillage pour pouvoir simuler des zéros dans leurs modèles, ainsi (De Montera et al. 2010, Verrier et al. 2011) ont appliqué un seuil sur des cartes générées avec le modèle multifractal FIF. Leblois et al. (2013) ont appliqué un seuil sur un champ généré avec un processus gaussien. Gires et al. (2014) ont proposé une version de modèle multifractal permettant de générer des zéros. Schleiss et al. (2014) ont simulé le support de pluie indépendamment des taux de pluie en utilisant une méthode de simulation séquentielle. Comme nous l'avons vu au chapitre 3, la simulation du support de pluie est primordial, nous devons donc le reproduire fidèlement. C'est donc ce type de méthode que nous avons retenue dans notre modèle. Elle nous permettra ainsi de décorrélérer la simulation du support et des taux de pluie.

### ***4.4.2 Méthode retenue (algorithme de simulation séquentielle) :Etape 1***

L'objectif de cette section est de montrer comment l'occurrence pluie/ non pluie peut être modélisée et quelles sont les informations nécessaires qui doivent être déduites des observations. Nous avons choisi l'algorithme de simulation séquentielle (Gomèz-Hernàndez, 1993) dont la version « Sequential indicator simulation » notée SIS est utilisée par Schleiss dans son générateur (Schleiss et al ; 2014). Ce choix est basée sur la méthode la plus réaliste après avoir implémenté les deux modèles de (Schleiss et al., 2014) et (Gires et al., 2014 ).

C'est une approche qui comme son nom l'indique simule le support point par point d'une façon séquentielle. Elle nécessite deux informations : la probabilité d'occurrence de pluie et la structure spatiale du champ à simuler.

La probabilité d'occurrence de pluie sur une carte ( $\mu$ ) est estimée sur les observations suivant la formule :

$$\mu = \frac{1}{N} \sum_{i=1}^N I(x_i) \quad \text{Équation 4.1}$$

Où  $I(x_i)$  est le processus binaire d'occurrence de pluie correspondant à  $R(x_i)$ .  $N$  est le nombre de pixels utilisés pour estimer l'occurrence de pluie (nombre de pixels total de la carte).

Quant à la structure spatiale du support de pluie, elle est caractérisée par le (semi-) variogramme empirique défini ci-après.

La simulation séquentielle consiste, comme signalé précédemment, à simuler le support point par point. Ainsi, cet algorithme nous permet, connaissant les réalisations en  $M$  points précédents, de générer  $N$  nouvelles réalisations de  $I(x)$  aux points  $x_{M+1}, x_{M+2}, \dots, x_{M+N}$  par une visite séquentielle et unique de chaque point. Notons que l'algorithme est indépendant de l'ordre de visite des points (Gomèz-Hernández, 1993). La seule condition pour pouvoir l'implémenter est de pouvoir calculer la loi de distribution de probabilité conditionnée par les valeurs des données originales et des données simulées auparavant. Cela revient à dire que pour chaque point visité  $x_i$  ( $i = M + 1$  à  $M + N$ ), nous faisons un tirage aléatoire d'une réalisation de la distribution conditionnelle de  $I(x_i)$  connaissant toutes les valeurs données ou/et simulées précédentes  $I(x_1), I(x_2), \dots, I(x_M)$ . Ainsi pour simuler le point  $x_{i+1}$ , la simulation est conditionnée par les valeurs  $I(x_1), I(x_2), \dots, I(x_M), \dots, I(x_i)$ .

Notons que la génération du support revient à faire des tirages de variables aléatoires de Bernoulli pour chaque point. Nous devons donc déterminer la probabilité que  $I(x_i)$  soit égal à 1 ou 0 en tenant compte de toutes les réalisations précédentes. En général, la valeur exacte de cette probabilité est difficile à déterminer, mais elle peut être estimée par la méthode de krigeage simple, ainsi la probabilité  $P(x_i)$  au point  $x_i$  :

$$P(x_i) = \mu + \sum_{j < i} w_j (I(x_j) - \mu) \quad \text{Équation 4.2}$$

Avec  $[w_1, \dots, w_{i-1}]$  le vecteur des coefficients obtenus en résolvant le système de krigeage suivant :

$$w = C^{-1} C_i$$

$C$  est une matrice de dimension  $(i - 1) \times (i - 1)$  et  $C_i$  est un vecteur de dimension  $(i - 1) \times 1$ .

S'il n'y a pas d'observations antérieures (c.-à-d. si  $i = 1$ ), le système de krigeage n'a pas besoin d'être résolu et  $P(x_i) = \mu$ .



Une fois la valeur de la probabilité  $P_i$  déterminée, le problème est réduit à un tirage aléatoire d'une valeur à partir d'une loi de Bernoulli de moyenne connue. La valeur estimée  $P_i$  peut en effet être inférieure à 0 ou supérieure à 1. Dans ce cas, elle est mise soit à 0 ou à 1. Pour plus de détails sur la méthode de calcul le lecteur peut se référer à (Albert, 1987; Gomèz-Hernàndez, 1993).

#### 4.4.2.1 Principe de l'analyse variographique

Le variogramme/semi-variogramme est une méthode mathématique utilisée en géostatistique pour résumer la structure et le degré de dépendance spatiale d'un champ aléatoire ou un processus stochastique  $Z$ . Il est notamment utilisé pour les méthodes de krigeage. L'analyse variographique ou l'analyse structurale est donc l'estimation et l'étude du variogramme d'un échantillon donné.

Un variogramme est défini comme la 1/2 variance de la différence entre deux réalisations du processus à des localisations  $(x, y)$ :

$$2\gamma(x, y) = Var(Z(x) - Z(y)) \quad \text{Équation 4.3}$$

Le champ d'occurrence des précipitations que nous avons nommé support de pluie est l'une des caractéristiques principale du champ de pluie. Il caractérise à la fois la fréquence et la position des régions pluvieuses et sèches dans la région étudiée. Dans le cadre de cette analyse variographique, nous avons étudié les variogrammes de la fonction  $I$  qui caractérise le support extrait des cartes d'observation de Palaiseau. Nous avons ainsi supposé que la fonction  $I$  est stationnaire au second ordre. Une hypothèse forte mais en cohérence avec la région d'étude qui est une région à peu près homogène qui est caractérisée, comme annoncé précédemment, par la quasi absence de relief. Notons qu'un variogramme peut être isotrope ou anisotrope. Selon (Tobin et al. 2011), un variogramme anisotrope reproduit correctement la structure spatiale des précipitations quel que soit la méthode géostatistique utilisée, en revanche, dans le modèle isotrope, quelques orientations peuvent ne pas être reproduites correctement. Dans le cas de notre

étude, une avons posé une deuxième hypothèse qui est l'isotropie à petite échelle, hypothèse classiquement posée dans certains autres simulateurs (e.g. Schleiss et al. 2014).

Le variogramme dépend ainsi seulement de la distance  $h$  entre deux points et non de leurs positions et/ou de la direction.

En termes d'équation statistique, le variogramme de l'indicateur correspondant au support  $I$  a donc la forme suivante :

$$2 \gamma_I(h) = \text{var}[I(x+h) - I(x)] = \mathbb{E}[(I(x+h) - I(x))^2] \quad \text{Équation 4.4}$$

Un variogramme, comme le montre la Figure 4.2 est caractérisé par trois paramètres qui sont :

La pépité : c'est la limite du variogramme quand la distance  $h$  tend vers zéro. Elle présente la variation des mesures à des localisations très proches. L'effet de pépité peut ne pas exister (le variogramme est continu à l'origine et le paramètre pépité est nul). Son existence peut être due en partie aux erreurs des instruments de mesure, ou encore à la variabilité réelle des mesures à des endroits très proches.

Le palier : valeur du variogramme qui présente un plateau. Dans ce cas, la covariance des données tend vers 0 à l'infini et  $\text{Palier}(\infty) = \text{var}[I]$ .

La portée : est la distance à partir de laquelle le variogramme atteint son palier. Autrement dit, la portée est la distance à partir de laquelle il n'y a plus de dépendance spatiale entre les mesures. En pratique elle est estimée par la distance à partir de laquelle 95% du palier est atteint.

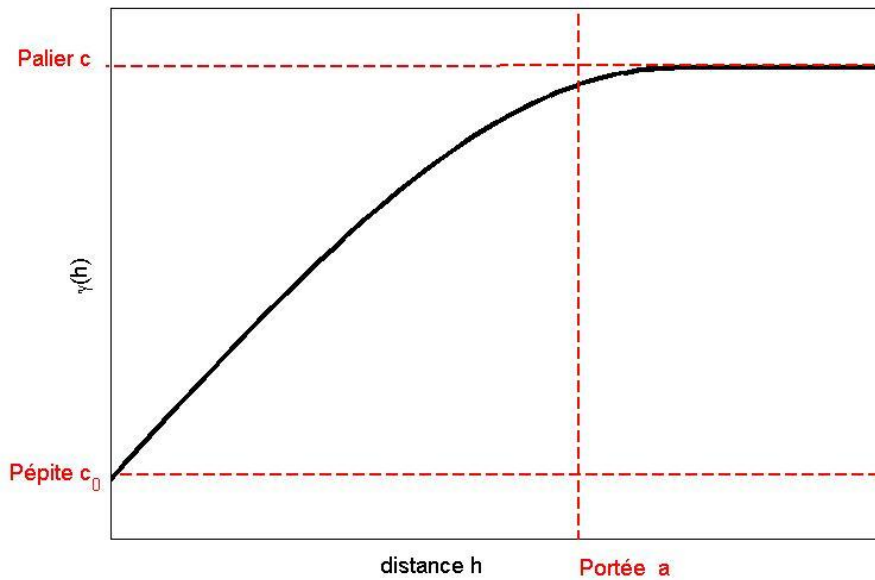


Figure 4.2 Exemple de variogramme empirique

Le variogramme empirique dans le cas isotrope est calculé de la façon suivante:

$$\gamma_I(h) = \frac{1}{2N_h} \sum_{x_i - x_j \sim h} [I(x_i) - I(x_j)]^2 \quad \text{Équation 4.5}$$

Avec  $N_h$  le nombre de pixel  $x_j$  dont la distance par rapport à  $x_i$  est égale à  $h$

Notons que la fonction  $I(x)$  est binaire et doit être stationnaire au second ordre (hypothèse faite), nous avons  $I(x) \in \{0,1\}$  de moyenne  $\mu$  et de fonction de covariance :

$$C(h) = Cov[C(x+h), C(x)] = C(0) - \gamma(h) \quad \text{Équation 4.6}$$

Avec  $C(0)$  la variance de  $I(x)$  et  $\gamma(h)$  son variogramme.

En pratique, les variogrammes empiriques sont modélisés par des modèles standards (modèles paramétriques). Le choix du variogramme modèle dépend des mesures. Dans le cas des précipitations les modèles utilisés sont les modèles sphériques, exponentiels ou linéaire (Chiles et Delfiner, 1999 ; Legorgeu, 2013). Plusieurs modèles peuvent aussi être combinés pour créer une structure adéquate aux mesures. Le Tableau 4.1 montre des exemples de modèles standards classiquement utilisés. Les modèles sphérique, exponentiel, linéaire avec palier, gaussien et

puissance présentés dans le tableau n'ont pas d'effet de pépite. Si besoin, il suffit de combiner le modèle en question avec l'effet de pépite pour avoir une discontinuité en zéro.

Comme nous le remarquons, tous les modèles sauf le modèle puissance possèdent un palier. Ils sont définis par trois paramètres : l'effet de pépite  $c_0$  qui peut ou ne pas exister, le paramètre  $c$  qui correspond à la valeur du palier et le paramètre  $a$  dit paramètre de portée représentatif de la longueur de la zone de transition.

Le modèle	$\gamma(h)$	distance $h$	Exemple de figure :
Effet de pépite	$0$ $c_0$	$h = 0$ $h > 0$	<p style="text-align: center;"><math>c_0 = 1</math></p>
Sphérique	$0$ $c \left( \frac{3h}{2a} - \frac{1}{2} \left( \frac{h}{a} \right)^3 \right)$ $c$	$h = 0$ $0 < h < a$ $h \geq a$	<p style="text-align: center;"><math>c = 1, a = 1</math></p>
Exponentiel	$c \left( 1 - \exp\left(\frac{-h}{a}\right) \right)$	$h \geq 0$	<p style="text-align: center;"><math>c = 1, a = 1</math></p>
Linéaire avec palier	$c \left( \frac{h}{a} \right)$ $c$	$0 \leq h < a$ $h \geq a$	<p style="text-align: center;"><math>c = 1, a = 1</math></p>

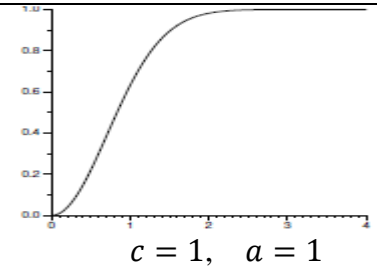
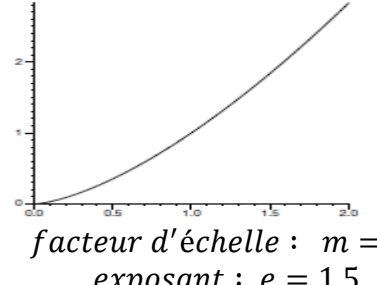
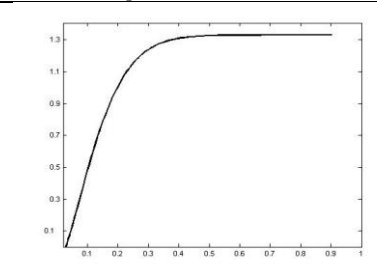
Gaussien	$c \left( 1 - \exp \left( - \left( \frac{h}{a} \right)^2 \right) \right)$	$h \geq 0$	 $c = 1, a = 1$
Puissance (pas de palier)	$m h^e$	$h \geq 0$ $0 < e \leq 2$	 <i>facteur d'échelle : m = 1</i> <i>exposant : e = 1.5</i>
Stable	$c \left( 1 - \exp \left( - \sqrt{\frac{h}{a}} \right) \right)$	$h \geq 0$	

Tableau 4.1 Quelques exemple de modèles de variogramme (cf. Chiles et Delfiner 1999)

### 4.4.3 Analyse et modélisation des variogrammes empiriques des supports de pluie observés

Les variogrammes empiriques comme le montre la Figure 4.3 sont très différents d'une carte à l'autre. En effet, un support de pluie avec une probabilité d'occurrence de pluie faible donne un variogramme avec une portée assez faible qui est de l'ordre de 10 Km dans l'exemple 1 de la Figure 4.3 ; un support avec une probabilité d'occurrence de pluie importante correspond à une portée plus importante (Exemple 2 de la Figure 4.3). Pour une probabilité d'occurrence donnée, des variogrammes de forme très différentes peuvent être également observés selon le type de la cellule de pluie. Ainsi sur les exemples 3 et 4 de la Figure 4.3 qui correspondent respectivement à des probabilités d'occurrence assez similaires (0.54 et 0.49), on observe des portées de 30 et 70 kilomètres et des plateaux à 0.11 et 0.3. Une grande cellule de pluie qui couvre toute la zone

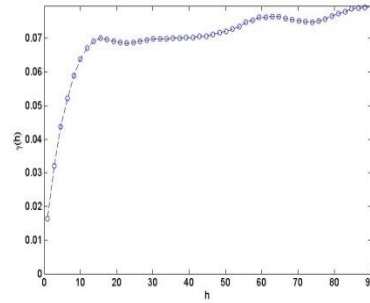
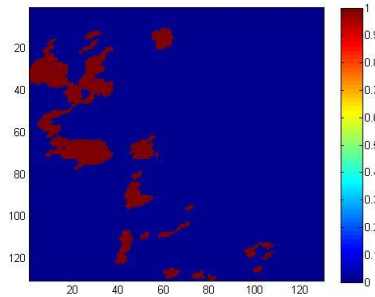
pluvieuse donne une portée assez importante, tandis que plusieurs petites cellules de pluie éparpillées sur toute la carte correspondent à une portée plus faible.

La portée (ou range en anglais) est donc une fonction de la dimension et de l'espacement des zones de pluvieuses. Le pallier (ou sill en anglais) correspond au rapport des pourcentages de pluie et de non pluie. La pépite (ou nugget en anglais) et la pente à l'origine sont, elles, fonctions de la complexité des limites des zones pluvieuses.

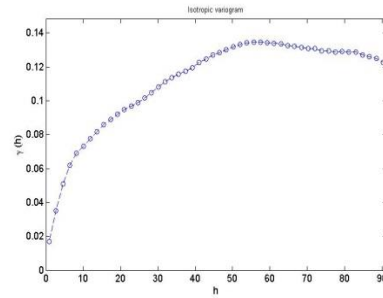
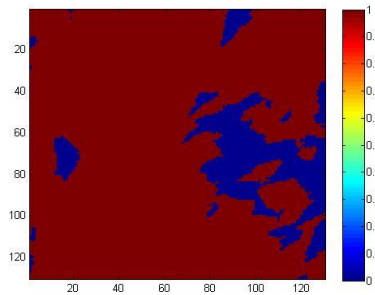
Support Réels

Variogramme de l'indicatrice

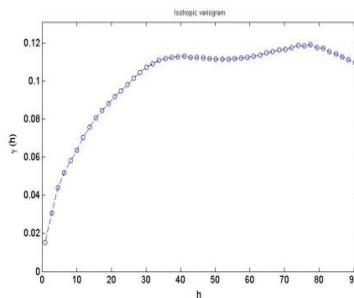
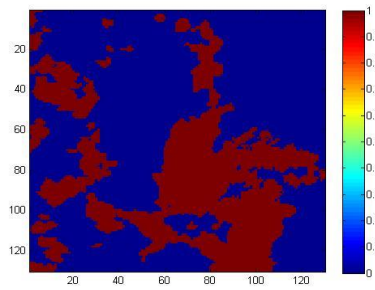
Exemple 1

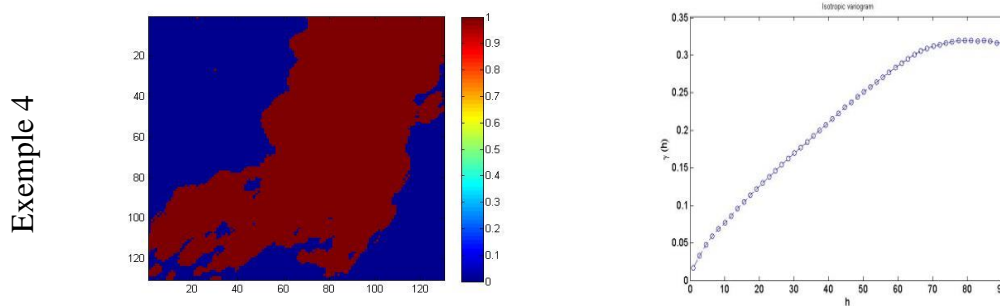


Exemple 2



Exemple 3



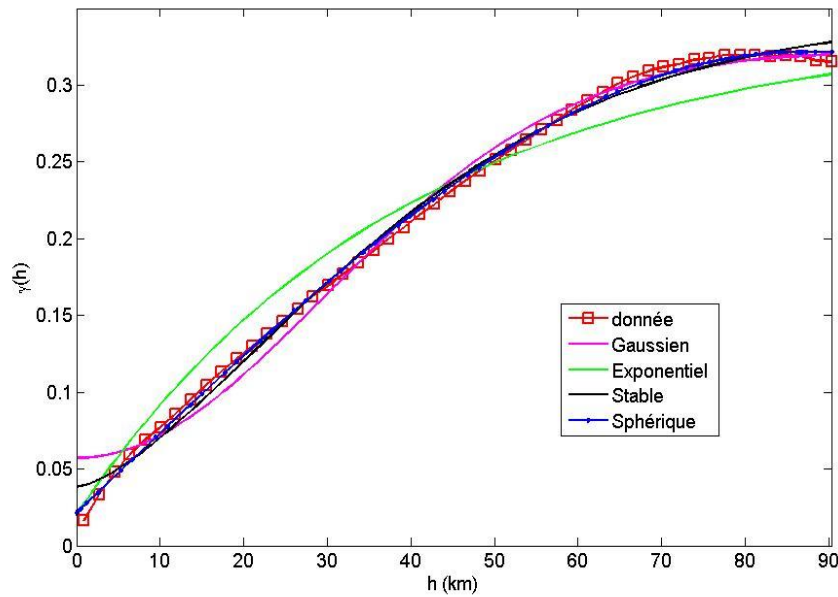


*Figure 4.3 Exemples de supports issus d’images radar, et des variogrammes empiriques correspondant à différents pourcentages de valeurs non nulles (exemple 1) 15%, (exemple 2) 86%, (exemple 3) 49%, (exemple 4) 54%*

#### 4.4.3.1 Choix du modèle paramétrique

Le choix d’un modèle paramétrique ajustant au mieux les mesures n’est pas évident. Il est à noter qu’il existe plusieurs approches pour ajuster un variogramme empirique, la plus utilisée est la méthode des moindres carrés visant à minimiser l’écart entre le variogramme empirique et le modèle paramétrique. Bien que les approches existantes donnent dans la plupart des cas une estimation, elles ne sont malheureusement pas toujours optimales. Il n’existe pas à notre connaissance de méthodes universelles qui justifie le choix optimal du modèle de variogramme. Dans la littérature, le modèle sphérique est très utilisé pour ajuster les variogrammes des structures spatiales des cartes des précipitations (Berne et al., 2004 ; Emmanuel et al., 2011 ; Schleiss et al., 2014 ; Legorgeu, 2013) ou du support de pluie (Schleiss et al. 2014). Le modèle exponentiel est aussi utilisé par certains auteurs (e.g. Legorgeu., 2013).

Parmi les modèles présentés dans le tableau 4.1, nous avons testé quatre modèles (les modèles : sphérique, exponentiel, gaussien et stable). Nous avons écarté le modèle puissance car il ne possède pas de palier (et qu’il diverge asymptotiquement) alors que tous les variogrammes empiriques calculés sur nos données en possèdent un. Après analyse de plusieurs cas de variogrammes différents, nous avons constaté que le modèle sphérique est celui qui ajuste au mieux les données. Un exemple est montré dans la Figure 4.4. Nous remarquons que les modèles gaussien et exponentiel ne peuvent pas être utilisés. Le modèle stable qui ajuste bien le variogramme pour les distances moyennes a un problème à la fois au niveau de la pépite et du palier.



*Figure 4.4 Exemples de variogrammes empirique et de modèles paramétriques ajustés (gaussien, exponentiel, stable et sphérique).*

#### 4.4.3.2 Estimation des paramètres des variogrammes observés

Dans l'objectif de simuler des supports de pluie synthétiques avec un pourcentage de pluie raisonnable, nous avons sélectionné dans un premier temps des cartes de précipitations observées en hiver entre janvier et avril 2012 ayant un pourcentage de pluie supérieur à 15%. Nous avons donc analysé au total 1277 cartes qui ne sont pas forcément successives dans le temps. Nous avons ensuite estimé sur ces données les paramètres du modèle de variogramme sphérique ajustant au mieux les variogrammes empiriques. Les paramètres estimés sont présentés à la Figure 4.5. Nous remarquons une forte variabilité des paramètres d'une carte à l'autre. L'effet de pépite est très faible, inférieur à 0.1 pour la plupart des cartes, avec une moyenne égale à 0.03.

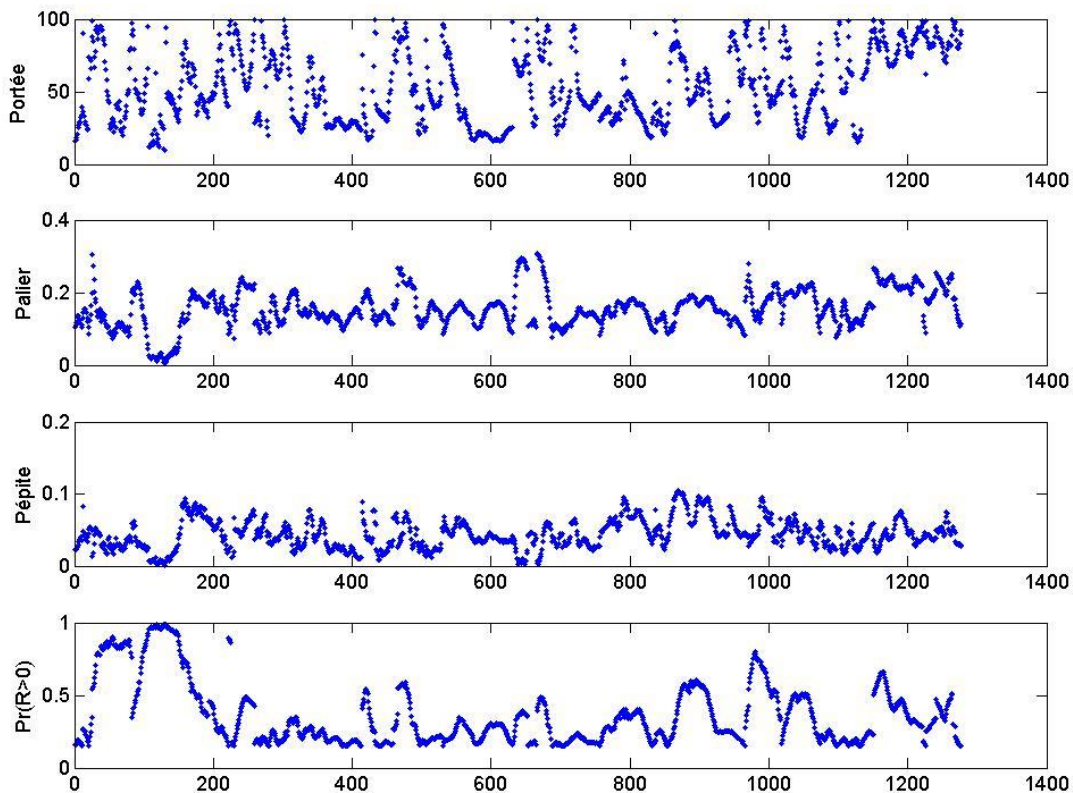
Pour les valeurs des paliers, 90% de valeurs sont comprises entre 0.1 et 0.3 et la moyenne des valeurs est égale à 0.17. Concernant le paramètre de portée, il varie fortement entre 10 et 100 km, avec une moyenne de 52 km.

La plupart des études publiées concernent des analyses variographiques de cartes de précipitations et non du support comme dans notre cas. Bien que le variogramme de la pluie et le variogramme de l'indicatrice (i.e. du support de pluie) soient différents. Ils nous renseignent tous les deux sur le processus de pluie. Dans une étude récente menée à partir de cartes observées à



partir du radar de Treillieres avec une résolution spatiale de 250 mètres et une résolution temporelle de 5 minutes Emmanuel et al. (2011) ont étudié les variogrammes des taux précipitants. Les cartes ont été classées en quatre groupes qui correspondent à des situations météorologiques différentes (pluie fine, orages organisés, orages moins organisés et averses). Les auteurs ont conclu qu'il n'y avait pas d'effet de pépite et que la portée dépend du type de pluie (5 km pour des pluies intenses et fortement variable, 15 km pour des pluies moins intenses et moins variables).

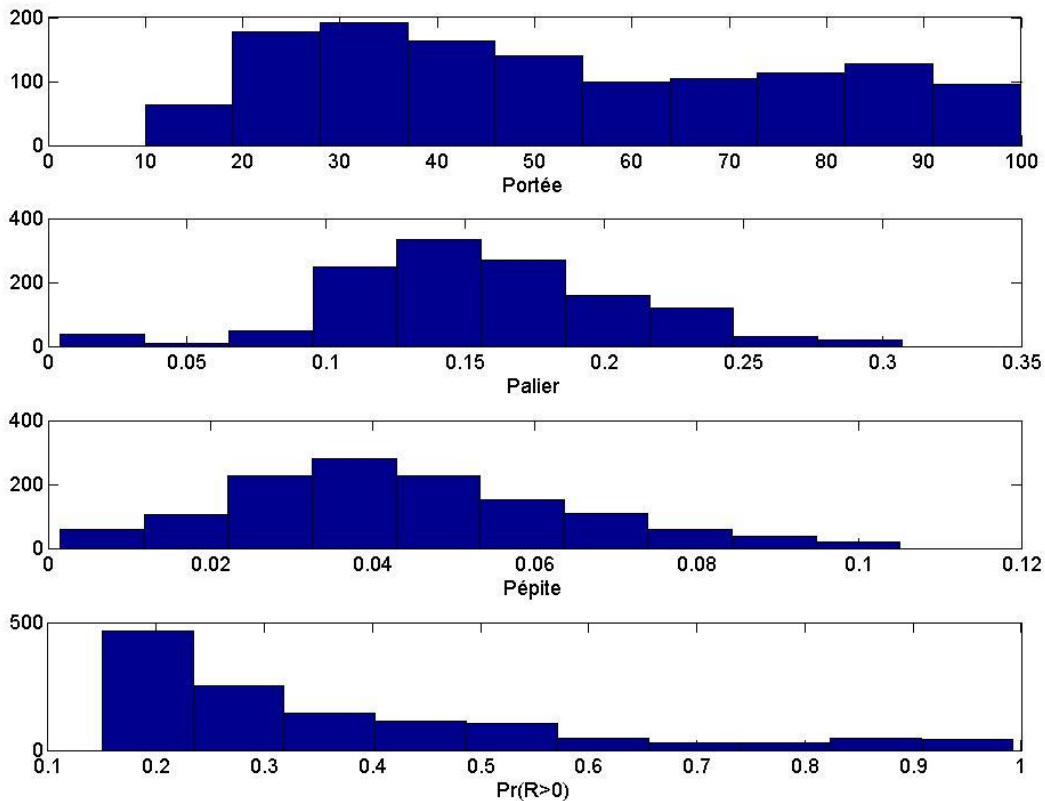
Dans l'article publié en 2014 relatif à la méthode présentée dans la section 4.4.2, Schleiss et al. (2014) utilisent comme paramètres de variogramme ceux estimés sur l'évènement convectif du 1er aout 2012 à Genève en Suisse. Concernant le variogramme du support de pluie, ils ont respectivement obtenu pour les trois paramètres pépite, portée et palier les valeurs 0.01, 40 et 0.3. Ils ont donc fixé ces trois paramètres dans leur modèle.



*Figure 4.5. Valeurs des paramètres du modèle de variogramme sphérique et probabilité d'occurrence de pluie en fonction du numéro de la carte sur une sélection de 1277 cartes (Hiver 2012 ayant un pourcentage de pluie supérieur à 15%)*

La Figure 4.6 représente les histogrammes des trois paramètres. La distribution de la portée est proche de la distribution uniforme entre 10 et 100 km. Il est à noter qu'il n'y a pas de valeurs inférieures à 10 km. Nous considérerons donc dans la suite de l'étude que ce paramètre est distribué uniformément dans l'intervalle [10, 100] kilomètres. Concernant le palier celui-ci possède une distribution en cloche légèrement dissymétrique, 95% des valeurs sont comprises entre 0.1 et 0.2. Nous avons donc choisi de prendre ce paramètre constant et égal à sa valeur moyenne (0.17). Concernant l'effet de pépite, nous remarquons que ce paramètre est distribué aux alentours de la valeur 0.04, valeur jugée suffisamment faible pour être fixée à sa valeur théorique, c'est à dire zéro.

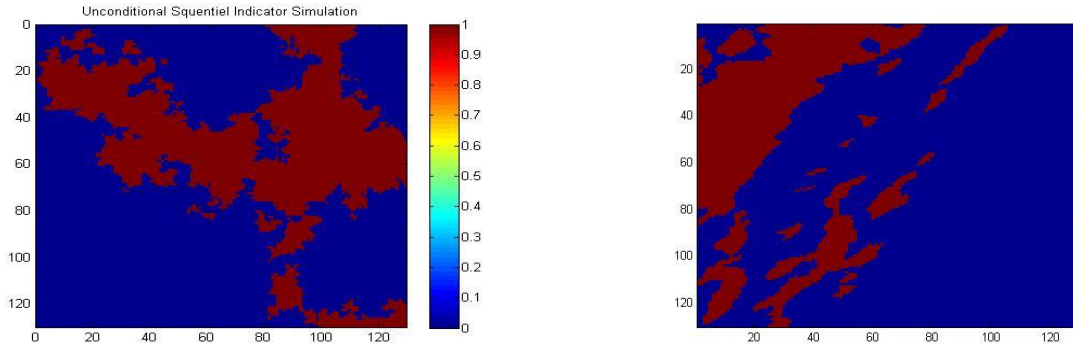
La probabilité d'occurrence/pourcentage de pluie est le quatrième paramètre nécessaire pour simuler le support de pluie. L'histogramme obtenu montre clairement que les pourcentages faibles sont majoritaires. Il n'est cependant pas judicieux de garder cette distribution dans notre simulateur. En effet si tel était le cas, la plupart des cartes générées ne contiendraient qu'un faible pourcentage de pluie. En pratique, on souhaite en général générer des cartes avec « suffisamment » de pluie. Nous prenons donc volontairement le parti de distordre la distribution naturelle afin d'augmenter artificiellement le nombre de cartes où la surface de pluie est significative. Pour cela, nous avons choisi de tirer aléatoirement ce paramètre suivant une loi uniforme entre 0.15 et 1. Le simulateur ne simulera donc que les cartes ayant un pourcentage de pluie supérieur à 15 % ce qui correspond en pratique sur nos données à 25% des cartes non nulles que nous avons observées et à 10% du total des cartes observées.



*Figure 4.6. Histogramme des paramètres des variogrammes empiriques et la probabilité d'occurrence de pluie calculés sur les 1277 cartes sélectionnées.*

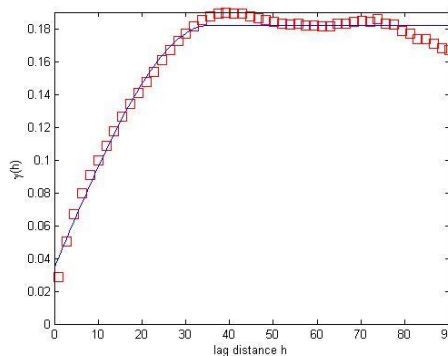
#### **4.4.4 Exemple de supports simulés**

Nous montrons dans cette section un exemple d'un support synthétique de dimension 128x128 généré avec la méthode présentée dans la section 4.4.2 et les paramètres estimés dans la section 4.4.3.2 : probabilité d'occurrence  $\mu=0.25$ , portée  $a=31$  km, palier  $c=0.17$  et un effet pépite égale à  $c_0=0$ . La Figure 4.7 montre à titre de simple comparaison visuelle une réalisation du support synthétique et un exemple de support réel dont les paramètres du variogramme sont proches des valeurs citées ci-dessus. Comme nous pouvons le remarquer sur cet exemple, dans chacune des deux cartes on observe un grand évènement qui occupe à peu près toute la zone de pluie avec quelques cellules de surfaces moindres de part et d'autres.



**Figure 4.7.** À gauche, exemple de support synthétique (128x128) généré avec les paramètres  $\mu=0.25$ ,  $a=31.44$ ,  $c=0.17$  et  $c_0=0$ . A droite, un support réel observé ayant des paramètres proches de ceux utilisés pour la simulation ( $\mu=0.29$ ,  $a=30.02$ ,  $c=0.16$  et  $c_0=0.02$ )

L'étude de la structure spatiale de la carte simulée via l'estimation du variogramme empirique (Figure 4.8) montre que celui-ci correspond à modèle sphérique de paramètres  $\mu=0.22$ ,  $a=33.12$  km,  $c=0.18$ ,  $c_0=0.03$  ce qui est conforme aux paramètres utilisés pour la simulation.



**Figure 4.8.** Variogramme empirique estimé sur la simulation présentée Figure 4.7 (courbe rouge). En bleu l'ajustement du variogramme par un modèle sphérique ( $\mu=0.22$ ,  $a=33.12$ ,  $c=0.18$ ,  $c_0=0.03$ ).

## 4.5 Simulation des lames d'eau dans les zones de pluie

### 4.5.1 Simulation des zones de pluie: étape 2

Il s'agit de générer des cartes de précipitation de taille  $N \times N$  comportant des zones sèches et des zones pluvieuses. La distinction entre les deux types de zone est réalisée grâce à la modélisation du support présentée précédemment. Il s'agit maintenant de « remplir » les zones pluvieuses afin de générer des cartes de précipitation réalistes. Pour cela, de la même manière que pour le 1D nous allons utiliser le modèle multifractal FIF étendu à la dimension 2. Il reste cependant une

question à traiter : doit-on utiliser une seule réalisation du modèle FIF de taille  $N \times N$  pour générer l'ensemble de la carte ou doit-on appliquer des réalisations indépendantes localement à chaque cellule présente dans la carte ? Dans le premier cas, il s'agit de générer à partir du modèle FIF une carte de dimension  $N \times N$  puis de la multiplier par le support (constitué d'une carte  $N \times N$  de 0 et de 1) afin de forcer à zéro les zones sèches. Cette solution outre sa simplicité présente l'avantage de préserver une certaine corrélation entre les différentes cellules pluvieuses présentes dans la carte compte tenu du caractère « mémoire longue » du modèle FIF. L'inconvénient d'une telle approche est le caractère brutal qu'elle implique où des pics de pluie peuvent être brutalement coupés lorsqu'ils sont à la frontière d'une zone sèche et pluvieuse. Les propriétés multifractales du champ simulé ne sont pas bien conservées par cette approche.

Une seconde approche consiste à faire une hypothèse similaire à celle faite en 1D concernant l'indépendance entre les périodes sèches et pluvieuse. Dans le cas du 2D cette hypothèse se traduit par l'indépendance entre les cellules. Cette hypothèse implique qu'il est possible de générer les cellules indépendamment les unes des autres, autrement dit, sur chaque région pluvieuse délimitée par un contour fermé, on simule un événement de pluie en utilisant le modèle FIF. Cette approche permet d'atténuer fortement l'inconvénient de la première approche concernant les transitions décrites dans le paragraphe précédent. C'est pour cette raison que nous avons retenu la dernière hypothèse d'indépendance entre les cellules pluvieuses.

En ce qui concerne les paramètres FIF nous avons utilisé les mêmes que ceux qui avaient été retenus dans le chapitre précédent relatif au 1D. Les raisons sont multiples : l'estimation des paramètres sur les cartes météo-France n'est pas satisfaisante du fait de la quantification/discrétisation « grossière » du produit « lames d'eau » fourni par Météo France, il n'est donc pas possible de proposer un jeu de paramètres fiable. (L'estimation des paramètres multifractales aux fines échelles, caractérisant des comportements non-conservatifs, est très sensible à la quantification.) Par ailleurs, des études menées par Verrier et al. (2011) ont montré que les paramètres multifractals sont constants aussi bien dans le domaine temporel que spatial dès lors qu'on ne considère que les relations d'échelle intra événement, c'est-à-dire pour des durées inférieures à une trentaine de minutes ou des distances inférieures à quelques dizaines de kilomètres. Il convient de signaler que la variabilité des paramètres proposés dans la littérature (domaine temporel essentiellement) est principalement due aux données elles-mêmes : certaines

analyses sont réalisées au moyen de données intégrées sur 5 minutes, d'autres sur 1 heures voire 24 heures. Ces données mélangent donc des périodes de pluie avec des périodes de sécheresse et induisent un fort biais sur l'estimation des paramètres multifractals. La région considérée dans ce chapitre (région parisienne) est identique à celle considérée pour les séries chronologiques, nous utiliserons donc le même jeu de paramètres, à savoir  $\alpha=1.60$ ,  $C_1=0.10$  et  $H=0.40$ .

Les valeurs des paramètres multifractals retenus sont proches de ceux estimées dans la littérature pour des cartes contenant 100 % de pluie. En effet, (Verrier et al. 2010) a trouvé  $\alpha = 1.78$ ,  $C_1 = 0.12$  sur des cartes mesurées par le radar RONSARD (Verrier et al., 2010) et ayant 97.4%. (Tessier et al. 1993) ont trouvé  $\alpha = 1.4$  sur des cartes radars tandis que (Lovejoy et al. 2008) ont trouvé 1.5 sur des cartes de réflectivités de TRMM. En ce qui concerne le paramètre  $C_1$  ces derniers ont estimé une valeur beaucoup plus forte (0.63) contrairement à Tessier et Verrier. Cependant comme nous l'avons signalé plus haut cette valeur est biaisée du fait du faible pourcentage de pluie de leurs observations. Concernant le paramètre  $H$ , la valeur que nous utilisons est cohérente à la fois avec celle estimée par (Verrier et al. 2010) qui ont trouvé  $H = 0.4$  mais aussi avec celle estimée par (Tessier et al. 1993) qui l'ont estimé à 0.32.

### ***4.5.2 Re-normalisation des lames d'eau simulées: Etape 3***

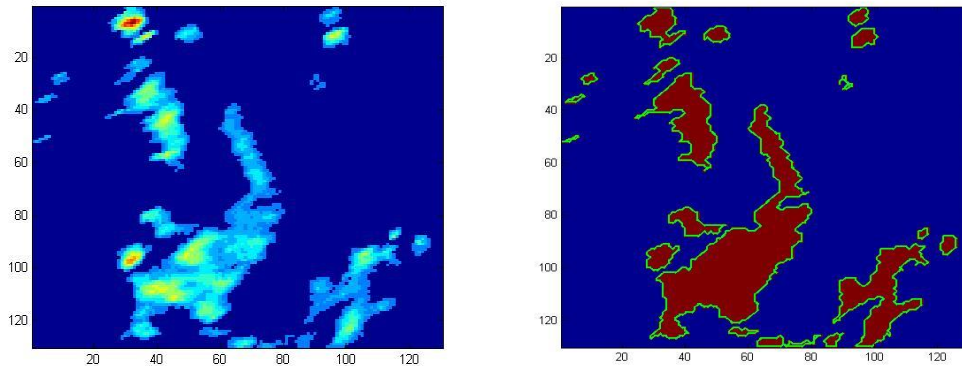
Comme en dimension 1 la méthodologie retenue pour réaliser le simulateur nécessite une étape de re-normalisation des lames d'eau générées par le modèle FIF.

Cette étape permet de dé-normaliser les différents événements générés par le modèle FIF afin qu'ils soient représentatifs de lames d'eau. Le principe retenu est de trouver une relation entre la valeur moyenne de lame d'eau de la cellule à sa taille (ou plus exactement sa longueur équivalente).

#### **4.5.2.1 Relation entre la surface et la lame d'eau moyenne des événements**

Dans un premier temps, j'ai extrait les cellules de pluie des 42 000 cartes disponibles. Cette extraction a permis d'isoler 330 346 cellules de surface comprise entre  $2.5 \times 2.5 \text{ Km}^2$  et  $18 \times 18 \text{ Km}^2$ .

La Figure 4.9 montre un exemple d'extraction des cellules de pluie. Cette extraction est basée sur l'algorithme Moore-Neighbor Tracing (Gonzalez, 2004). Un évènement de pluie est un ensemble de pixels voisins non nuls qui définit un contour fermé. La surface d'une cellule est égale à la somme des surfaces élémentaires de chaque pixel la constituant. La lame d'eau moyenne par cellule est la moyenne des lames d'eau des pixels qui la constituent.



*Figure 4.9. Exemple d'extraction des évènements (cellules) de pluie : à gauche la carte radar des précipitations, à droite les évènements extraits entourés du vert*

Pour chaque cellule, nous avons considéré sa longueur équivalente  $L$  définie comme la longueur d'un côté d'un carré de même surface : ( $L = \sqrt{\text{surface de la cellule}}$ ). Nous avons également calculé la lame d'eau moyenne correspondante. Le choix des classes est effectué de manière à avoir un nombre de cellules significatif dans chaque classe. La Figure 4.10 présente les lames d'eau moyennes par classe en fonction de leur longueur équivalente. Une relation linéaire par morceau apparaît plus ou moins clairement avec une cassure aux alentours de 8 kilomètres. La pente est moins forte au-delà de la cassure.

Pour chaque classe de distance  $L$  (i.e. chaque taille de cellule) la lame d'eau peut être considérée comme une variable aléatoire continue et distribuée de manière asymétrique. Nous avons pu correctement modéliser cette distribution à l'aide d'une loi  $\alpha$ -stable. Dans des contextes proches, d'autres auteurs ont utilisé la distribution d'extremum généralisée (e.g. Barnolas et al. 2010). Les auteurs ont analysé des cellules convectives de surface supérieure à  $32 \text{ Km}^2$  dans la région de la Catalogne en Espagne avec une résolution spatiale de 2 Km et une résolution temporelle de 10 minutes. Dans l'étude proposée par Von Hardenberg et al. (2003), les auteurs ont choisi une loi exponentielle pour les cellules convectives. Notre étude ne se focalise pas sur les cellules convectives mais cherche à représenter l'ensemble des cellules observées pendant l'année 2012.

Parmi les lois testées à l'aide d'un ajustement par maximum de vraisemblance (gamma, gev, exponentiel,  $\alpha$ -stable) cette dernière déjà utilisée pour la modélisation 1D s'avère être la plus à même de s'ajuster à nos données.

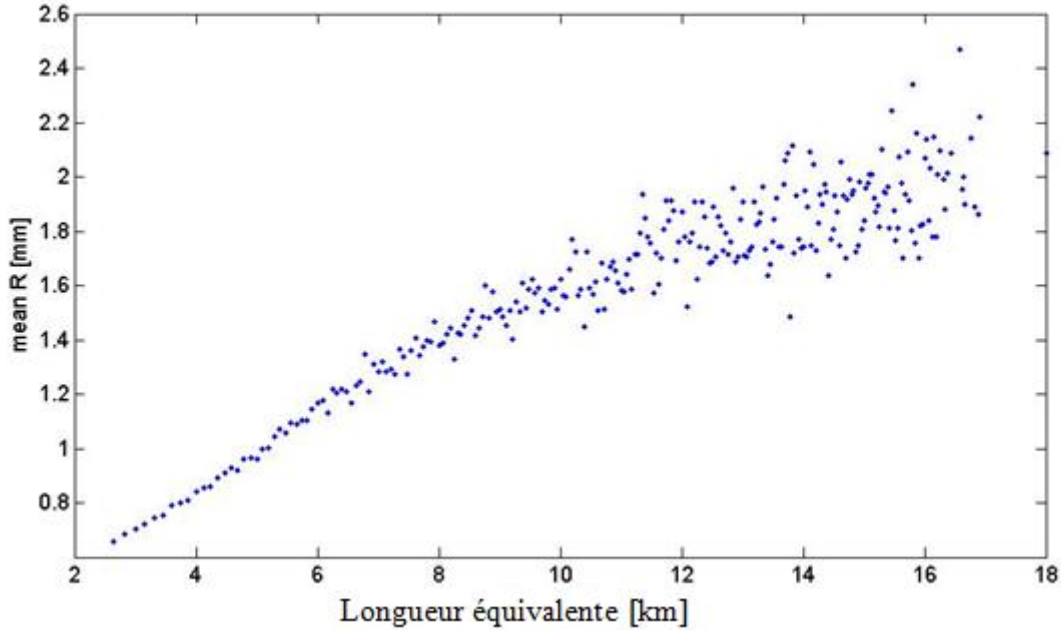


Figure 4.10. Moyenne des lames d'eau par cellule en fonction de la dimension des cellules de pluie.

La

Figure 4.11 présente un exemple de distribution de la lame d'eau moyenne par événement pour des événements dont la longueur équivalente est comprise entre 9 et 10 kilomètres (soit des surfaces comprises entre 81 et 100 km<sup>2</sup>).

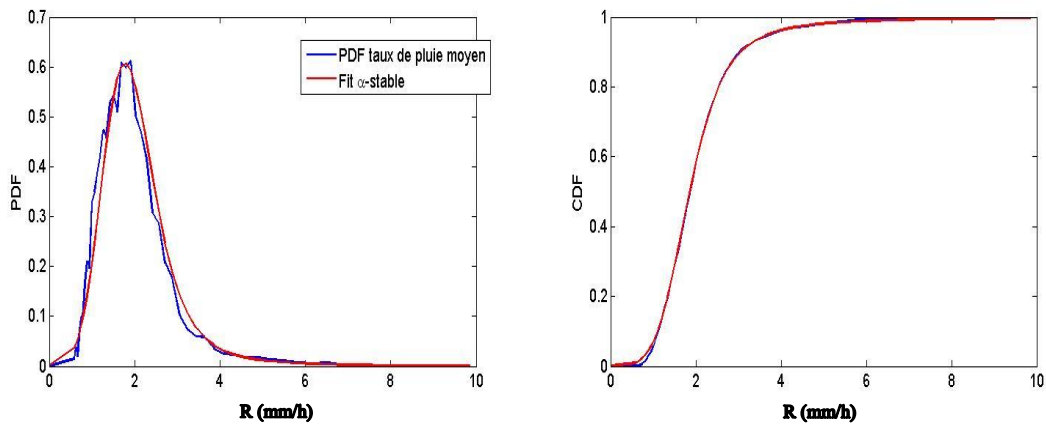


Figure 4.11. Exemple d'ajustement sur nos données de la loi  $\alpha$ -stable sur la distribution empirique des lames d'eau par évènement pour une longueur équivalente comprise entre 9 et 10 km



Cet ajustement de la loi  $\alpha$ -stable sur les lames d'eau observées est réalisé pour l'ensemble des évènements par classe de même longueur équivalente. Pour chaque classe de longueur équivalente, on calcule une lame d'eau moyenne  $R$  (Figure 4.10) ainsi que les quatre paramètres de la distribution  $\alpha$ -stable (

Figure 4.11). Les quatre paramètres sont présentés sur la Figure 4.12 en fonction de la longueur équivalente. Nous remarquons que le paramètre de stabilité ( $\alpha$ ) est très peu variable, il est dans 99% des cas supérieur à 1, ce qui veut dire que la moyenne est définie et qu'elle peut être égale au paramètre de localisation ( $\mu$ ). Le paramètre d'asymétrie ( $\beta$ ) quant à lui possède des valeurs entre 0 et 1, mais 60% de valeurs sont égales à 1. Dans le processus de modélisation, le but est d'avoir un modèle réaliste avec un minimum de paramètres possible.

Nous considérons donc  $\alpha$  et  $\beta$  constants, soit  $\alpha=1.18$ , la moyenne des valeurs estimés et  $\beta=1$  puisque la distribution est fortement asymétrique. Pour les paramètres d'échelle ( $\gamma$ ) et de location ( $\mu$ ), nous remarquons qu'ils augmentent d'une façon quasi linéaire en fonction de la longueur équivalente. Nous avons donc fait une régression linéaire pour chacun des deux paramètres. Au final on estime les 4 paramètres de la loi  $\alpha$ -stable dont deux sont conditionnés par la longueur équivalente  $L$  :

$$\left\{ \begin{array}{l} \alpha = 1,18 \\ \beta = 1 \\ \gamma = 0.03 \times L \\ \mu = 0.09 \times L + 0.65 \end{array} \right. \quad \text{Équation 4.7}$$

Ces équations définissent notre relation lame d'eau par événement / longueur équivalente. Pour chaque événement de longueur équivalente  $L$ , ce modèle stochastique nous permet de tirer aléatoirement une valeur de lame d'eau de l'événement conditionnée à sa longueur équivalente. Cette valeur est ensuite utilisée à l'étape de re-normalisation des champs multifractals générés par le modèle FIF.

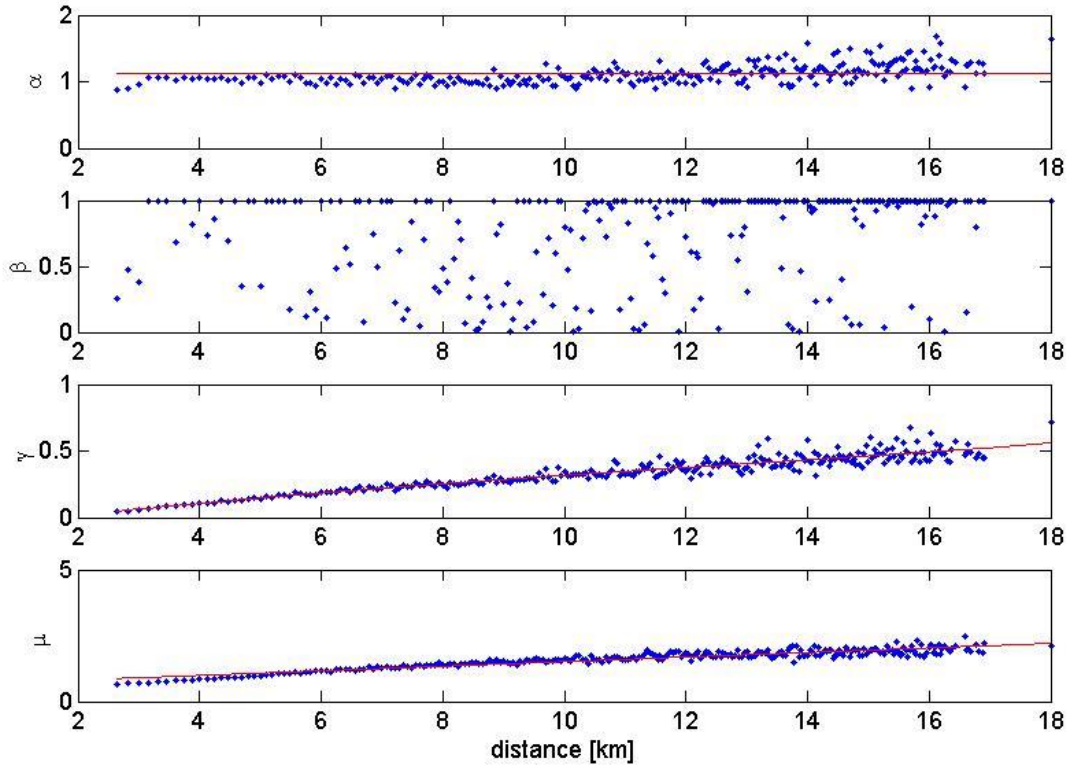


Figure 4.12. Paramètres de stabilité, d'asymétrie, d'échelle et de location estimé pour les lois  $\alpha$ -stable ajustés sur les lames d'eau des évènements observés par classe de distance

La procédure de dé-normalisation est la suivante :

- Pour chaque cellule de pluie  $i$ , on détermine sa longueur équivalente  $L = \sqrt{S \times N_i}$  ( $N_i$  : nombre de pixel de la cellule  $i$ ,  $S$  : surface élémentaire d'un pixel).
- On tire aléatoirement une valeur de lame d'eau moyenne  $R_{mg}$  conditionnée à sa longueur  $L$  à l'aide d'une loi  $\alpha$ -stable de paramètres.
- Pour chaque pixel  $k$  de la cellule  $i$  (noté  $R_{ik}$ ) on calcule sa valeur de lame d'eau correspondante

$$R_{ik} = \frac{R_{ik}^{FIF}}{R_{i,m}} R_{mg} \quad \text{Équation 4.8}$$

$R_{ik}^{FIF}$  est la valeur du  $k$  ième pixel généré par le modèle FIF de l'évènement  $i$ .

$R_{i,m}$  est la valeur moyenne des pixels de l'évènement  $i$  générés par le modèle FIF :

$$R_{i,m} = \frac{1}{N_i} \sum_{k=1}^{N_i} R_{ik}^{FIF}$$

### 4.5.3 Traitement des transitions pluie/non pluie

Les étapes 1 à 3 permettent d'obtenir des cartes de précipitation ayant des caractéristiques statistiques similaires à celles observées. Néanmoins, en visualisant ces cartes, nous nous rendons compte que bien que l'intérieur des évènements soit réaliste, les transitions pluie/non pluie ne le sont pas. Pour cela, nous avons rajouté au processus de simulation une étape supplémentaire qui a pour but le traitement des transitions. Nous nous sommes inspirés de la méthode « dry drift » développée par (Schleiss et al. 2014) qui, pour chaque pixel d'une zone pluvieuse, définit sa distance de la zone sèche suivant la formule :

$$d(z) = \min(\|z - y\|) \quad \text{Équation 4.9}$$

Tel que  $z \in \{x, I(x) = 1\}$  et  $y \in \{x, I(x) = 0\}$  .

Suite à l'étude de notre jeu de données, nous supposons également qu'au-delà d'une distance de  $d_0=7$  kilomètres la zone de transition est terminée, on définit la fonction  $F(d)$  :

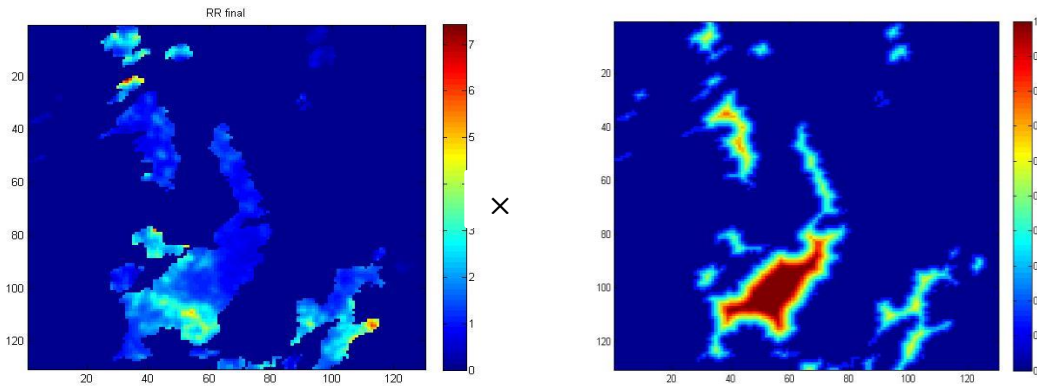
$$\begin{cases} F(d) = \frac{d}{d_0} & \text{si } d < d_0 \\ F(d) = 1 & \text{si } d > d_0 \end{cases} \quad \text{Équation 4.10}$$

En calculant  $F(d)$  pour chaque pixel, on déduit une carte de transition. Le traitement ultime consiste à multiplier pixel à pixel la carte des lames d'eau obtenue à l'étape 3 par la carte de transition.

L'avantage de la méthode de (Schleiss et al, 2014) est qu'elle ne modifie pas la structure spatiale de la carte. En effet, ce traitement n'est pas un filtrage linéaire de type filtre passe bas (il n'y a pas de produit de convolution), les pics intenses (pour peu qu'ils soient un peu éloignés du bord) se sont pas lissés. Les Figure 4.13 et Figure 4.14 présentent l'ensemble du processus de traitement des transitions. La Figure 4.13 de gauche montre un exemple de carte de lame d'eau avant l'étape de traitement de transition tandis que la Figure 4.13 de droite montre la carte de transition associée. Il apparait que seul l'évènement situé dans la partie centrale inférieure est peu affecté par le traitement hormis sur ses bords tandis que tous les autres évènements se voient

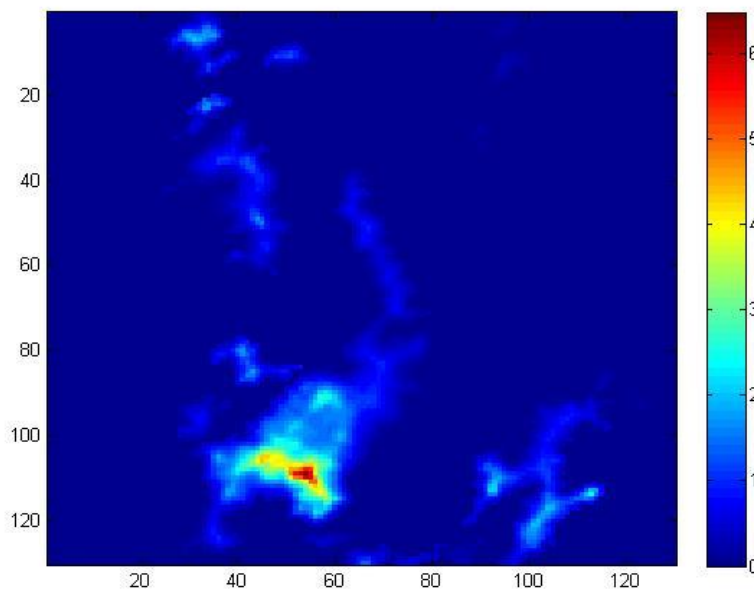
pondérés par le traitement. Cet effet de « rognage » est particulièrement important pour des cellules de faible dimension, c'est le cas par exemple pour la cellule dont le maximum est situé aux coordonnées (40,20) sur la Figure 4.13 de gauche qui présente un maximum aux alentours de 7 mm et qui après traitement voit son maximum réduit à 2 mm (Figure 4.14). Nous verrons plus loin quelle conséquence cela implique en termes de distribution de lame d'eau.

Finalement, la carte finale de lame d'eau est présentée à la Figure 4.14, il appartient maintenant de vérifier les caractéristiques statistiques des cartes générées.



*Figure 4.13. Exemple de carte simulée avant l'application du traitement des transitions. A droite carte des transitions  $F(d)$  calculée sur la carte simulée*

Le produit de ces deux cartes, est le résultat final présenté sur la Figure 4.14 suivante :



*Figure 4.14. Résultat final du processus de simulation*

En résumé, la procédure de simulation consiste à :

1. *Générer un support de pluie  $I(x)$  de dimension  $N \times N$ . (Etape 1)*
2. *Pour chaque événement (cellule) de pluie présent dans le masque (Etape 2 et 3) :*
  - a. *Détecter le sous masque de forme carré ou rectangulaire  $IR(x)$  qui circonscrit la cellule de pluie.*
  - b. *Simuler une carte  $RR$  de la même taille que le sous masque  $IR$  au moyen du modèle FIF en utilisant le jeu de paramètres cité plus haut*
  - c. *Re-normalisation de la carte  $RR$  afin d'obtenir des lames d'eau 5 minutes*
  - d. *Copier la carte  $RR$  normalisée dans la carte  $R(x)$  à l'emplacement du sous masque  $IR(x)$*
3. *Appliquer sur l'ensemble de la carte un pseudo-filtre pour lisser les transitions pluie/non pluie (Etape 4)*

La méthodologie de simulation des évènements pluie est présentée dans la Figure 4.15:

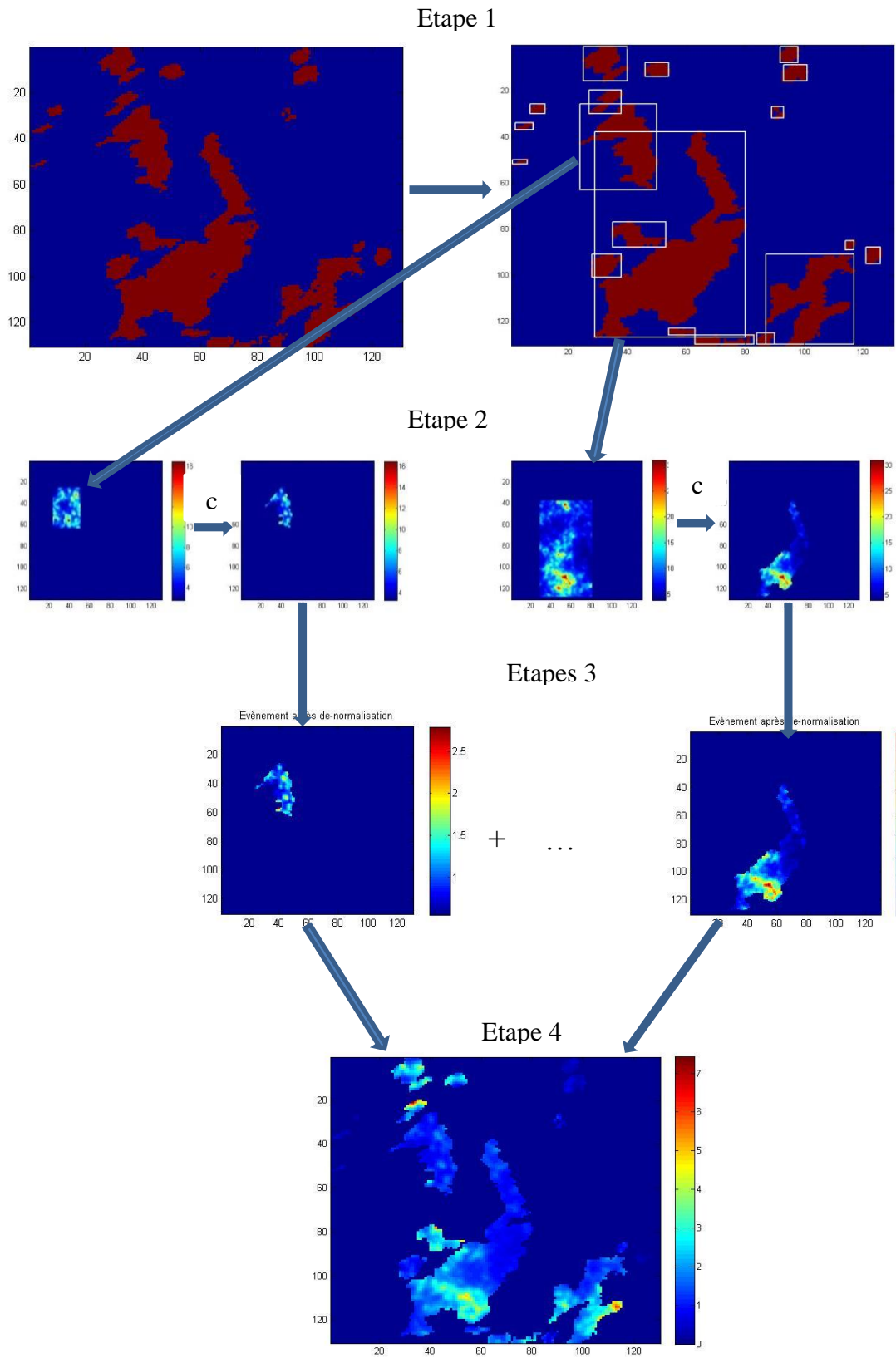


Figure 4.15 Méthodologie de simulation d'une carte de lame d'eau

## 4.6 Validation des cartes simulées

### 4.6.1 Propriétés multifractales des simulations

Nous vérifions que les cartes obtenues présentent les propriétés attendues, en particulier en ce qui concerne les propriétés multifractales micro-échelles qui caractérisent la variabilité de la lame d'eau à l'intérieur des cellules de pluie.

Afin d'avoir une idée des paramètres multifractal des cartes simulées, j'ai fait une analyse multifractale sur une sélection de 50 d'évènements de  $40 \times 40 \text{ km}^2$  contenant 100% de pluie. J'ai calculé ensuite les paramètres multifractals moyens. La Figure 4.16 de gauche montre la fonction d'échelle des moments  $K(q)$  obtenue sur les 50 évènements (courbe bleu) ainsi que son ajustement (Equation 2.4) (courbe rouge). Les paramètres estimés moyen valent :  $\alpha = 1.71$ ,  $C_1 = 0.10$ . Concernant le paramètre de non conservativité  $H$  la Figure 4.16 de droite montre la fonction de structure d'ordre 1. Une regression linéaire pour des valeurs de  $\delta x$  comprises entre 1 et 16 km fournit une valeur de  $H$  égale à 0.38.

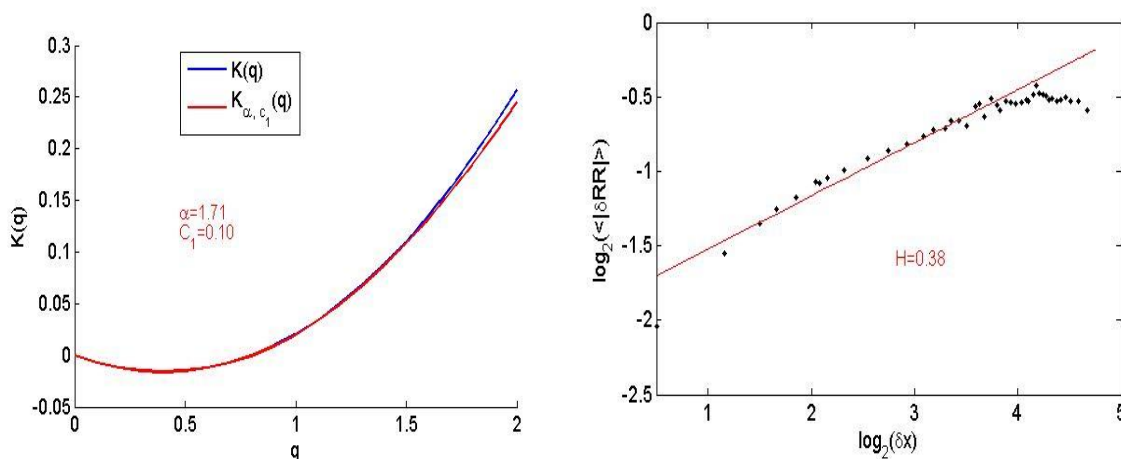


Figure 4.16. Paramètres multifractals estimés sur une sélection de 50 évènements de  $40 \times 40 \text{ Km}^2$  extraits des cartes simulées.

Bien que ces paramètres estimés ne représentent qu'un faible échantillon de nos simulations, nous pouvons dire d'après cette analyse que notre modèle permet de générer des cartes ayant les caractéristiques multifractales attendues et que le processus de « rognage » de l'étape de traitement des transitions n'affecte pas trop les relations d'échelle. En effet les paramètres  $C_1$  et

H sont bien restitués. Le paramètre  $\alpha$  est surestimé (valeur estimée de 1.71) mais il reste cohérent avec la valeur de (1.6) utilisée pour la simulation.

## 4.6.2 Propriétés spectrales

### 4.6.2.1 Analyse des propriétés spectrales des observations

L'analyse spectrale permet d'étudier les propriétés du moment d'ordre deux du champ. Comme déjà indiqué pour les séries chronologiques, les propriétés d'invariance d'échelle des précipitations se traduisent par un spectre en loi de puissance.

Les spectres de l'ensemble des 42000 cartes 2D de Palaiseau sont présentés en bleu sur la Figure 4.18. L'estimation de ces spectres repose sur une hypothèse d'isotropie qui permet de passer d'un spectre 2D en un spectre 1D par intégration à rayon constant. Chaque courbe bleue correspond à une carte de précipitation. Le spectre moyen annuel calculé sur l'ensemble des cartes est présenté en rouge.

Comme nous le confirme ces courbes, ces cartes de précipitation présentent des propriétés d'invariance d'échelles par morceau (une droite en coordonnées log-log). En considérant le comportement moyen annuel, deux gammes d'invariance d'échelle peuvent être mises en évidence, l'une pour des échelles inférieures ou égale à 7 km environ et l'autre au-delà. Ces deux régimes « scalant » observés sur nos données sont liés à la structure spatiale des précipitations. Ainsi (Waymire et Gupta, 1981) ont suggéré que le régime au-delà de 7 km serait plutôt représentatif des processus à méso-échelle alors qu'en deçà il s'agirait de comportements représentatifs de la variabilité intra cellulaire (Figure 4.17), notamment de cellules convectives.

L'estimation de la pente sur le régime de scaling aux plus fines échelles (<7 km) fournit une valeur  $\beta = 1,95$ , c'est-à-dire nettement supérieure à 1. Compte tenu de la relation qui relie la pente à la valeur de H :

$$\beta = 1 - K(2) + 2H \tag{Equation 4.11}$$

Et en remarquant que  $K(2)$  est positif, il en découle que H est positif et que le champ est non conservatif. Cela est conforme à notre hypothèse de non conservativité (  $H=0.4$ ).



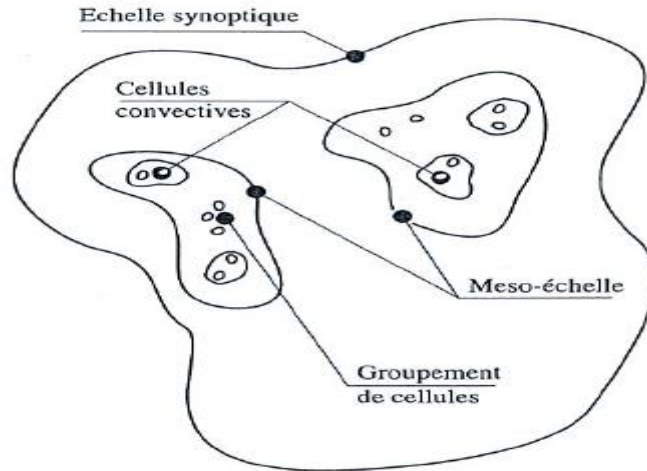


Figure 4.17. Schéma de la structure spatiale des précipitations à l'intérieur de l'échelle synoptique (Gupta et Waymire 1979; 1981; 1993)

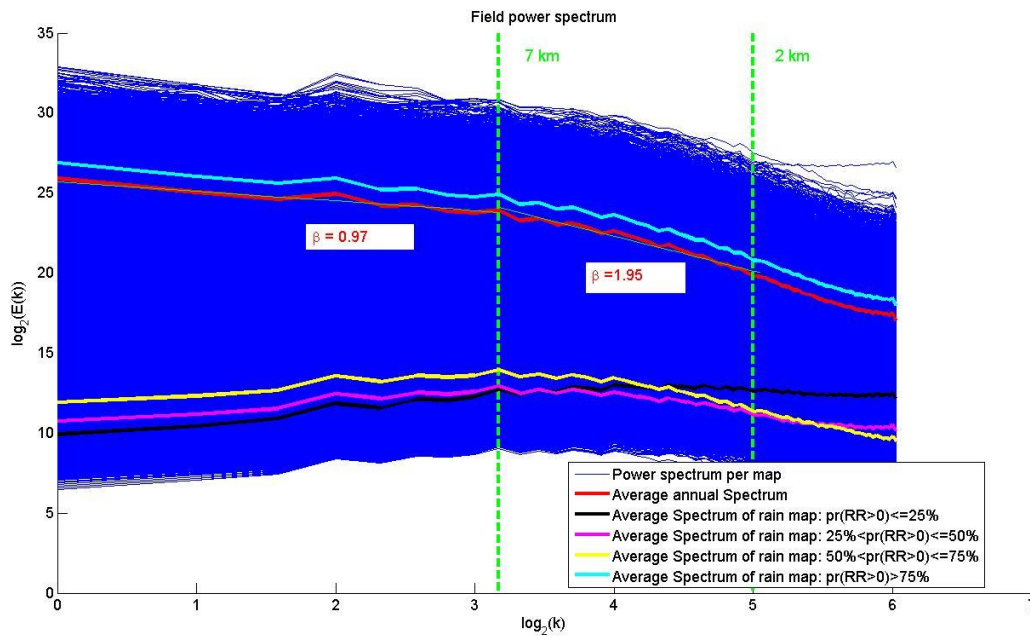


Figure 4.18. Variabilité spectrale de 42000 cartes radar Météo France de taille 130\*130 km observées sur la région de Palaiseau. La moyenne de tous les spectres est présentée en rouge, les courbes noir, rose, jaune et bleu claire sont les spectres moyen des cartes ayant respectivement un pourcentage de pluie <25%, entre 25% et 50%, entre 50% et 75% et au de-là de 75% .

Cette valeur de pente est à comparer avec la pente obtenue sur des cartes de précipitation de mousson africaine dans la gamme d'échelle 1km à 10 km obtenue par (Verrier et al. 2010). Ce dernier avait obtenu une valeur sensiblement plus faible ( $\beta = 1.47$ ) laissant présager un

processus moins « lissé ». Il en est de même pour la pente estimée par (Tessier et al. 1993) ( $\beta = 1.45$ ) estimée sur les cartes de réflectivités radar dans la gamme d'échelle 75m à 10km. Ces auteurs ont estimé les pentes sur des cartes comportant un faible pourcentage de zéros tandis que notre étude est faite sur les cartes de précipitation avec un pourcentage de zéro variable et qui peut atteindre 95%. Afin d'apprécier l'influence des zéros, le calcul de la pente a été réalisé sur des sous ensemble de cartes sélectionnés en fonction de leur pourcentage de zéros et pour les deux régimes. Le Tableau 4.2 et Figure 4.18 résument les valeurs des pentes estimées et le spectre correspondant en fonction du pourcentage de zéros. En ce qui concerne le régime à fine échelle (<7 km) les différentes courbes de la Figure 4.18 montrent que la valeur de la pente croît avec le pourcentage de zéros passant d'une valeur de 0.06 pour un pourcentage de pluie inférieur à 25% à 2.17 lorsque le pourcentage de pluie est supérieur à 75%. De plus, il semblerait que la cassure soit située vers 4 km lorsque le pourcentage de pluie est faible (<50%).

En ce qui concerne la pente aux échelles plus larges (> 7 km), sa valeur estimée est négative pour les faibles valeurs du pourcentage de zéros (-0.96). Compte tenu de la formule précédente (Equation 4.11) cette valeur de pente pourrait conduire à conclure sur une valeur de H négative. Lorsque le pourcentage de zéros augmente la pente devient positive et passe brusquement à une valeur positive située aux alentours de l'unité.

Pentes Pourcentage de pluie	$\beta$ moyen dans la gamme 2km – 7 km	$\beta$ moyen dans la gamme 7km – 130 km
$\Pr(RR > 0) \leq 0.25$	0.06	-0.96
$0.25 < \Pr(RR > 0) \leq 0.5$	0.74	-0.85
$0.5 < \Pr(RR > 0) \leq 0.75$	1.36	-0.67
$\Pr(RR > 0) \geq 0.75$	2.17	1.05
Moyenne sur toutes les cartes	1.95	0.97

**Tableau 4.2 les pentes moyennes des spectres obtenues en fonction du pourcentage de pluie dans les cartes de précipitation**

(Rysman et al. 2013) ont fait une étude similaire sur des cartes contenant des zéros issues du radar Hydrix situé à 1,5km de Nice. Ils ont estimé des pentes de  $\beta = 2,29$  dans la gamme 3-7 km et  $\beta = 0.94$  dans la gamme 20- 70 km pour les cartes observées en été. En hiver, ils ont obtenu une pente sensiblement plus faible  $\beta = 1.55$  dans la gamme 3-7 km et  $\beta = 1.55$  dans la gamme 20- 70 km. Les pentes que nous avons estimées sur la région parisienne sont en cohérence avec les pentes estimées par (Rysman et al. 2013) si on considère les pentes calculées sur la moyenne

de l'année (moyenne hiver + été) :  $\beta = 1,92$  dans la gamme 3-7 km et  $\beta = 1,24$  dans la gamme 20- 70 km.

Le Tableau 4.3 présente quelques grandeurs statistiques de la pente calculées sur les 42000 cartes de 130 Km de côté et contenant au minimum 15% de pluie. Comme nous pouvons le remarquer l'écart type est élevé pour les deux régimes, ce qui traduit la grande variabilité observée sur la Figure 4.18.

	$\beta$ dans la gamme 2km – 7 km					$\beta$ dans la gamme 7km – 130 km				
	Mean	std	Q1	Q2	Q3	mean	Std	Q1	Q2	Q3
42000 cartes 2D	1.95	0.7	1.50	1.88	2.30	0.97	0.5	0.33	0.80	1.15

**Tableau 4.3** Moyenne, écart type et quartiles de  $\beta$  calculés sur l'ensemble de 42000 cartes observées dans les deux gammes d'échelles 2-7 km et 7-11 km

Il est intéressant dans le cas où les cartes analysées ont un pourcentage de zéros inférieur à 100% d'analyser l'influence du support de pluie sur le spectre. Pour cela, nous avons analysé le spectre moyen des spectres des supports de pluie  $I(x)$  de ces 42000 cartes afin de le comparer au spectre moyen obtenu directement avec les spectres des cartes. La courbe bleue de la Figure 4.19 montre le spectre moyen obtenu sur l'ensemble des spectres des cartes de pluie tandis que la courbe rouge montre le spectre moyen obtenu à partir des spectres du support. On peut remarquer que le spectre du support ne semble pas présenter de cassure à 7 km aussi nette que dans le cas du spectre des cartes de pluie. En effet, la différence de pente entre les deux régimes vaut (0.64) pour le support alors que cette différence est égale à pratiquement 1 pour les cartes. Dans la gamme d'échelle comprise entre 2 et 7 km nous remarquons que la pente du support (1.69) est plus faible que celle estimée sur les cartes (1.95) ce qui signifie que ce régime est lié à la variabilité de la lame d'eau à l'intérieur des événements. Au-delà de 7 km, la différence entre les deux pentes est négligeable, ce qui revient à dire qu'aux grandes échelles le support gouverne en grande partie la pente du spectre.

Le décalage vertical entre les deux courbes n'est pas à prendre en compte, il montre seulement que l'énergie moyenne est différente, ce qui est logique dans la mesure où le support présente des valeurs maximales égales à 1 alors que les lames d'eau présentent des valeurs pouvant atteindre 200 mm.

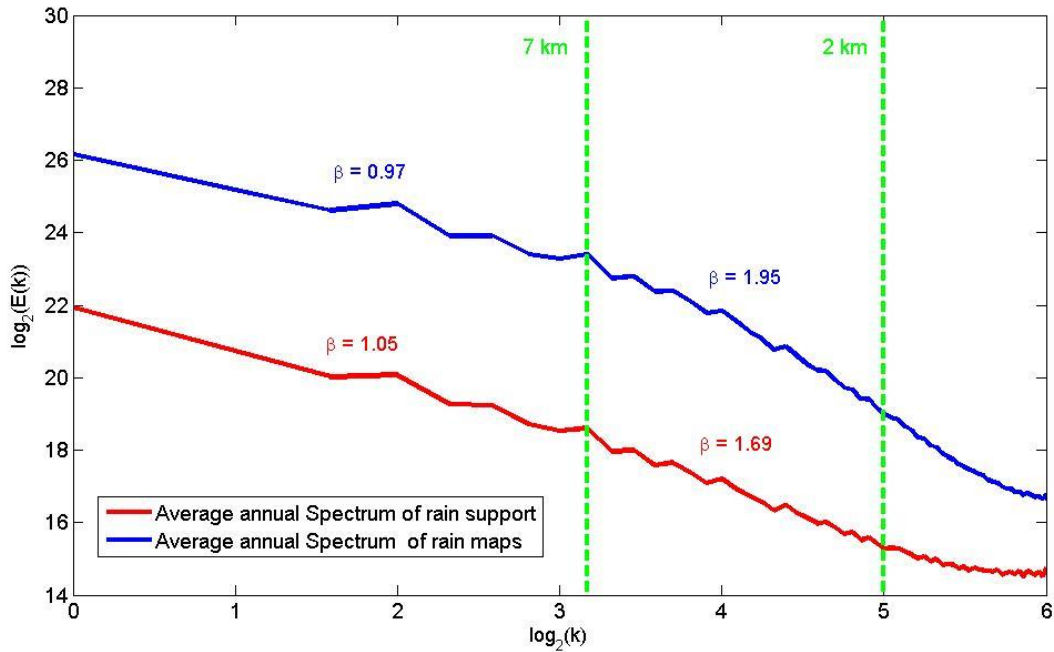


Figure 4.19. Spectre annuel moyen des cartes précipitation en bleu et le spectre moyen annuel de leurs supports

#### 4.6.2.2 Analyse des propriétés spectrales des simulations obtenues avec support conditionnel (réel)

Nous utilisons ici les supports de pluie issus des 42000 cartes observées dans la région de parisienne pour l'année 2012 et ayant un pourcentage de pluie excédant 5%. A partir de ces supports de pluie nous avons simulé, comme expliqué précédemment dans la méthodologie, des cartes de précipitation. Au total nous avons simulé l'équivalent de 4 ans de données, c'est à dire 4 jeux de 42000 cartes chacun. Nous avons ensuite fait une analyse spectrale année par année afin d'avoir une idée de la variabilité. La Figure 4.20 présente les spectres moyens annuels obtenus sur les quatre jeux de données simulées et sur les 42000 cartes observées tandis que le Tableau 4.4 résume les différentes valeurs de la pente et les écarts-types associés pour les 2 régimes. En ce qui concerne la grande échelle les valeurs moyennes des 4 pentes sont comprises entre 0.8 et 0.95 soit une variabilité relativement faible, et qui de plus est, très inférieure à l'écart type des moyennes estimées. La variabilité est encore plus faible si on considère la médiane (0.75-0.80). Ce résultat paraît logique dans la mesure où nous avons montré précédemment qu'aux grandes échelles, le support conditionne en grande partie la pente du spectre.

En ce qui concerne le régime 2-7 km, les valeurs moyennes des 4 pentes varient entre 1.8 et 1.93 (1.90 et 1.97 pour la médiane) pour une valeur moyenne de 1.95 sur les cartes observées.

Nous pouvons conclure que toutes les grandeurs calculées (moyenne, écart type, quartiles) sont cohérentes avec celles obtenues sur les observations et par voie de conséquence que notre modèle arrive à restituer les régimes observés.

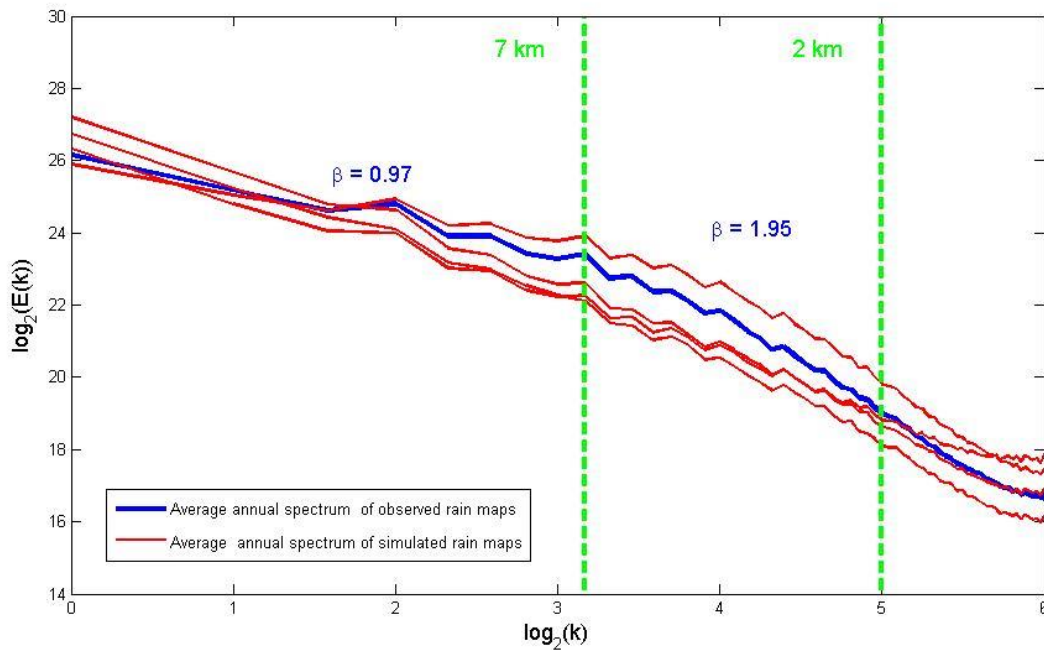


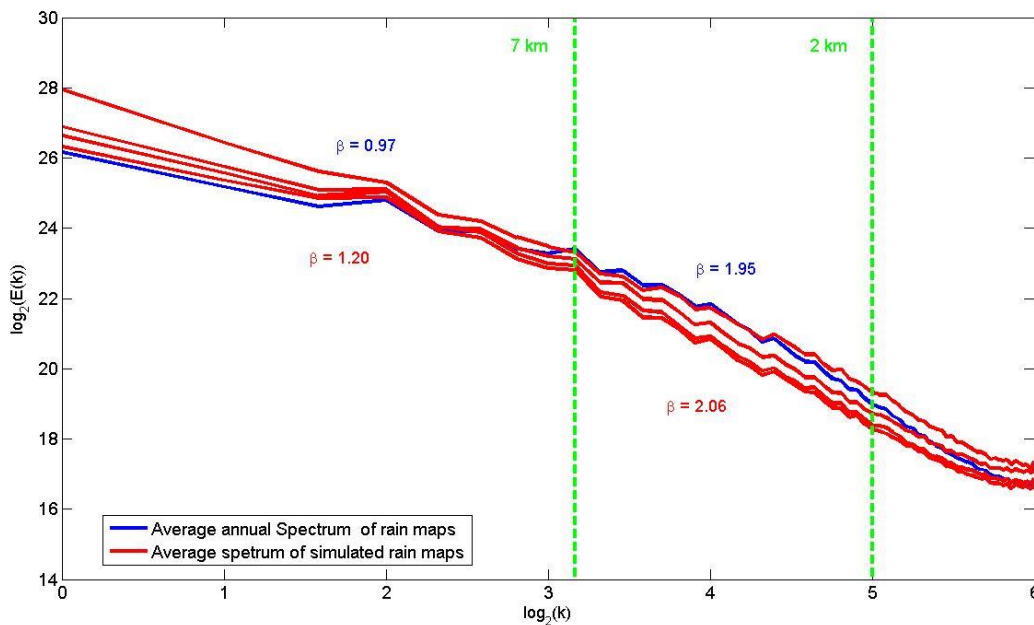
Figure 4.20. Spectre moyen des données observées en bleu (1 an soit 42 000 cartes) et des données simulées en rouge (4 ans soit 4\*42000) dans le cas conditionnel (support réel)

	$\beta$ dans la gamme 2km – 7 km					$\beta$ dans la gamme 7km – 130 km				
	Mean	std	Q1	Q2	Q3	mean	Std	Q1	Q2	Q3
Observation	1.95	0.7	1.50	1.88	2.30	0.97	0.5	0.33	0.80	1.15
1 <sup>ere</sup> année de simulation	1.80	0.8	1.30	1.96	2.28	0.80	0.6	0.32	0.75	1.10
2 <sup>eme</sup> année de simulation	1.85	0.78	1.28	1.97	2.34	0.86	0.52	0.32	0.77	1.12
3 <sup>eme</sup> année de simulation	1.93	0.8	1.41	1.90	2.31	0.94	0.6	0.33	0.77	1.13
4 <sup>eme</sup> année de simulation	1.87	0.72	1.35	1.91	2.32	0.95	0.55	0.32	0.80	1.14

Tableau 4.4 Résumé des grandeurs statistiques calculées sur les pentes des spectres des observations et des simulations dans le cas conditionnel.

### 4.6.2.3 Analyse des propriétés spectrales des simulations obtenues avec support inconditionnel (synthétique)

Contrairement à la simulation conditionnelle nous utilisons dans la simulation inconditionnelle des supports de pluie générés avec la méthode décrite dans la section 4.4.2 avec les paramètres estimés sur les observations de la section 4.4.3, soit une loi de probabilité d'occurrence uniforme entre 0.15 et 1 et les paramètres de la structure spatiale (portée uniforme entre 10 et 100 km, palier 0.17 et effet de pépite nul). Vu le temps de simulation qu'il faut pour simuler  $42000 \times 4$ , nous avons réduit le nombre de carte à 1000 pour chaque réalisation. Nous avons ainsi simulé 4 réalisations, chacune contenant 1000 cartes de pluie, avec la même méthodologie que celle décrite en section 4.3. Les 4 spectres moyens issus des 4 réalisations sont présentés sur la Figure 4.21 ci-dessous.



**Figure 4.21. Spectre moyen des données observées en bleu (42000 cartes) et des données simulées en rouge (soit 4\*1000) dans le cas inconditionnel**

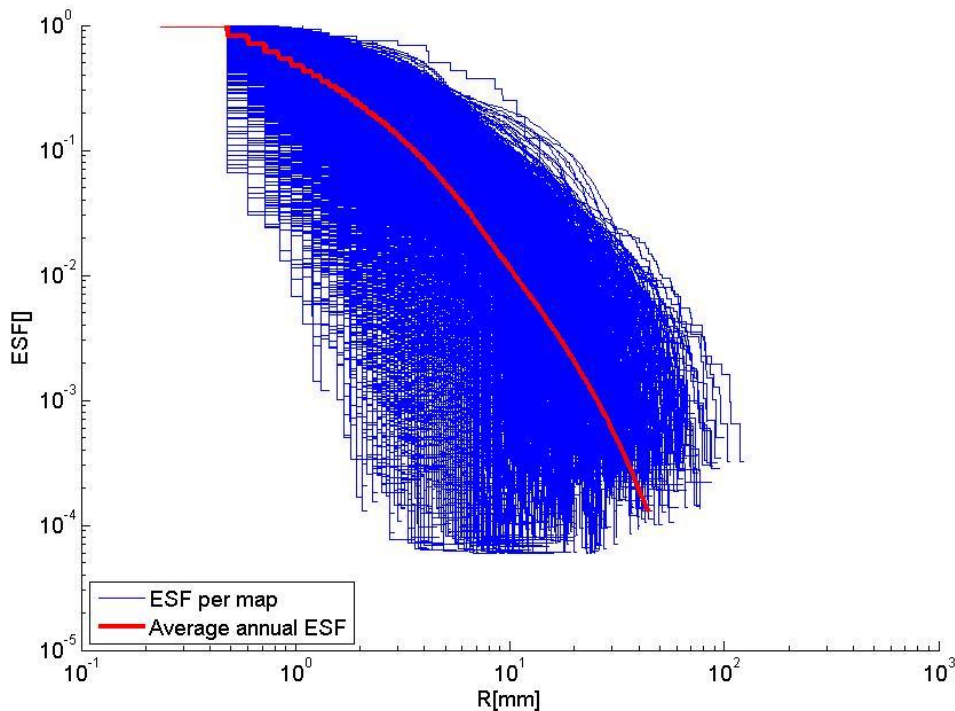
Nous remarquons que le régime de scaling 2-7 km est bien restitué avec une pente moyenne (2.06) cohérente avec celle observée (1.95). En revanche en ce qui concerne le régime au-delà de 7 km, qui comme nous l'avons vu précédemment est fortement conditionné par le support, la pente moyenne est plus forte (1.20) que la pente observée (0.97). Nous en déduisons donc que cette surestimation est due au modèle de génération du support de pluie. Il convient de relativiser cet écart qui est somme toute relativement faible. Qui plus est, la valeur obtenue reste cohérente

avec d'autres observations, ainsi Rysman et al. (2013) ont estimé une pente  $\beta = 1.55$  dans la gamme 20-70 km.

### 4.6.3 Distributions statistiques

#### 4.6.3.1 Analyse statistiques des observations

Quelle que soit l'échelle spatiale sur laquelle on travaille, les précipitations présentent une forte hétérogénéité et des gradients importants aussi bien spatialement que temporellement (Molini et al. 2002) existent. La Figure 4.22 présente les fonctions de survie empiriques (ESF) des 42000 cartes observées (courbes bleues) et la moyenne de ces courbes (courbe rouge). Cette figure met en évidence cette importante variabilité.



**Figure 4.22.** Fonction de survie empirique des lames d'eau des 42000 cartes observées à la résolution 1x1 km<sup>2</sup> à Palaiseau en 2012, la ESF moyenne est présentée en rouge

Je rappelle que la valeur minimale de lame d'eau considérée ici est de 0.48 mm. La valeur maximale enregistrée, très variable d'une carte à l'autre, peut atteindre 200 mm sur certaines cartes. La moyenne, l'écart type et trois quartiles des maxima sont présentés au Tableau 4.5. Comme nous pouvons le remarquer les maxima des lames d'eau observés durant l'année 2012

ont une moyenne de 12.8 mm et 75% des maxima observés ne dépassent pas 15mm, soit une intensité de pluie équivalente de 180 mm/h.

	Moyenne	L'écart type	Q1	Q2	Q3
Valeurs maximales	12.85 mm	20.31 mm	2.52 mm	6.12 mm	14.76 mm

Tableau 4.5 Caractéristiques statistiques des valeurs maximales observées sur les cartes radar de Palaiseau

#### 4.6.3.2 Analyse statistique des simulations avec support conditionnel (réel)

La Figure 4.23 montre la variabilité annuelle des fonctions de survie des quatre réalisations de données simulées (42000 \* 4) ainsi que la fonction de survie annuelle des données observées pendant l'année 2012 (courbe en bleu). Une sous-représentation des valeurs simulées entre 3 et 10 mm est clairement visible alors que les valeurs supérieures à 10 mm présentent une occurrence cohérente avec les observations. La cause probable provient de l'étape de traitement des transitions qui rappelle le consiste à « rogner » les bords des cellules. Cela a pour conséquence de diminuer les valeurs faibles situées en bordure alors que les valeurs fortes situées généralement au centre des cellules sont inchangées.

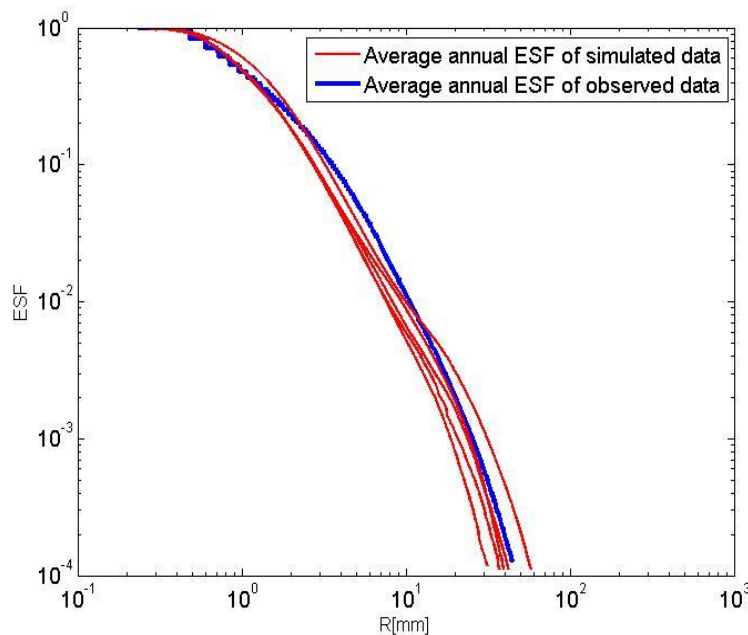


Figure 4.23. Fonction de survie empirique des 4 simulations conditionnelles (4 courbes rouge correspondant à chacune à la moyenne de 42000 cartes) et des observations de 2012 (en bleu)



### 4.6.3.3 Analyse statistiques des simulations avec support inconditionnel (synthétique)

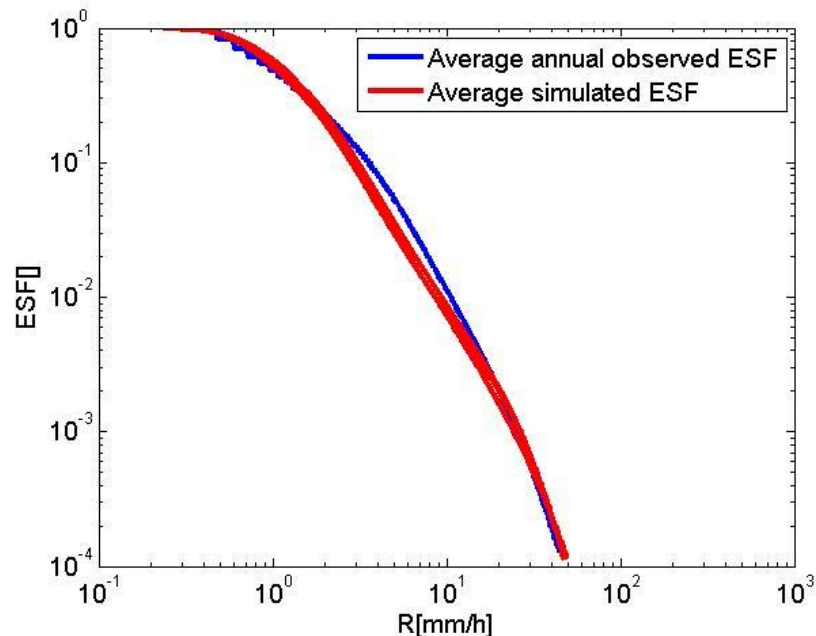
De la même manière que précédemment, mais en générant des supports synthétiques nous calculons les fonctions de survie des lames d'eau générées pour 4 réalisations (4 \* 1000) de données (Figure 4.24).

Afin d'avoir une cohérence dans la méthodologie de validation, j'ai sélectionné 1000 cartes parmi les 42000 cartes ayant en moyenne le pourcentage de pluie proche de pourcentage moyen des cartes simulées et nous avons tracé la courbe en bleu.

La Figure 4.24 montre la variabilité des fonctions de survie des quatre réalisations de données simulées ainsi que la fonction de survie de la sélection des données observées pendant l'année 2012 (courbe en bleu).

La même remarque concernant une sous-représentation des lames d'eau entre 3 et 10 mm peut être faites, la cause étant identique.

Il convient toutefois de noter que la variabilité des simulations est plus faible que dans le cas des simulations conditionnelles, cela pourrait être dû au nombre réduit de chaque réalisation.



*Figure 4.24. . Fonction de survie empirique 4 moyennes annuelles des simulations inconditionnelles en rouge (4 courbes rouge correspondant à chacune à la moyenne de 1000 cartes) et des observations de 2012 (en bleu)*

## 4.7 Conclusion

L'objectif de ce chapitre était de présenter l'étude d'un générateur 2D de lames d'eau basé sur des cartes de lames d'eau issues de radars météorologiques. D'un point de vue pratique nous nous sommes focalisés sur des cartes radar de Météo France centrées sur la région parisienne afin de préserver une cohérence régionale avec le simulateur 1D présenté au chapitre précédent et basé sur des observations collectées sur le site du SIRTA à Palaiseau. Le modèle présenté est une extension du modèle 1D, il s'articule sur 4 étapes principales : la simulation du support, la simulation de lames d'eau dans chaque cellule (simulation FIF), la re-normalisation et le traitement des transitions pluie / non pluie. L'analyse statistique d'un ensemble de réalisations montre que les propriétés d'invariance d'échelle sont correctement restituées sur toute la gamme d'échelle. La distribution des lames d'eau reste cohérente avec les observations bien que l'occurrence des pixels de lame d'eau compris entre 3 et 10 mm/h soit sensiblement inférieure à celle observée. Le processus de traitement des transitions semble être la source du problème. Une amélioration de ce traitement pourrait avantageusement être envisagée. Au final, la particularité de ce processus de génération de cartes de lames d'eau radar est qu'il restitue correctement un ensemble de propriétés statistiques, à savoir : la structure spatiale, l'invariance d'échelle et les différents régimes de scaling, la distribution des lames d'eau. La cohérence simultanée de l'ensemble de ces propriétés statistiques en fait un modèle original comparé aux modèles existant pour lesquels seule une partie des propriétés sont vérifiées (e.g. Schleiss et al. 2014) qui valident la structure spatiale mais avec des taux de pluie parfois aberrant et aucune information sur les régimes de scaling ou (Gires et al., 2014) qui proposent un modèle qui génère des zéros mais avec des régimes de scaling qui ne sont pas réalistes).

Pour clore ce chapitre il convient de discuter des diverses améliorations et évolutions possibles de ce simulateur.

Actuellement le générateur est calé pour simuler des cartes à la résolution de  $1 \times 1 \text{ km}^2$  en région tempérée. Pour cela, 11 paramètres sont nécessaires, parmi ces paramètres certains sont dépendants de la région climatique, d'autres non. Il est clair que les 4 paramètres pour la génération du support dépendent de la région considérée (variogramme, pourcentage de pluie). De même la relation lame d'eau moyenne d'une cellule / longueur équivalente est susceptible de dépendre de la région considérée. Seuls les paramètres multifractals peuvent à priori être considérés comme constants.

La prise en compte des fines échelles a permis d'améliorer la simulation de la pluie 1D. Une réflexion reste à mener concernant la génération de carte à une résolution plus fine que celle à laquelle celui-ci a été développé (1km). Quels sont les paramètres susceptibles de changer ? Il est tentant de penser que certains des paramètres pourraient demeurer identiques. Nous avons vu que les paramètres multifractals sont constants, il convient donc d'apprécier l'évolution des paramètres du variogramme et ceux relatifs à la dé-normalisation. Concernant les premiers on ne s'attend pas à voir un changement important de l'allure du variogramme pour des résolutions plus fines, quant aux seconds ils relient la hauteur moyenne de la lame d'eau de la cellule à sa surface (longueur équivalente). Bien évidemment cette dernière est dépendante de la résolution. Il conviendrait donc de disposer d'un nouveau jeu d'observation à une résolution plus fine comme par exemple les observations à 250 m des radars de Météo France. L'utilisation de données observées à une meilleure résolution garantirait de plus, une meilleure caractérisation du support, la frontière pluie/non pluie étant ainsi décrite de manière plus fine.

Le simulateur proposé permet de générer des cartes indépendantes les unes des autres représentatives d'une situation donnée. L'étape suivante consisterait à développer un simulateur 3D (2D spatial + 1D temporel) afin de générer des cartes temporellement dépendantes. Cette étape est loin d'être triviale. Plusieurs pistes peuvent être envisagées par exemple l'extension en 3D du modèle FIF. D'un point de vue purement mathématique cette approche ne pose pas de réel problème, un paramètre supplémentaire est toutefois nécessaire pour la paramétrisation du noyau de convolution du générateur  $g_\alpha(\vec{x})$  (cf. chapitre 2 paragraphe 2.2.1). Cependant ici encore se pose le problème de la représentation des zéros non pris en compte par cette modélisation et son corollaire relatif à la modélisation du support.

# Chapitre 5

## APPLICATION A L'ESTIMATION DE CARTES DE PRECIPITATIONS A PARTIR DES LIAISONS MICRO-ONDES

### Sommaire

---

5	APPLICATION A L'ESTIMATION DE CARTES DE PRECIPITATIONS A PARTIR DE LIAISONS MICRO-ONDES .....	149
5.1	Introduction : .....	149
5.2	Les données .....	152
5.2.1	Principe de simulation d'une liaison .....	153
5.2.2	Configurations retenues .....	156
5.3	Méthodologie pour l'estimation des champs précipitants .....	159
5.3.1	Assimilation variationnelle .....	159
5.3.1.1	3D-Var.....	160
5.3.1.2	4D-Var.....	160
5.3.2	Hypothèses retenues.....	162
5.3.2.1	Modèle advection pure .....	162
5.3.2.2	Vents.....	164
5.3.2.3	Bruit de mesure .....	164
5.3.2.4	Résolution temporelle.....	164
5.4	Résultats.....	164
5.4.1	Influence du nombre et de la position des antennes et de la direction du vent.....	165
5.4.2	Influence du bruit de mesure sur la qualité de la restitution.....	169
5.4.3	Echelles temporelles d'observation.....	171
5.4.4	Résultats en fonction de taux précipitants .....	174
5.5	Conclusion .....	177

---



## **5 APPLICATION A L'ESTIMATION DE CARTES DE PRECIPITATIONS A PARTIR DE LIAISONS MICRO-ONDES**

### **5.1 Introduction :**

La mesure des précipitations (intensité, taux de pluie, lame d'eau) est réalisée à l'échelle du globe essentiellement par des réseaux de pluviomètres et de radars météorologiques. Cependant ces derniers ont l'inconvénient de n'effectuer que des mesures ponctuelles. Il est donc intéressant de déployer des réseaux denses pour rendre compte de manière satisfaisante la variabilité de la pluie. Ces dernières années, on observe que le nombre de pluviomètres diminue de façon importante un peu partout dans le monde. D'après Overeem et al. (2013) il apparait un déclin de 50% du nombre de pluviomètres dans la période 1989-2006 pour le réseau GPCC en Afrique et une baisse de 50% du nombre de relevés journaliers valides dans la période 2000-2007 en Asie.

Les observations provenant des systèmes de télédétection spatiale ne présentent pas pour l'instant des caractéristiques compatibles avec la plupart des applications concernant les précipitations. En effet, les valeurs restituées le sont avec une précision qui est relativement faible, une résolution spatiale qui est importante (allant de 25 km<sup>2</sup> à 5km<sup>2</sup>) et surtout leur répétitivité temporelle (de l'ordre de 3 à 6 passages par jours sur une zone) ne permet pas la caractérisation des précipitations instantanées.

Les réseaux de radars météorologiques peuvent du point de vue de la qualité des observations fournies être une alternative pour de nombreuses applications. Dans la perspective, d'une couverture globale, ils présentent cependant deux inconvénients majeurs : leur utilisation est limitée par les reliefs et leur coût d'installation et de maintenance est important. Il en résulte que peu de pays disposent d'une couverture conséquente sur l'ensemble de leur territoire et se limitent bien souvent à couvrir les régions les plus densément peuplées (capitale, grandes villes).

Des moyens d'observation alternatifs restent à développer. Les caractéristiques de ces nouveaux systèmes d'observation devraient prendre en compte non seulement le coût de l'équipement et de sa maintenance mais devraient permettre une vision moins ponctuelle de la pluie qu'un pluviomètre. Depuis l'avènement des systèmes radios dans le domaine des

télécommunications, notamment pour les faisceaux hertziens puis pour les liaisons avec les satellites de télécommunications, les opérateurs se sont heurtés au problème de l'atténuation des ondes liée à la propagation en période de pluie. Ceci a été d'autant plus vrai que les bandes de fréquences allouées sont de plus en plus élevées. De nombreuses études ainsi que des expériences sur le sujet ont foisonné depuis le début des années 90, on peut citer par exemple les expériences de propagation entre des terminaux au sol et des satellites géostationnaires telles que Italsat (1991-2001) (Polonia et al., 1998) aux fréquences 20, 40 et 50 GHz ou Olympus-1 (1989-1993) (Arbesser et Brussaard, 1993) aux fréquences 20 et 30 GHz. Depuis quelques années a surgi l'idée de mettre à profit l'atténuation des liaisons radios utilisées par les opérateurs pour la mesure de la pluie. Ce concept est notamment à l'étude en France, en Italie, au Pays Bas et en Israël. La plupart de ces travaux sont basés sur la mesure de l'atténuation issue des réseaux capillaires pour la téléphonie mobile (Giuli. et al., 1991 ; 1999; Zinevich et al., 2008 ; 2009 ; Overeem et al.,2013). Ces réseaux sont composés d'un ensemble relativement dense de faisceaux hertziens assurant des liaisons de quelques centaines de mètres à une dizaine de kilomètres entre les stations de base. Les fréquences utilisées, comprises entre 6 et 40 GHz, sont sensibles aux précipitations (tout au moins celles supérieures à une dizaine de GHz). Dans de nombreuses régions du globe il n'est pas possible pour le moment d'obtenir les données nécessaires à la restitution des champs précipitants auprès des opérateurs de téléphonie mobile. L'alternative, envisagée dans le cadre de cette thèse, consiste, plutôt que d'utiliser les stations des réseaux mobiles, de mettre à profit les émissions en provenance des satellites de télécommunications ou de TV-Sat (Maitra et Chakravarty, 2005; Adhikari et al., 2011). Durant des travaux antérieurs, Barthès et Mallet (2013) ont montré que l'utilisation des ondes émises par les satellites commerciaux TV-SAT dans la bande 10,7-12,7 GHz (bande Ku) permettent d'estimer l'intensité de pluie moyenne le long de la partie de la liaison située dans la troposphère. Ainsi, l'utilisation opportuniste des signaux émis par les satellites commerciaux pourrait permettre de développer un système d'observation à bas coût. Par la suite ce type de mesure a été étendu à plusieurs liaisons en provenance de différents satellites afin de sonder simultanément différentes directions. Un algorithme a été développé afin de restituer un champ de pluie à partir des différentes liaisons. Cet algorithme de restitution a été développé et testé dans un premier temps sur des données simulées puis sur des données réelles issues d'une campagne de mesure réalisée en Ardèche dans le cadre du projet européen HyMex (HYdrological cycle in Mediterranean EXperiment) durant

l'année 2013. Les lames d'eau restituées ont été comparées à celles d'un réseau de pluviomètres HPiconet colocalisé (Mercier et al, 2015). Les résultats obtenus sont assez satisfaisants, le modèle reproduisant bien les variations de la pluviométrie le long de l'axe est/ouest. Les auteurs ont néanmoins constaté quelques difficultés à reproduire les variations de pluviométrie selon l'axe nord/sud. Ceci s'explique par la configuration géométrique du système (les antennes étant toutes dirigées dans un secteur Sud-Est, Sud-Ouest) et l'utilisation d'un seul récepteur visant seulement 4 satellites.

La méthode de restitution mise en œuvre pour obtenir un champ précipitant à partir des affaiblissements atmosphériques observés en bande Ku est une méthode d'assimilation de type variationnelle (4D Var). Elle utilise un modèle direct pour confronter les sorties du modèle avec les observations afin de déterminer (ou de corriger) les variables physiques qui les ont inférées. Un modèle direct simple d'advection décrivant le déplacement de cellules de pluie induit par le champ de vitesse du vent est utilisé pour décrire le déplacement horizontal. En utilisant ce modèle direct, la technique de l'assimilation de données permet de relier entre eux les enregistrements successifs et améliore ainsi le champ de pluie reconstitué en le forçant à présenter une cohérence spatiale et temporelle.

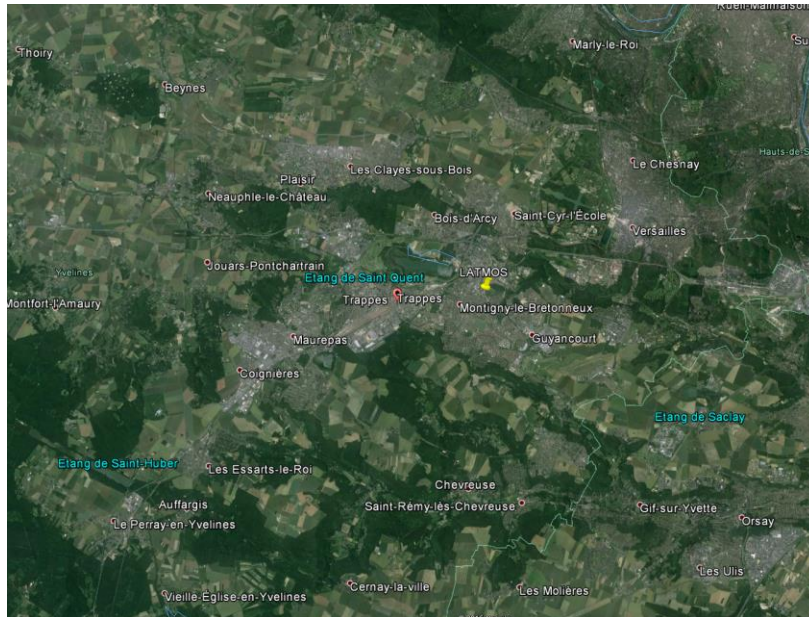
L'objectif de ce chapitre est l'étude d'un système de mesure opportuniste innovant pour améliorer la description spatiale des précipitations en milieu urbain ; autrement dit étudier la faisabilité de mesurer la pluie en utilisant les affaiblissements des liaisons satellite-TV à partir d'un certain nombre de récepteurs. La présence de postes de télévision reliés à des antennes paraboliques un peu partout dans le monde assure que ces mesures seront techniquement possibles y compris dans les zones où il n'existe pas de radars météorologiques.

L'hypothèse de départ consiste à supposer une ville dans laquelle un certain nombre d'utilisateurs (quelques dizaines) possède un petit terminal branché en dérivation sur le câble coaxial de leur antenne TV-SAT. Ce terminal mesurera en temps réel l'affaiblissement dû à la pluie et le transmettra via internet à un serveur de calcul. Nous allons donc réaliser une étude sur des données simulées. Il s'agit donc de savoir dans quelle mesure il est possible de reconstituer les champs de pluie à partir de l'ensemble des affaiblissements potentiellement mesurés par ces terminaux distribués sur une zone géographique donnée. Pour se faire nous allons envisager un nombre restreint de configurations (nombre de liaisons satellites, nombre de récepteurs, direction du vent, taille/intensité des cellules pluvieuses) et étudier les champs restitués.



## 5.2 Les données

L'étude est réalisée sur des données simulées car nous ne disposons pas à ce stade de données réelles contrairement à l'étude réalisée précédemment (Mercier et al. 2015) dans laquelle il n'y avait qu'un seul récepteur. Nous avons choisi comme région d'étude, la même que celle étudiée au chapitre précédent relative à la simulation de champ de pluie (Ile de France). Ce choix nous permet ainsi d'utiliser notre simulateur 2D pour simuler des cartes de précipitation de taille  $30 \times 30 \text{ km}^2$  (Figure 5.1).



*Figure 5.1 image satellitaire de la région d'étude*

Parmi l'ensemble des cartes simulées, nous avons sélectionné les trois cartes ci-dessous (dénommées carte 1, 2 & 3 dans la suite). Ces trois cartes possèdent des caractéristiques différentes en termes d'intensité de pluie et de surface pluvieuse.

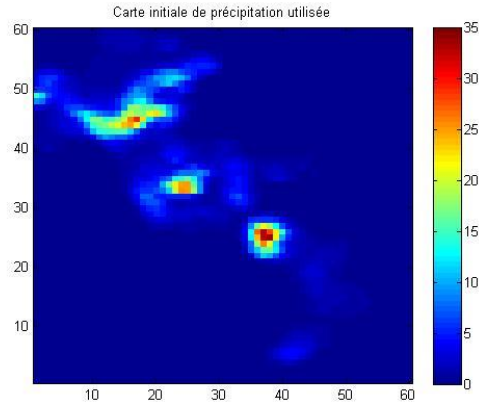


Figure 5.2 Carte 1(simulée avec le modèle 2D utilisé pour toutes les configurations testées dans ce chapitre.

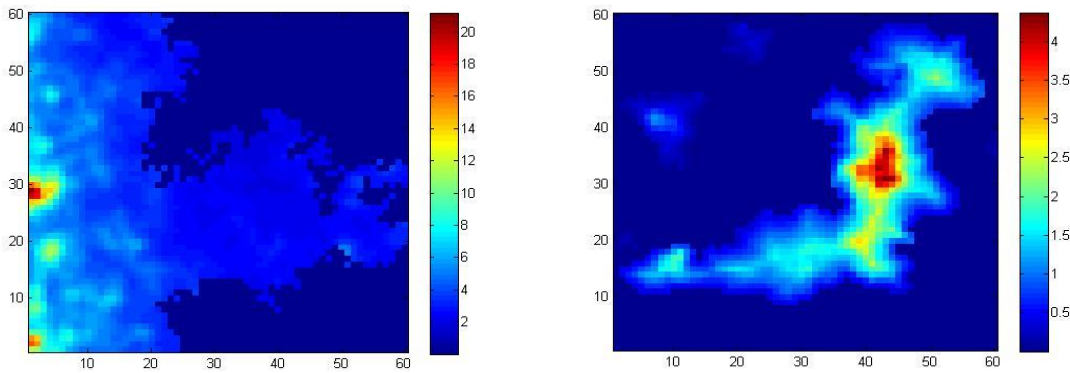


Figure 5.3. Cartes 2 & 3 simulées avec le modèle 2D utilisées pour la configuration « optimale ».

Le Tableau 5.1 fournit certaines caractéristiques des trois cartes simulées.

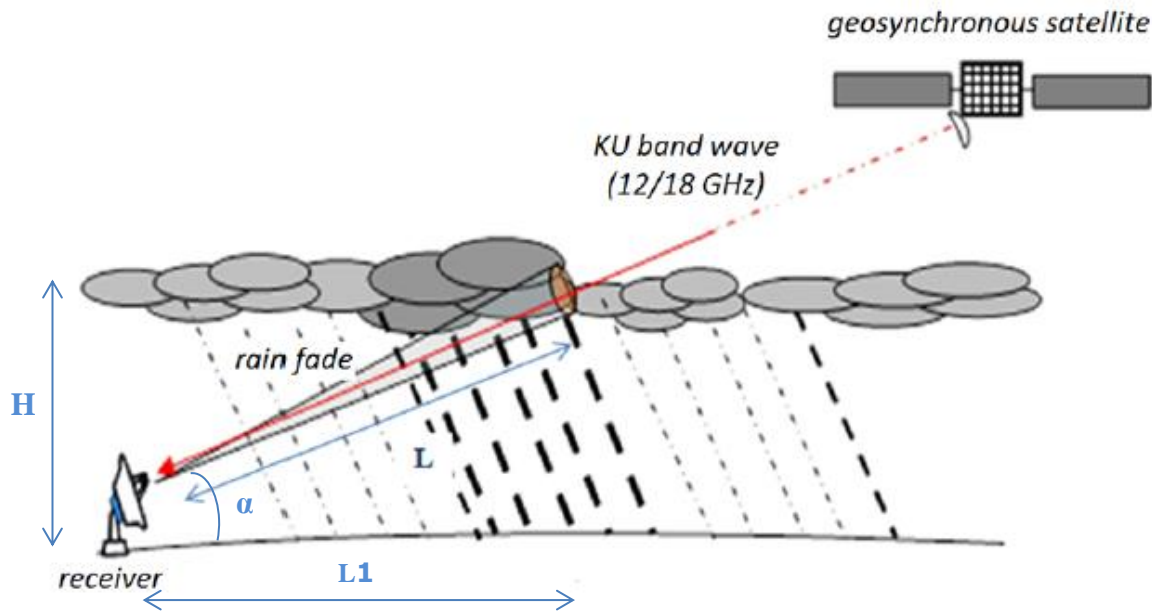
	Surface pluvieuse (pourcentage)	RR maximum (mm/h)	RR moyen de toute la carte (mm/h)	RR moyen de de la surface pluvieuse (mm/h)
Carte 1	37%	33.78	1.12	3.09
Carte 2	57.89%	21.17	2.26	3.92
Carte 3	34%	4.38	0.38	1.10

Tableau 5.1 Caractéristiques des cartes simulées utilisées.

### 5.2.1 Principe de simulation d'une liaison

Le principe de ce système de mesure de l'intensité des précipitations est d'utiliser les micro-ondes se propageant depuis des satellites jusqu'à des antennes au sol. En effet, les satellites géostationnaires de télécom ou de télévision émettent vers le sol des ondes électromagnétiques qui traversent la troposphère avant d'être captées par des antennes. La bande de fréquence utilisée ici se situe en bande Ku et plus précisément entre 10.7 et 12.7GHz. Ces ondes sont affaiblies

lorsqu'elles traversent des zones pluvieuses pour deux raisons principales. D'une part il existe un phénomène d'absorption par effet Joule d'une partie de l'énergie de l'onde électromagnétique et d'autre part, une partie de l'énergie est redistribuée dans toutes les directions par diffusion sur les hydrométéores (diffusion de Mie). Notons par ailleurs que la glace (présente dans les nuages au-dessus de l'isotherme  $0^{\circ}$ ) ne présente pas ce phénomène pour la gamme de fréquence considérée et contribue peu à l'affaiblissement de l'onde. Toutefois, il convient de signaler que la couche de fonte contribue un peu à l'affaiblissement du fait de la présence d'une pellicule d'eau liquide enrobant les hydrométéores glacés. Cet effet est partiellement géré par l'algorithme de traitement qui fait appel à un recallage dynamique du signal satellite. Cependant, comme signalé dans Barthes et Mallet (2013) cette couche bien que pouvant apporter une atténuation linéique importante est peu épaisse, sa contribution à ces fréquences-là reste donc modeste dans la plupart des situations. La Figure 5.4 présente un schéma général du dispositif.



**Figure 5.4 Principe du dispositif de mesure des affaiblissements des ondes de la bande KU**

On suppose ici que l'isotherme  $0^{\circ}\text{C}$  se trouve à l'altitude  $H$ . L'angle par rapport au sol de la liaison récepteur / satellite est noté  $\alpha$ .

Lors de la campagne de mesure HYMEX, une antenne multi focus équipée de 4 têtes de réception permettait de recevoir 4 satellites. Celle-ci était reliée à un analyseur de champ permettant ainsi

de mesurer le niveau reçu de façon séquentielle sur les 4 têtes de réception. On disposait donc d'une mesure d'affaiblissement à haute résolution temporelle (10 secondes) sur 4 liaisons satellites. Pour plus de détail, le lecteur peut consulter l'article Mercier et al., (2015).

Pour une couche de pluie homogène, on peut relier l'affaiblissement linéique  $K$  du signal (en dB/km), qui correspond à l'affaiblissement de l'onde par unité de distance dû à la pluie, à l'intensité des précipitations  $RR$  par la relation :

$$K = a RR^b \tag{Équation 5.1}$$

$RR$  est le taux de pluie en mm/h tandis que  $a$  et  $b$  sont des coefficients dépendant de la fréquence, de la polarisation et de la distribution du diamètre des gouttes (ITU-R P.838)<sup>4</sup>, (Jameson, 1991 ; Leijnse et al., 2008).

Nous avons utilisé ce modèle pour simuler les affaiblissements à partir des cartes de pluie simulées. Pour cela nous avons estimé pour chaque pixel de pluie l'affaiblissement linéique correspondant via l'Équation 5.1 avec les paramètres  $a$  et  $b$  préconisés par l'Union Internationale des Télécommunications (ITU). La contribution de chaque pixel est alors obtenue par multiplication de son affaiblissement linéique par la longueur de liaison le traversant. L'affaiblissement total est ensuite obtenu en sommant l'ensemble des contributions de chaque pixel faisant partie de la liaison qui est considérée (Figure 5.5 et Équation 5.2).

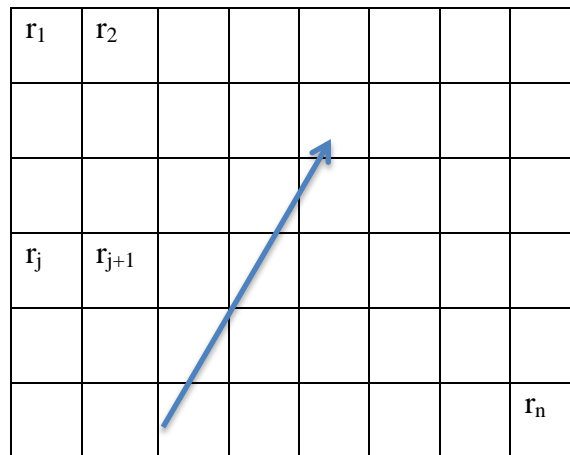


Figure 5.5 Grille de pixels d'une carte de lames d'eau avec un exemple de liaison présenté en bleu: les taux de pluie des pixels sont notés de  $r_1$  à  $r_n$

<sup>4</sup> <https://www.itu.int/rec/R-REC-P.838/fr>

$$A_{liaison} = \sum_{j \in liaison} l_j a r_j^b \quad \text{Équation 5.2}$$

Avec  $l_j$  longueur de la liaison dans le pixel  $j$

Des exemples de séries temporelles d'affaiblissement sont donnés par la Figure 5.7

### 5.2.2 Configurations retenues

L'objectif de l'étude ci-après ne prétend pas être exhaustive, il s'agit simplement à ce stade d'appréhender les performances d'un tel système à partir de simulations suffisamment réalistes. Pour cela, définissons dans un premiers temps le nombre de pixels  $N_{pix}$  contenant des antennes. Nous avons ensuite estimé visuellement la densité d'habitations pour les différentes zones de la carte présentée Figure 5.1. A partir de la densité d'habitations nous avons tiré aléatoirement un nombre  $N_{pix}$  de pixels. Ainsi plus un pixel est dans une zone dense, plus il a de chance d'être tiré. Nous avons enfin tiré pour chacun des  $N_{pix}$  pixels le nombre de liaisons satellites présentes dans le pixel. En région Ile de France il est possible de recevoir plusieurs satellites TV-SAT. Nous avons retenu les 7 satellites présentés dans le tableau ci-après. L'algorithme de génération d'une configuration est le suivant :

- Tirer aléatoirement le nombre de pixel  $N_{pix}$  contenant des récepteurs
- Choisir aléatoirement  $N_{pix}$  pixels en fonction de la densité urbaine
- Pour chaque pixel tirer aléatoirement  $N$  fois un numéro de satellite (Tableau 5.2, colonne 3) (nombre de récepteur par pixel)
- Supprimer les satellites en doublons

Ainsi on dispose au final de  $N_{pix}$  pixels contenant chacun entre 1 et 7 liaisons satellites.

Satellites	Position orbitale	Numéro
NSS7	20.0°W	1
Eutelsat 12 West A	12.5°W	2
Eutelsat 5 West A	5.0°W	3
Intelsat 10-02	0.8°W	4
Eutelsat 9A	9.0°E	5
Hot bird	13.0°E	6
Astra	19.2°E	7

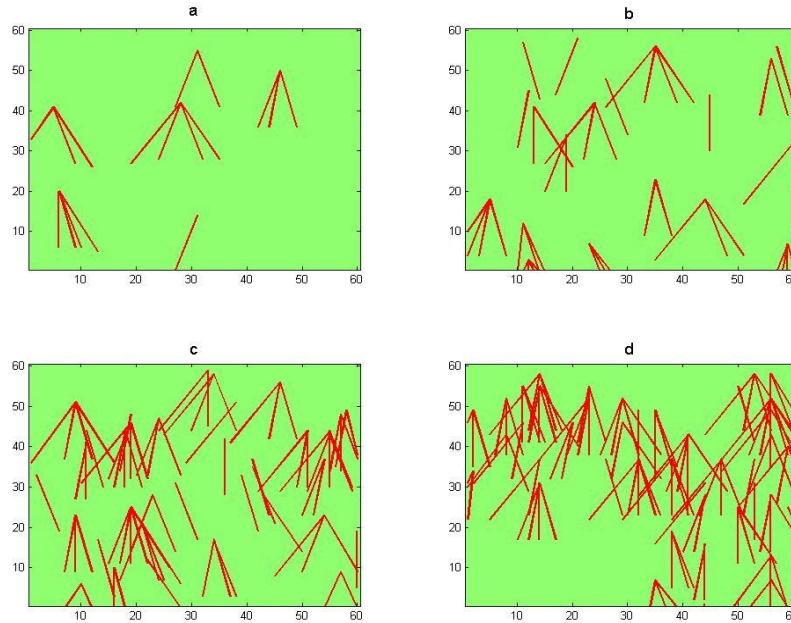
*Tableau 5.2 Liste des satellites géostationnaire utilisés dans notre étude et position orbitale*

Au final parmi les 8 configurations que nous avons testées nous présenterons nos résultats sur 4 configurations. Le Tableau 5.3 fourni pour ces 4 configurations le nombre de pixels contenant des antennes de réception ainsi que le nombre total de liaisons. Les 4 figures (Figure 5.6) schématisent les configurations retenues. La configuration 1 possède peu de liaisons, il y a par conséquent de nombreuses zones non couvertes alors que la configuration 4 est relativement dense excepté dans la partie Sud-Ouest moins urbanisée.

Le Tableau 5.3 présente le nombre de pixels contenant des antennes et le nombre de liaisons utilisées pour chaque configuration.

	Nombre de pixels contenant des antennes ( $N_{pix}$ )	Nombre total de liaisons
Configuration 1	6	17
Configuration 2	19	45
Configuration 3	36	79
Configuration 4	45	122

*Tableau 5.3 Résumé des quatre configurations sélectionnées.*



*Figure 5.6 Les quatre configurations étudiées dans ce chapitre.*

Un paramètre important à considérer est la vitesse de déplacement des cellules de pluie. Nous supposons ici que ces dernières sont uniquement advectées par le vent (horizontal), en d'autres termes nous faisons l'hypothèse dite de « l'atmosphère gelée ». Il est cependant bien connu qu'une telle hypothèse est loin d'être vérifiée sauf lorsque les échelles temporelle et spatiale sont relativement fines, ce qui est notre cas ( $< 1$  heure et  $< 30$  km). Par la suite, nous considérerons deux directions d'advection, un vent Ouest-Est et un vent Nord-Sud avec une vitesse de déplacement constante de 10 m/s.

En advectant les champs de pluie simulés il est possible d'obtenir les séries temporelles d'affaiblissement « vues » par les différents récepteurs. Ces séries représentent nos observations sensées être fournies par les différents récepteurs. C'est donc à partir de ces observations que nous allons devoir estimer le champ de pluie qui les a produites. En d'autres termes : quel est le champ de pluie initial qui, lorsqu'il traverse la zone urbaine que nous avons considérée, reproduit au mieux les observations obtenues sur les différents récepteurs ?

La Figure 5.7 montre un exemple de séries temporelles obtenues à partir de la configuration n°1 et du champ de pluie n°1. Le vitesse d'advection Ouest-Est est fixé à 10m/s. Nous obtenons pour chacune des 17 liaisons un série d'observations d'affaiblissements. Comme on pouvait s'y

attendre, certaines liaisons sont très affectées par le passage des cellules pluvieuses, c'est le cas des liaisons situées dans la partie au nord de la carte. Les liaisons les plus à l'Ouest subissent l'affaiblissement avant celles à l'est de la carte, le décalage temporel correspondant à la distance entre les liaisons

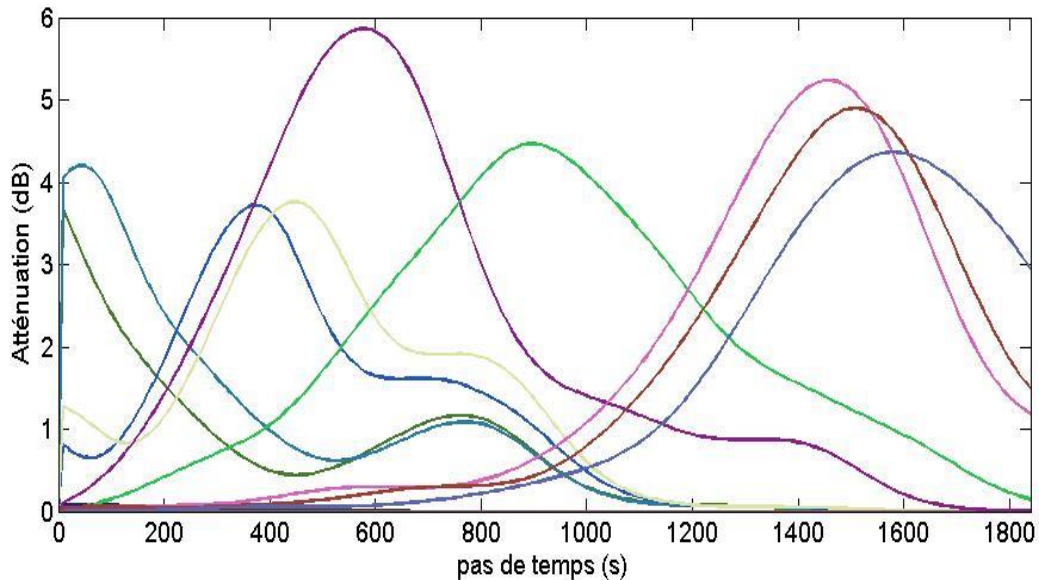


Figure 5.7 Exemple de série temporelles d'affaiblissement vues par les liaisons de la configuration 1 en advectant de l'est vers l'ouest le champ de pluie de la Figure 5.2 avec une vitesse =10m/s

## 5.3 Méthodologie pour l'estimation des champs précipitants

On cherche à reconstituer des champs de pluie à partir des observations en utilisant les principes de l'assimilation de données variationnelle. Cette dernière méthode consiste à intégrer les observations dans un modèle numérique (ici un modèle d'advection de cellules pluvieuses) et à minimiser une fonction de coût évaluant l'écart entre le champ propagé par le modèle et les observations disponibles.

### 5.3.1 Assimilation variationnelle

Dans cette partie, nous faisons quelques rappels concernant les méthodes d'assimilation variationnelle : 3D-var et 4D-var (utilisée ci-après). Le lecteur intéressé par plus de détails peut se référer au cours de Bocquet; (Bocquet, 2005) ou encore à (Daget, 2007)



### 5.3.1.1 3D-Var

Dans le cadre de l'assimilation 3D-var, on ne prend pas en compte la dimension temporelle. On cherche donc, à une date donnée, à minimiser une fonction de coût  $J$ , fonction prenant en compte les écarts entre l'état  $x$  du système et d'une part l'état d'ébauche  $x_b$  (obtenu généralement par intégration d'un modèle numérique depuis la date précédente) et d'autre part le vecteur des observations  $y$ . On peut écrire cette fonction :

$$J(x) = \frac{1}{2}(x - x_b)^T \mathbf{B}^{-1}(x - x_b) + \frac{1}{2}(y - \mathbf{H}x)^T \mathbf{R}^{-1}(y - \mathbf{H}x) \quad \text{Équation 5.3}$$

Où  $x$  est le vecteur contenant les paramètres à estimer (dans notre cas le champ de pluie);  $x_b$  est l'ébauche, c'est-à-dire l'estimation à priori de  $x$  ;

$y$  est le vecteur contenant les observations à la date considérée ;

$\mathbf{H}$  est l'opérateur de projection de l'espace des états du modèle sur celui des observations (ces dernières étant généralement beaucoup moins nombreuses) ;

$\mathbf{B}$  est la matrice de variance/covariance de l'erreur de l'ébauche et

$\mathbf{R}$  la matrice de variance/covariance de l'erreur des observations, ces deux dernières matrices permettant de pondérer plus ou moins les différents termes selon la confiance qu'on leur accorde.

On minimise ensuite la fonction  $J$  en annulant son gradient. Celui-ci vaut :

$$\nabla J(x) = \mathbf{B}^{-1}(x - x_b) - \mathbf{H}^T \mathbf{R}^{-1}(y - \mathbf{H}x) \quad \text{Équation 5.4}$$

Où  $\mathbf{H}^T$  est l'adjoint de l'opérateur linéarisé de l'opérateur  $\mathbf{H}$ .

Dans l'assimilation 3D-var, on propage via un modèle numérique  $\mathbf{M}$  le système de la date  $t-1$  à la date  $t$  (on obtient ainsi un état d'ébauche à la date  $t$  :  $x_b^t$ ) puis on corrige cet état d'ébauche grâce aux observations disponibles à la date  $t$  (on obtient ainsi un état analysé à la date  $t$  :  $x_a^t$ ) puis l'on recommence en propageant le vecteur  $x_a^t$  vers la date  $t+1$  en intégrant le modèle ...

### 5.3.1.2 4D-Var

L'assimilation 4D-var permet d'introduire le temps dans le processus d'assimilation. En effet, le modèle étant déterministe, le système est connu pour tout  $t$  dès lors que l'on connaît son état initial. Notre objectif est donc de contrôler l'état initial du modèle permettant de rendre compte

au mieux (Équation 5.4) des observations disponibles à chacun des pas de temps. La fonction de coût de l'assimilation 4D-var s'écrit ainsi :

$$J(x) = \frac{1}{2}(x - x_b)^T \mathbf{B}^{-1}(x - x_b) + \frac{1}{2} \sum_{i=0}^N (y_i - \mathbf{H}_i x_i)^T \mathbf{R}^{-1}(y_i - \mathbf{H}_i x_i) \quad \text{Équation 5.5}$$

$x$  est donc désormais l'état initial du champ de précipitations (variable pronostique du modèle). On note ici  $x_i$  l'état du système à la date  $t_i$ . Celui-ci n'est autre que l'état initial  $x$  intégré jusqu'à la date  $t_i$  par le modèle numérique.

En notant  $\mathbf{M}_i$  l'opérateur qui propage le vecteur  $x_{i-1}$  de  $t_{i-1}$  à  $t_i$  par le modèle et

$\mathbf{M}_i = \mathbf{M}_{0i} = \mathbf{M}_{01} \mathbf{M}_{12} \dots \mathbf{M}_{i-1i}$  l'opérateur propageant le vecteur  $x$  de  $t_0$  à  $t_i$  on peut réécrire la fonction de coût  $J$  :

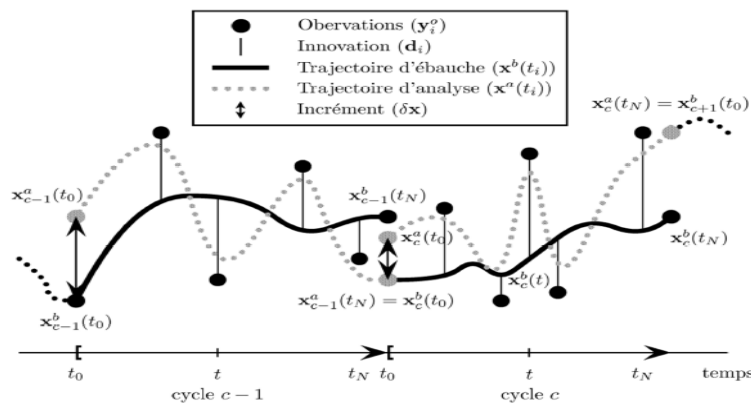
$$J(x) = \frac{1}{2}(x - x_b)^T \mathbf{B}^{-1}(x - x_b) + \frac{1}{2} \sum_{i=0}^N (y_i - \mathbf{H}_i \mathbf{M}_i x)^T \mathbf{R}^{-1}(y_i - \mathbf{H}_i \mathbf{M}_i x) \quad \text{Équation 5.6}$$

Notons que l'opérateur  $\mathbf{M}_i$  n'est a priori pas linéaire, ce qui va compliquer singulièrement les calculs dans la suite.

Le gradient de la fonction de coût  $J$  s'écrit :

$$\nabla j(x) = \mathbf{B}^{-1}(x - x_b) - \sum_{i=0}^N \mathbf{M}_i^T \mathbf{H}_i^T \mathbf{R}^{-1}(y_i - \mathbf{H}_i \mathbf{M}_i x) \quad \text{Équation 5.7}$$

Où  $\mathbf{M}_i^T$  est l'adjoint de l'opérateur linéarisé de l'opérateur  $\mathbf{M}_i$ .



**Figure 5.8** Schéma présentant le principe de l'assimilation 4D-Var. 2 cycles de minimisation sont représentés ici (cycles  $c-1$  et  $c$ ). Dans chaque cas, on a représenté l'état d'ébauche (trait noir), les observations disponibles (ronds noirs) et leurs écarts à l'ébauche (traits noirs verticaux), l'incrément (flèches), c'est-à-dire la modification de l'état initial apportée par le minimiseur et l'état analysé qui en découle (pointillés) (Daget, 2007).

L'adjoint est particulièrement délicat à calculer « à la main » explicitement, puisqu'il s'agit d'inverser le code du modèle numérique. Pour cela, nous avons utilisé le logiciel YAO qui nous facilite tous les calculs liés à l'assimilation. Pour plus de détails, voir (Thiria et al., 2006).

Le modèle d'assimilation utilisé est présenté par la Figure 5.9 (schéma extrait de l'article de Mercier et al. (2015))

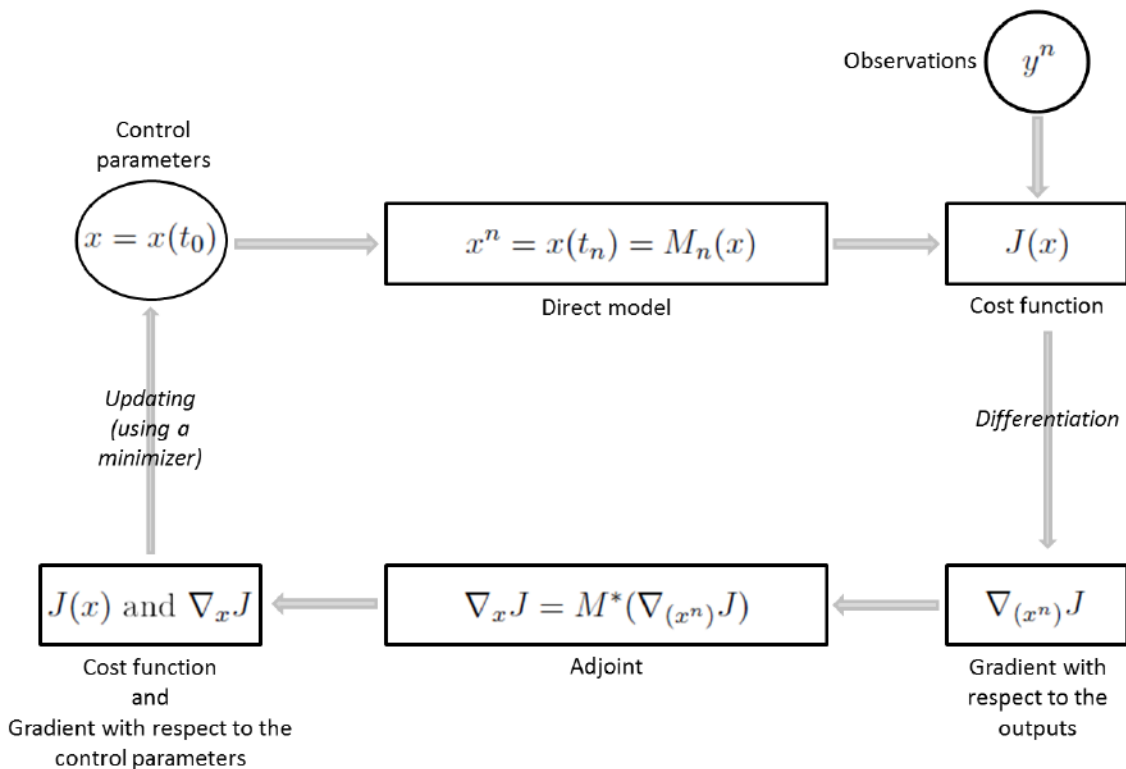


Figure 5.9. Schéma du modèle (figure extraite de l'article de (Mercier et al., 2015))

## 5.3.2 Hypothèses retenues

### 5.3.2.1 Modèle advection pure

Le choix d'un schéma numérique de résolution de l'équation d'advection. Rappelons tout d'abord que celle-ci est définie, en dimension 1, par :

$$\frac{\partial RR}{\partial t} + V \frac{\partial RR}{\partial x} = 0$$

Où  $RR$  désigne le champ inconnu (l'intensité pluvieuse mesurée en  $mm/h$  ou encore dans notre cas les affaiblissements linéiques),  $V$  désigne la vitesse d'advection,  $t$  le temps et  $x$  la position.

Comme vu précédemment, on modélise des champs de pluie de courte durée de vie, et donc supposés non déformés. Il s'agit donc d'utiliser un schéma d'advection limitant au maximum la diffusion numérique. Par ailleurs, nos champs de pluie présentent de très forts gradients en espace. Dans ce cas, la diffusion numérique est à éviter au maximum pour que notre champ ne soit pas lissé en quelques itérations seulement. Ces gradients très forts sont cependant susceptibles de poser des problèmes de stabilité.

Plusieurs schémas d'advection existent dans la littérature listons par exemple schéma décentré amont, schéma de Lax-Wendroff. Dans le cas de notre étude, nous avons retenu le schéma de Smolarkiewicz (Smolarkiewicz, 1983) qui a été testé et utilisé par (Mercier et al., 2015) dans un cadre similaire.

Les équations présentant ce schéma sont présentées ci-dessous :

$$A_i^{n+1} = A_i^* - [F(A_i^*, A_{i+1}^*, \tilde{V}) - F(A_{i-1}^*, A_i^*, \tilde{V})]$$

$$A_i^* = A_i^n - [F(A_i^n, A_{i+1}^n, V) - F(A_{i-1}^n, A_i^n, V)]$$

$$\tilde{V} = \frac{(|V|\Delta x - \Delta t V^2)(A_{i+1}^* - A_i^*)}{(A_i^* + A_{i+1}^* + \varepsilon)\Delta x}$$

$$F(A_i, A_{i+1}, V) = \frac{\Delta t}{2\Delta x} [(V + |V|)A_i + (V - |V|)A_{i+1}]$$

Il s'agit d'un schéma conçu pour l'équation d'advection et plus particulièrement pour la modélisation atmosphérique (Smolarkiewicz, 1983). De manière générale, les schémas d'ordre 1 étant trop diffusifs et les schémas d'ordre 2 trop instables dès qu'il y a des gradients un peu trop

forts, on utilise ici un schéma d'ordre 1 en première étape (calcul du  $A_i^*$ ) afin d'assurer la stabilité, puis on retire à l'étape 2 (calcul du  $A_i^{n+1}$ ) la diffusion induite par la 1<sup>ère</sup> étape.

### 5.3.2.2 Vents

L'objectif ici est une restitution en temps réel des champs précipitants notamment pour des applications telles que la prévision des crues en milieu urbain. De ce fait, les conditions d'advection ne sont pas connues et devront in fine être contrôlées par le modèle. Dans un premier temps cependant pour des raisons de simplicité, nous avons considéré que la vitesse et la direction du vent sont connues.

### 5.3.2.3 Bruit de mesure

En pratique, toute mesure est entachée d'erreur. Nous étudierons donc l'influence d'une erreur de mesure sur la qualité de la restitution des champs. Pour cela nous réaliserons des assimilations avec des observations non bruitées puis avec un bruit additif gaussien d'écart type 0.5 dB et 1dB.

### 5.3.2.4 Résolution temporelle

Pour chaque simulation, on sauve les séries temporelles qu'enregistrent les capteurs avec une période d'échantillonnage de 10 secondes. Cette période d'échantillonnage correspond à celle utilisée durant la campagne de mesure HYMEX. Toutefois, il n'y a aucune raison que les caractéristiques des capteurs installés chez les particuliers soient identiques. Nous avons donc choisi 3 périodes d'échantillonnage : 10 s, 1 minute et 5 minutes afin d'estimer l'influence de la période d'échantillonnage sur la qualité de la restitution.

## 5.4 Résultats

Afin d'évaluer les performances du modèle d'assimilation, nous présenterons dans la suite les cartes obtenues après assimilation pour diverses configurations ainsi que leur écart avec la carte d'origine. Nous avons également calculé l'écart quadratique moyen entre les pixels estimés et les valeurs « vraies » (RMSE) et le taux précipitant moyen pour les diverses configurations sur l'ensemble de la carte, sur les zones pluvieuses de la carte d'origine ainsi que sur les zones sèches

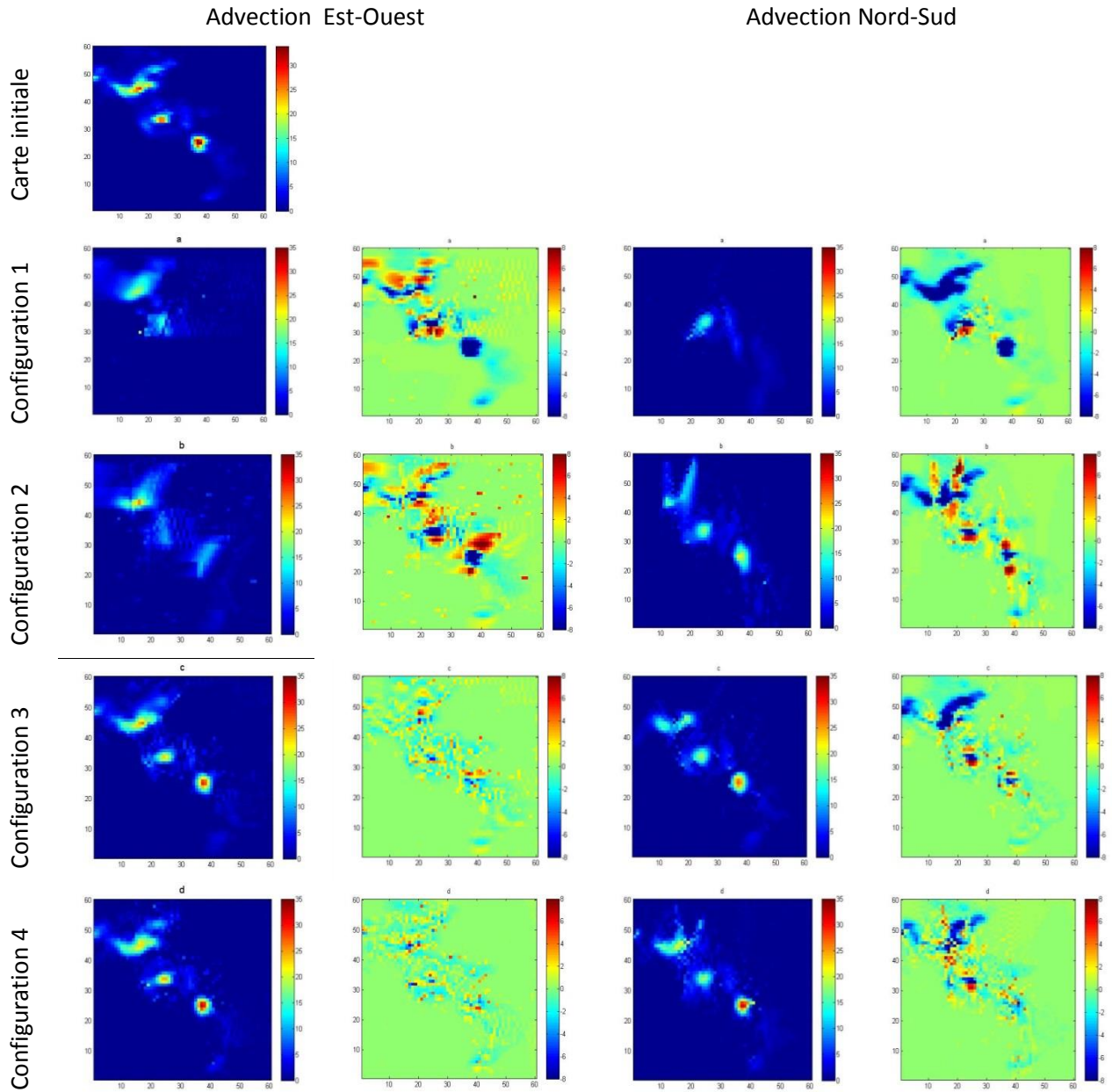
de la carte d'origine. Le but étant de vérifier dans quelle mesure il a plu « au bon endroit » et quelle quantité.

#### ***5.4.1 Influence du nombre et de la position des antennes et de la direction du vent***

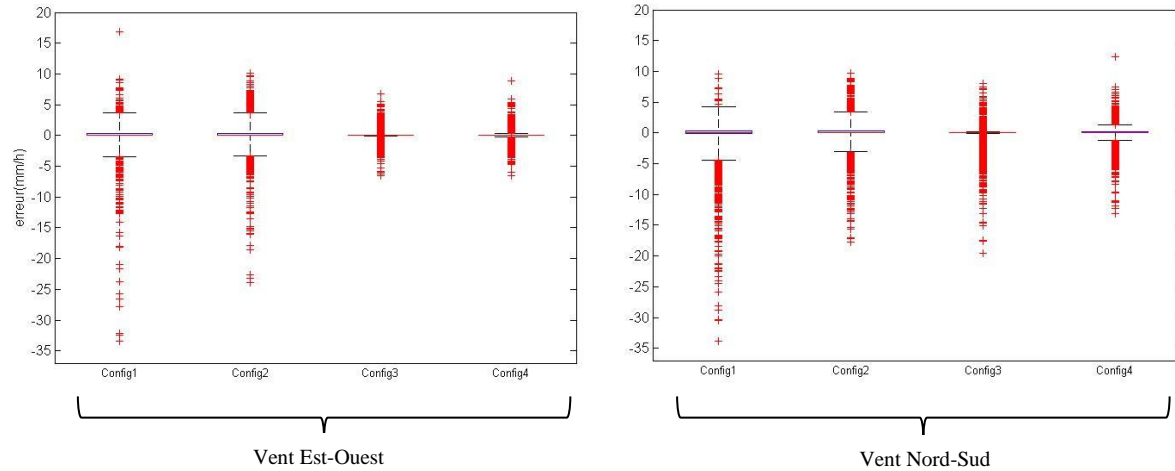
Nous avons considéré la carte de précipitation n°1 et les 4 configurations géométriques (Figure 5.6). Nous avons également considéré des vents Est – Ouest puis Nord – Sud. Après assimilation, nous avons comparé les cartes obtenues par assimilation des observations d'affaiblissement avec la carte initiale. Les résultats sont présentés à la Figure 5.10. La colonne 1 permet de comparer le champ initial (Carte n°1) (ligne 1) avec les champs obtenus pour les 4 configurations géométriques (lignes 2 à 5) et pour une advection Est –Ouest de 10 m/s. La colonne 2 montre l'écart entre la carte initiale (Carte n°1) et les cartes obtenues. Nous pouvons constater que l'erreur est la plus forte pour les deux configurations 1 et 2, les pixels en bleu foncé montrent que la pluie n'est pas restituée à cet endroit tandis que les pixels en rouge indiquent la présence de fausses détections de la pluie.

Il apparait que seules les configurations géométriques 3 & 4 permettent de restituer les 3 cellules de pluie initialement présentes. Ce résultat est logique dans la mesure où aucune liaison ne traverse à un instant donné la cellule située le plus au Sud. De même, l'erreur sur chaque pixel tend à décroître lorsque le nombre de liaison augmente (colonnes 2). Concernant une advection Nord – Sud, la configuration géométrique n°1 ne permet de retrouver qu'une seule cellule. La configuration géométrique n°2 fait apparaître les 3 cellules mais avec des erreurs importantes en termes de taux de pluie. Les configurations géométriques 3 et 4 permettent comme précédemment d'obtenir de meilleures restitutions mais tout en étant un peu moins bon que dans le cas de vent Est- Ouest. Ces résultats vont dans le même sens que l'étude de Mercier et al., (2015) toutefois il semblerait que cet effet soit atténué compte tenu du nombre plus élevé de capteurs.

La Figure 5.11 (gauche) illustre les boîtes à moustache de l'erreur pour les quatre configurations avec une advection Est-Ouest, et (droite) pour l'advection Nord-Sud. Elles confirment que la carte est mieux restituée dans le cas d'un vent est-ouest, notamment pour les configurations 3 et 4.



*Figure 5.10 Résultats de l'assimilation lorsque la carte 1 est utilisée comme champ de pluie initial pour les 4 configurations géométriques et pour un vent Est – Ouest (colonne 1). Ecart pixel à pixel entre la carte 1 et les cartes obtenues (colonne 2). Colonnes 3 & 4 : Idem aux colonnes 1 et 2 mais pour un vent Nord - Sud*



**Figure 5.11. Boîtes à moustaches illustrant la différence entre la carte initiale et les résultats d'assimilation pour les 4 configurations avec le vent Est-Ouest et le vent Nord-Sud.**

Les remarques précédentes peuvent être confirmées par les résultats présentés dans le Tableau 5.4. Pour les configurations 1 et 2, on obtient les plus fortes valeurs du RMSE qui sont de l'ordre de 0.8, 0.9 mm/h soit le double de celles des configurations 3 et 4. Concernant l'intensité de pluie moyenne, elle est sous-estimée pour les deux premières configurations et à peu près bien restituée pour les configurations 3 et 4, et ce, que ce soit sur la totalité de la carte ou sur les zones de pluie.

		RMSE   Biais (mm/h)	RMSE   Biais des zones de pluie (mm/h)	RMSE   Biais des zones sèches (mm/h)	RR moyen de toute la carte (mm/h)	RR moyen sur les zones de pluie (mm/h)
<b>Carte initiale</b>					1.12	3.09
<b>Configuration 1</b>	Vent E-O	0.88   -0.09	1.95   -0.75	0.27   0.27	1.02	2.33
	Vent N-S	0.98   -0.63	2.41   -2.04	0.16   1.16	0.48	1.05
<b>Configuration 2</b>	Vent E-O	0.82   0.10	1.71   -0.24	0.30   0.30	1.23	2.85
	Vent N-S	0.81   -0.13	1.93   -0.69	0.18   0.18	0.98	2.40
<b>Configuration 3</b>	Vent E-O	0.40   0.01	0.99   -0.08	0.07   0.07	1.13	3.01
	Vent N-S	0.59   -0.23	1.51   -0.78	0.07   0.07	0.98	2.30
<b>Configuration 4</b>	Vent E-O	0.37   -0.01	0.93   -0.11	0.05   0.04	1.10	2.98
	Vent N-S	0.55   -0.07	1.34   -0.39	0.10   0.10	1.04	2.69

**Tableau 5.4 Grandeurs statistiques sur la qualité des champs restitués par assimilation**



De façon générale, pour les configurations 3 et 4, on remarque une bonne restitution des taux précipitants ainsi que de la forme des évènements et de leur emplacement y compris dans le cas d'un vent nord-sud

La Figure 5.12 illustre la RMSE des pixels de taux précipitants en fonction du nombre de liaisons passant dans le pixel. La première conclusion à tirer de cette figure, est que l'erreur est quasiment nulle sur un pixel quel que soit la configuration si le pixel est traversé par plus de 2 liaisons. La RMSE diminue en fonction du nombre de liaisons. La comparaison entre les 4 configurations géométriques montre que pour les configurations 3 et 4, la RMSE des pixels ayant 0 liaison est plus faible par rapport aux configurations 1 et 2. En d'autres termes l'estimation du taux précipitant dans les zones non couvertes par les liaisons est d'autant meilleure que le nombre total de liaisons est élevé.

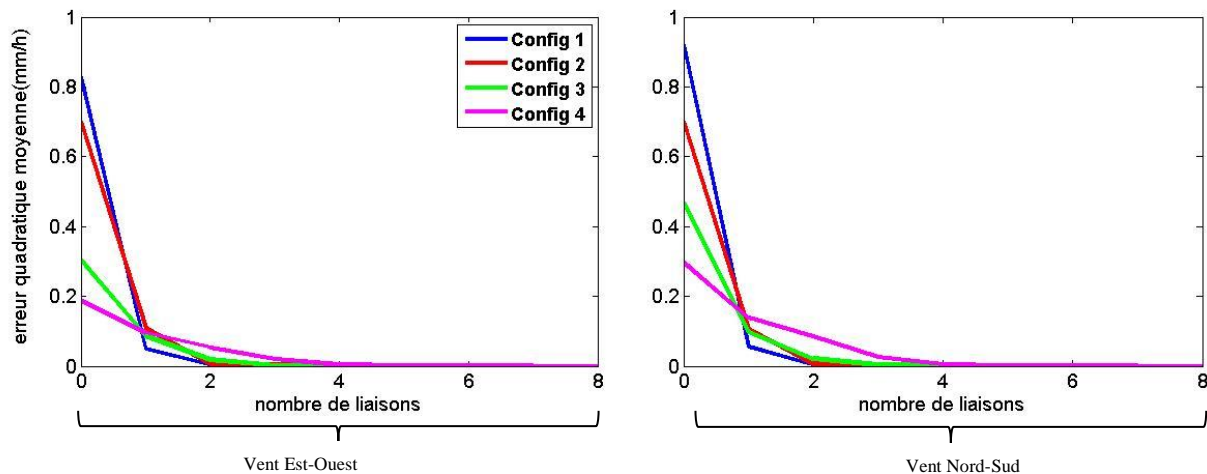


Figure 5.12. RMSE des pixels en fonction de nombre de liaisons qui les traversent.

Bien que loin d'être exhaustive, des premières conclusions peuvent être faites à partir de l'ensemble des résultats présentés ci-dessus. Ces derniers montrent que le nombre de liaisons est primordial afin de restituer correctement le champ initial, et qu'une couverture suffisamment homogène est préférable si on désire des performances correctes quel que soit la direction du vent. Cependant, il convient de remarquer que la configuration 4 qui comporte 122 liaisons présente des performances à peine supérieures à celle de la configuration 3 qui en comporte 79. Il convient aussi de remarquer que quel que soit la configuration les pixels traversés par 4 liaisons

ont une erreur quadratique très faible. Dans l'étude réalisée les récepteurs ont été disposés de manière aléatoire en fonction de la densité, l'amélioration des performances peut être obtenue non pas en augmentant le nombre de récepteurs mais en choisissant leur position dans la carte de la manière la plus régulière possible.

Dans la suite de notre étude, nous utiliserons seulement la configuration 3 qui présente un bon compromis du point de vue des performances et du nombre de récepteurs

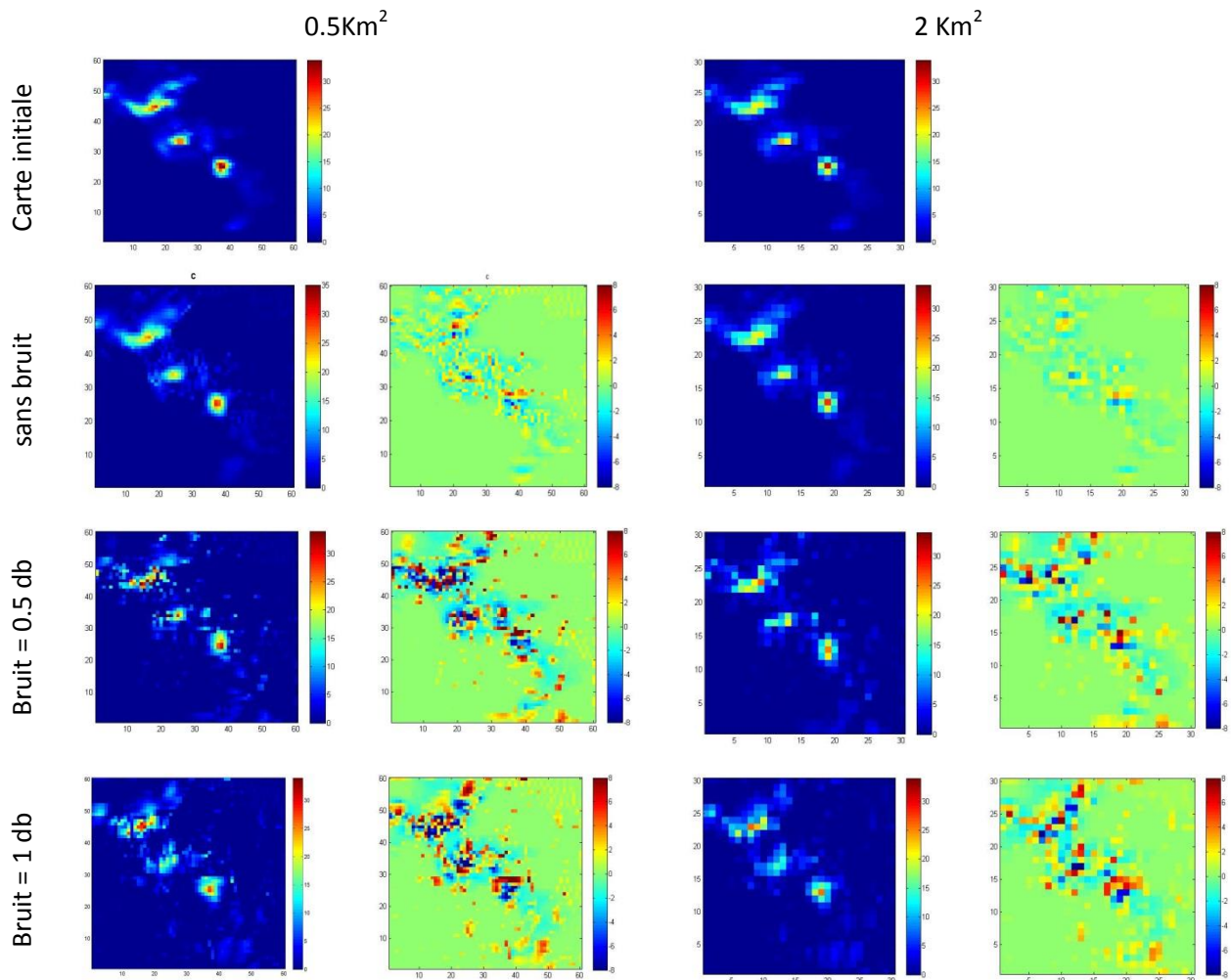
#### ***5.4.2 Influence du bruit de mesure sur la qualité de la restitution***

Nous allons illustrer l'influence du bruit de mesure sur les observations en utilisant la configuration géométrique n°3 et un vent Est – Ouest de 10 m/s. Pour cela, les séries temporelles d'affaiblissement sont bruitées avec un bruit additif gaussien d'écart type 0.5 dB puis 1 dB. Les Figure 5.13 et Figure 5.14) montrent les résultats obtenues d'une part à la résolution spatiale initiale (0.5 km) et d'autre part lorsque les pixels sont agrégés pour obtenir des cartes à la résolution de 2 km. En effet, il est également intéressant de comparer les cartes à une échelle plus grossière mais qui reste toutefois compatible pour des applications en hydrologie urbaine.

A première vue, nous remarquons que les cartes obtenues avec les observations bruitées correspondent à la carte initiale. On retrouve en effet les trois cellules initiales et elles sont correctement positionnées. On aperçoit toutefois la présence de quelques pics sporadiques. Comme on pouvait s'y attendre, l'ajout d'un bruit de mesure dégrade les performances, cependant il convient de nuancer cela. Si on s'intéresse aux performances globales (Tableau 5.5), l'intensité de pluie moyenne sur l'ensemble des zones pluvieuses reste identique (voire meilleures) lorsque les données sont bruitées. Concernant l'intensité moyenne sur l'ensemble de la carte, une dégradation apparaît et est d'autant plus importante que le bruit augmente, l'intensité moyenne passant de 1.13 mm/h lorsque les observations ne sont pas bruitées à 1.29 mm/h pour un bruit de 1 dB, soit une variation de 15%. En dégradant la résolution à 2 km, la forme et la position des cellules de pluie sont bien restitués. Les diagrammes à moustache présentés à la Figure 5.14 montrent que les performances ne sont pas fondamentalement différentes entre un bruit de 0.5 dB et 1 dB. En revanche, l'intensité de pluie moyenne calculée sur toute la carte (Tableau 5.5), est nettement mieux estimée avec un bruit de 0.5 dB que de 1 dB où une

surestimation des taux de pluie est visible. Le phénomène inverse est observé sur l'intensité de pluie moyenne calculée sur les zones de pluie.

Nous pouvons conclure que la présence de bruit de mesure n'empêche pas d'estimer correctement le champ de pluie. Toutefois il apparait qu'avec un bruit de 0.5 dB les zones de pluie sont à peu près restituées avec quelques déformations des zones de pluie tandis qu'avec un bruit à 1 dB il y a apparition de quelques petits évènements de pluie faible. Ce phénomène semble cependant s'atténuer lorsqu'on considère les cartes reconstituées à la résolution de 2 km



**Figure 5.13** Colonnes 1 : Cartes estimées par l'assimilation à partir d'observations non bruitées (ligne 2), avec un bruit de 0,5 dB (ligne 3) et un bruit de 1 dB (ligne 4). Colonne 2 : erreurs correspondantes. Colonnes 3 et 4 : Idem mais à la résolution de 2 km

Configuration 3 Vent Est-ouest		RMSE   Biais (mm/h)	RMSE   Biais des zones de pluie (mm/h)	RMSE   Biais des zones sèches (mm/h)	RR moyen de toute la carte (mm/h)	RR moyen sur les zones de pluie (mm/h)
Carte initial	$\lambda = 0.5 \text{ km}^2$				1.12	3.09
	$\lambda = 2 \text{ km}^2$				1.12	2.80
Sans bruit	$\lambda = 0.5 \text{ km}^2$	0.40   0.01	0.99   -0.08	0.07   0.07	1.13	3.01
	$\lambda = 2 \text{ km}^2$	0.28   0.01	0.60   -0.06	0.06   0.07	1.13	2.74
Bruit =0.5 dB	$\lambda = 0.5 \text{ km}^2$	0.95   0.06	2.19   -0.04	0.24   -0.07	1.19	2.86
	$\lambda = 2 \text{ km}^2$	0.69   0.06	1.39   -0.15	0.22   0.22	1.19	2.64
Bruit =1 dB	$\lambda = 0.5 \text{ km}^2$	0.91   0.17	2.04   0.05	0.27   -0.04	1.29	3.09
	$\lambda = 2 \text{ km}^2$	0.76   0.17	1.50   0.05	0.25   0.25	1.29	2.85

Tableau 5.5 Biais, RMSE et intensité de pluie moyenne calculés à partir des cartes estimées par assimilation à la résolution pour des bruits de mesure de 0, 0.5 dB et 1 dB

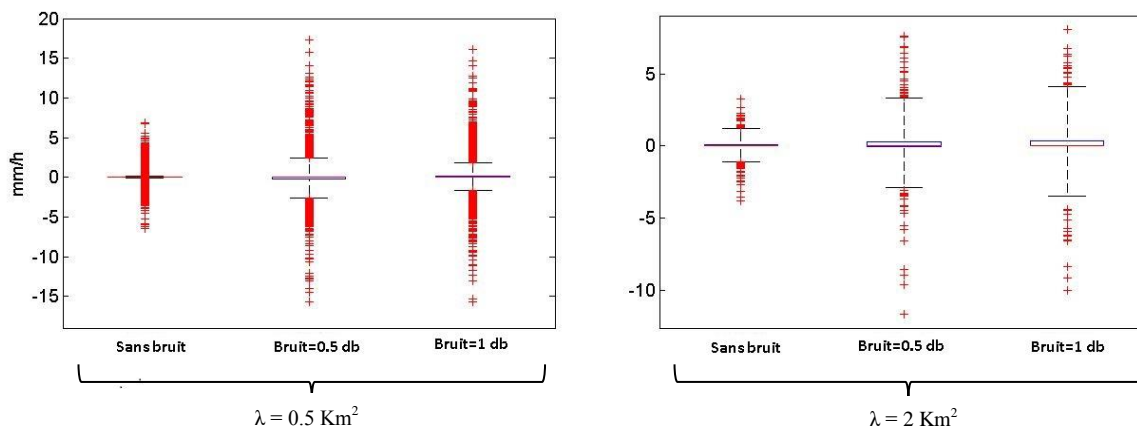


Figure 5.14. Boîtes à moustaches illustrant la différence entre la carte initiale et les résultats d'assimilation avec un bruit de 0.5 dB et 1 dB.

### 5.4.3 Echelles temporelles d'observation

Le principe de ce système de mesure est qu'à partir d'observations régulièrement espacées dans le temps et à suffisamment fine résolution temporelle (10 s), il est possible de reconstituer le champ précipitant. Il s'agit d'évaluer l'influence de la cadence d'échantillonnage des observations sur la qualité de la restitution. En effet, si des récepteurs sont placés chez des particuliers, il semble plus réaliste d'utiliser des cadences plus lentes. Pour cette raison, nous

supposerons trois périodes d'échantillonnage : 10 secondes, 1 minute et 5 minutes. Nous avons utilisé pour nos tests la configuration géométrique n°3 avec un vent Est-Ouest de vitesse 10 m/s et un bruit de 0.5 dB.

La Figure 5.15 illustre les résultats obtenus à la résolution de 0.5 km, puis à la résolution de 2 km. Il apparaît que quelle que soit la résolution l'algorithme de restitution retrouve les 3 cellules à leur bonne position. A la résolution de 1 minute, la carte reconstituée n'est pas très différente de celle reconstituée à 10 secondes, l'algorithme reconstitue correctement les trois cellules de la carte initiale. Avec des observations à 5 minutes de résolution, les cellules, bien que restituées n'ont plus la même forme que celles d'origine, nous remarquons un éclatement et apparition de plusieurs petits événements d'intensités variables. Cela est confirmé par les indicateurs calculés dans le Tableau 5.6 où nous remarquons pour les zones sèches que le biais sur les pixels, quasi nul pour un échantillonnage à 10 secondes, passe à 0.66 mm/h à la résolution de 5 minutes. Il en est de même pour la RMSE qui voit sa valeur triplée. La Figure 5.16 qui présente des boîtes à moustache relatives aux écarts entre pixels estimés et valeur vraie du pixel montre également un accroissement important de la taille des moustaches (1<sup>er</sup> et 3<sup>ème</sup> quartile) à la résolution de 5 minutes. Concernant l'intensité de pluie moyenne sur l'ensemble de la carte, l'augmentation de l'intervalle d'échantillonnage accroît substantiellement la valeur estimée. Cependant, nous n'avons pas actuellement, faute de temps, pu approfondir cet aspect et nous n'avancerons donc pas d'explication à ce comportement.

En conclusion de ce paragraphe, il semble que des observations à 5 minutes d'intervalles ne permettent pas de restituer correctement le champ de pluie. Toutefois il convient de pondérer cette affirmation, les performances dépendent également d'autres paramètres, notamment la configuration géométrique des récepteurs. On peut en effet, sans prendre trop de risque, affirmer que le fait d'augmenter le nombre de récepteurs permettrait sans aucun doute d'estimer correctement le champ de pluie avec une période d'échantillonnage de 5 minutes.

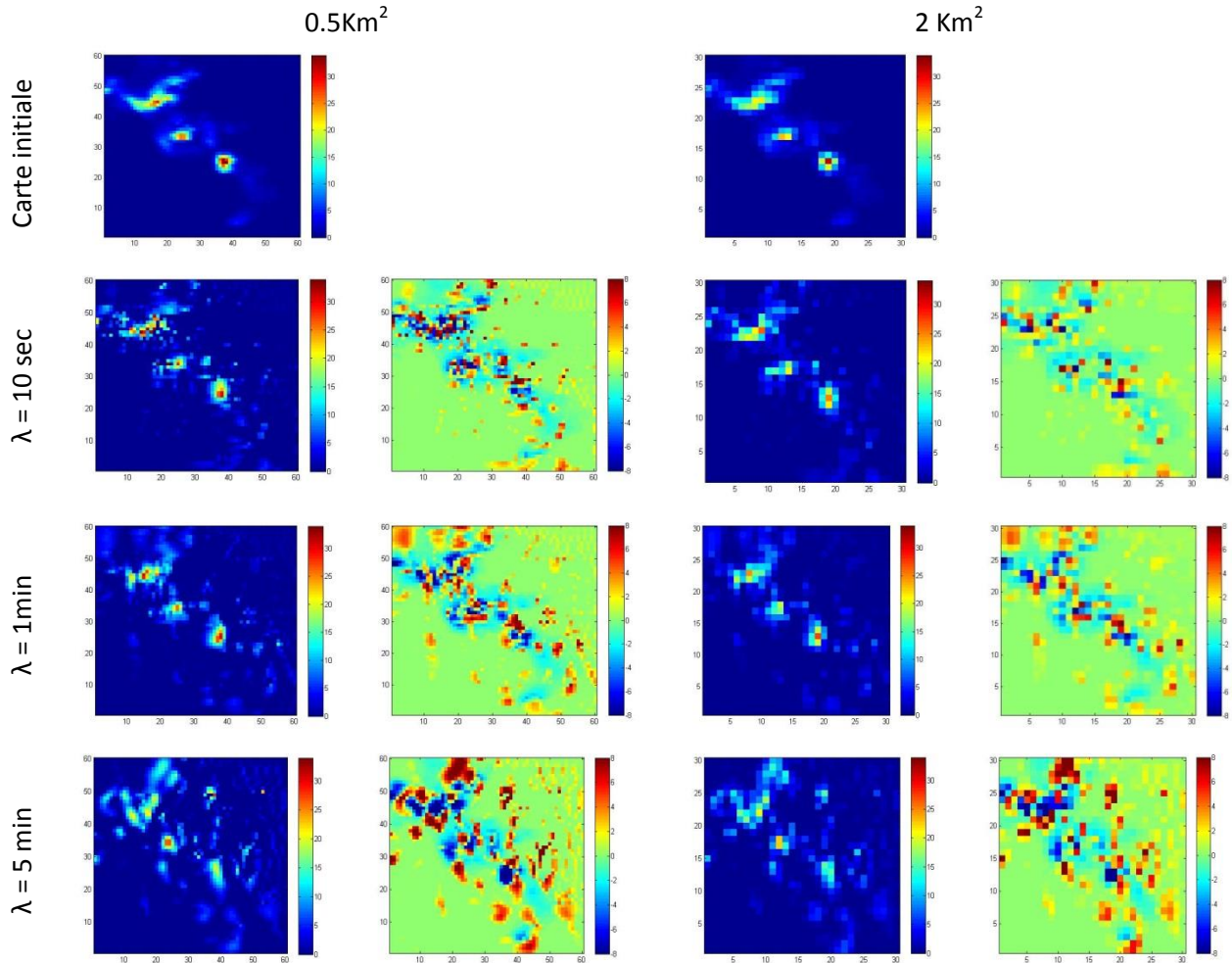


Figure 5.15 Cartes estimées pour trois résolutions temporelles (10s, 1 min et 5 min) à la résolution spatiale 0.5 km et 2 km

Configuration 3 Vent Est-ouest Bruit =0.5 dB		RMSE   Biais (mm/h)		RMSE   Biais des zones de pluie (mm/h)		RMSE   Biais des zones sèches (mm/h)		RR moyen de toute la carte (mm/h)	RR moyen sur les zones de pluie (mm/h)
Carte initial	λ= 0.5 km <sup>2</sup>							1.12	3.09
	λ= 2 km <sup>2</sup>							1.12	2.80
λ = 10 sec	λ= 0.5 km <sup>2</sup>	0.95	0.06	2.19	-0.04	0.24	-0.07	1.19	2.86
	λ= 2 km <sup>2</sup>	0.69	0.06	1.39	-0.15	0.22	0.22	1.19	2.64
λ = 1 min	λ= 0.5 km <sup>2</sup>	1.03	0.18	2.17	-0.18	0.39	0.39	1.30	2.90
	λ= 2 km <sup>2</sup>	0.90	0.18	1.68	-0.12	0.39	0.39	1.30	2.67
λ = 5 min	λ= 0.5 km <sup>2</sup>	1.52	0.46	2.99	0.07	0.69	0.69	1.58	3.17
	λ= 2 km <sup>2</sup>	1.38	0.46	2.48	0.17	0.66	0.66	1.58	2.97

Tableau 5.6 Biais, RMSE et intensité de pluie moyenne calculés à partir des cartes estimées par assimilation à la résolution de 10 secondes, 1 minute et 5 minutes

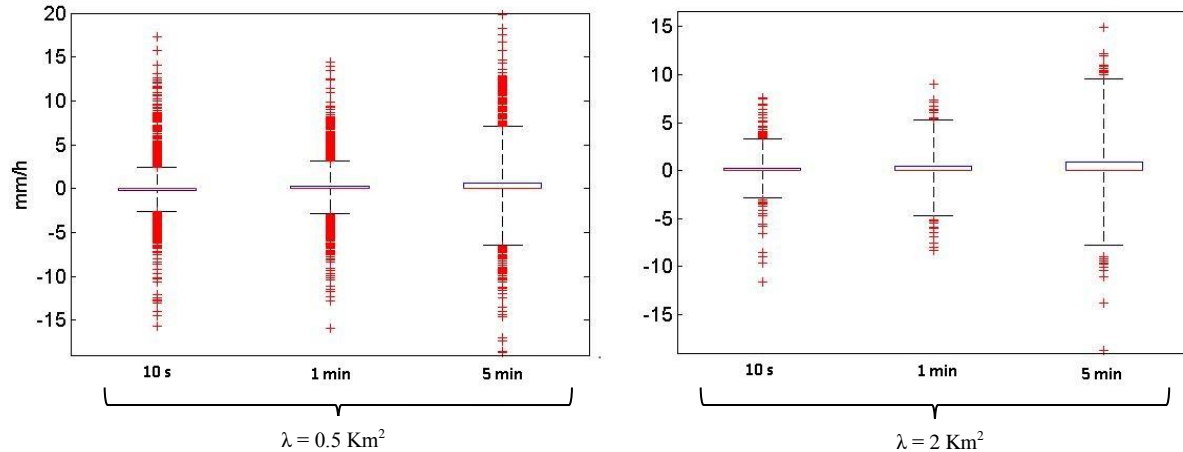


Figure 5.16. Boîtes à moustaches des erreurs entre la carte initiale et les cartes estimées par assimilation pour 3 résolutions temporelles (10s, 1min et 5 min)

#### 5.4.4 Résultats en fonction de taux précipitants

Ce paragraphe est dédié à l'analyse des résultats obtenus sur d'autres cartes différentes de celle utilisée dans le reste de ce chapitre que ce soit en termes d'intensité de pluie ou encore de structure des évènements présents sur la carte. Dans cette partie, nous avons décidé d'utiliser la configuration jugée optimale dans les tests précédents c'est-à-dire, la configuration 3 avec le vent Est-ouest constant. La Figure 5.17 illustre les champs restitués avec et sans bruit. L'analyse de cette figure et du Tableau 5.7 montre que les structures, les taux de pluie moyens des zones de pluie et des cartes sont bien restitués sur les trois champs assimilés sans bruit. Avec un bruit de 0.5 dB, les taux de pluie moyen sont bien restitués tandis que les structures des évènements ne semblent être bien restituées que sur les résultats des deux cartes 1 et 3.

En s'intéressant plus en détail à la carte 2, nous remarquons que par advection le taux de pluie maximal de la carte se déplace vers l'Ouest et ne traverse que peu de liaisons. Cette situation, dans le cas où les données sont bruitées, a tendance à créer des pics dans toute la région pluvieuse sans pouvoir localiser correctement le taux de pluie maximal.

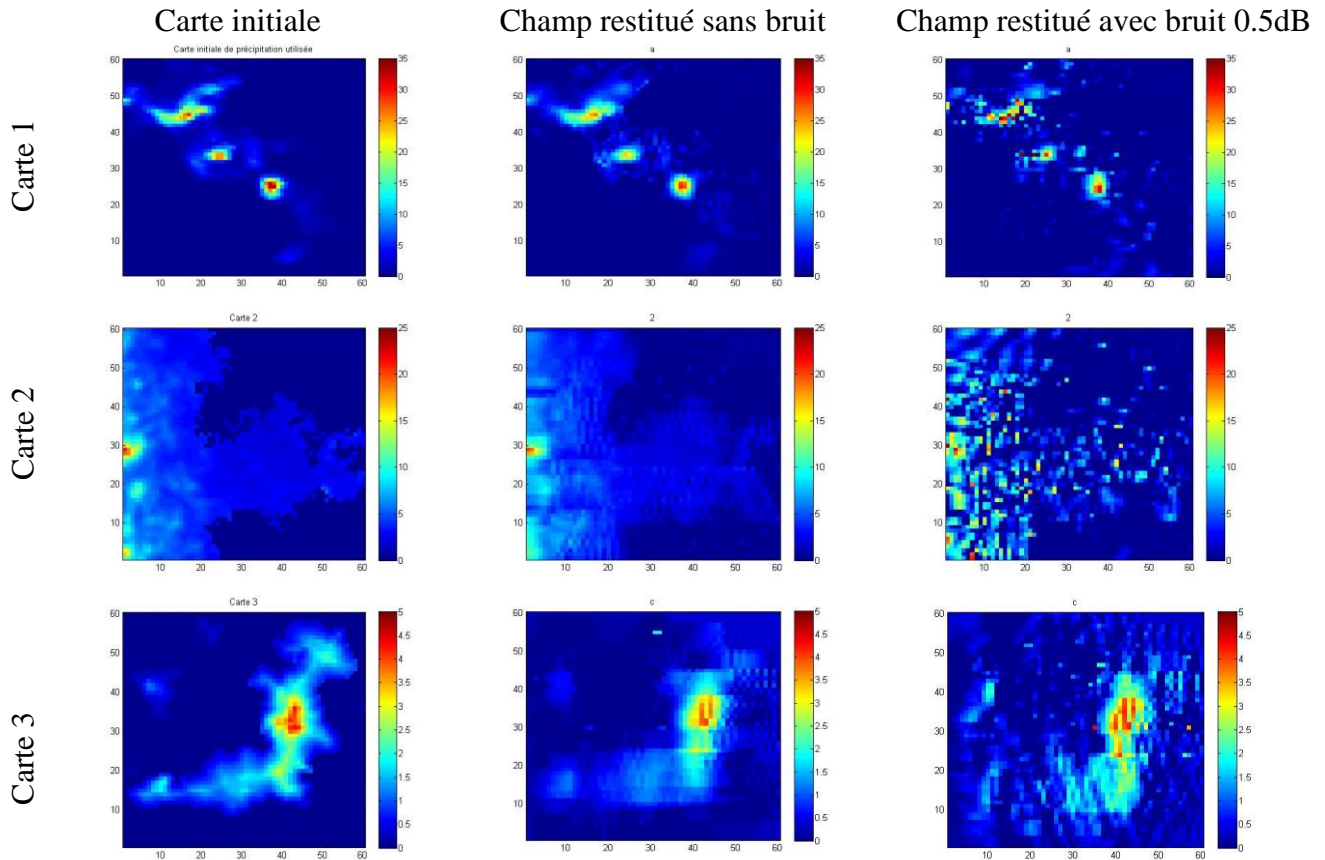


Figure 5.17. Cartes restituées avec et sans bruit, en utilisant la configuration 3 avec un vent Est-Ouest

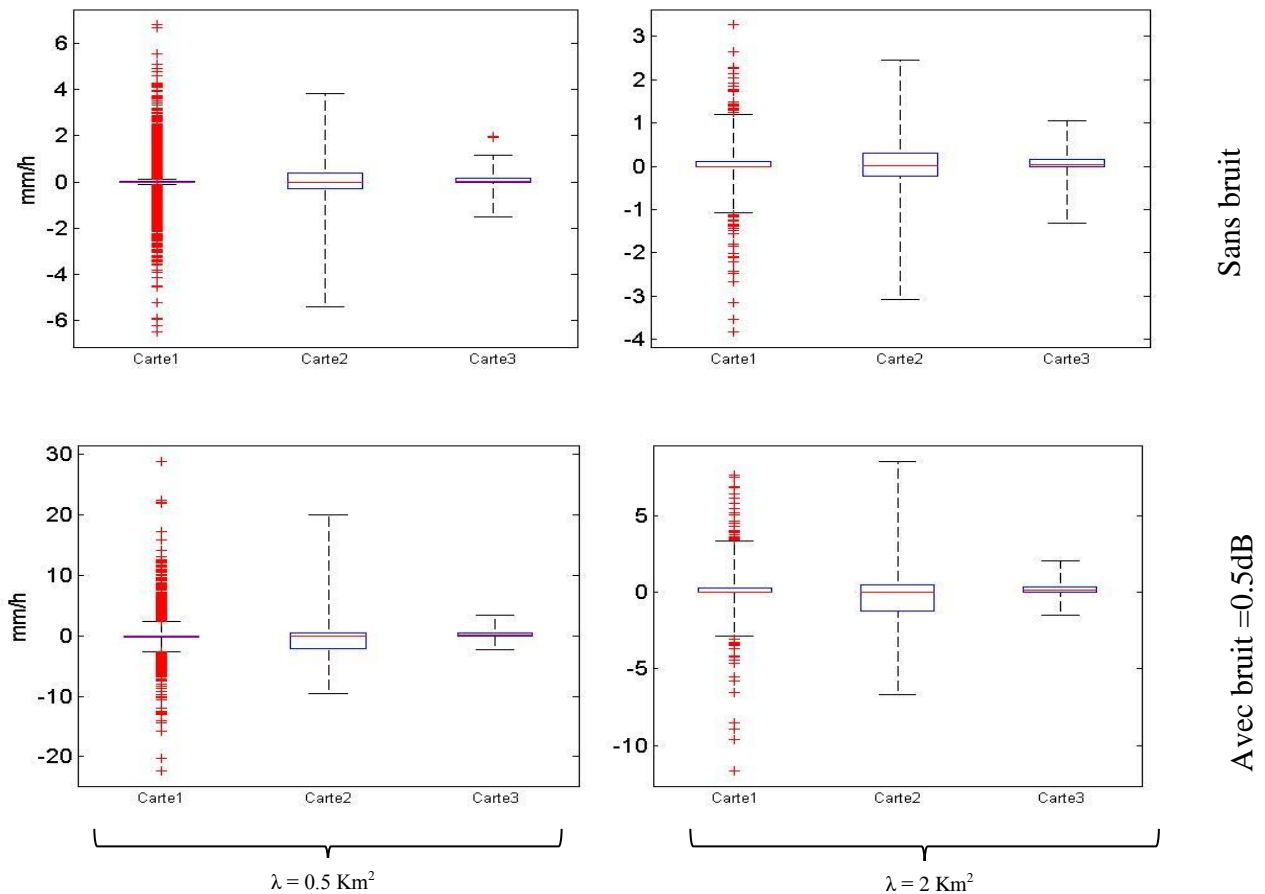
Configuration 3 Resolution = 10s Vent Est-Ouest		RMSE   Biais (mm/h)	RMSE   Biais des zones de pluie (mm/h)	RMSE   Biais des zones sèches (mm/h)	RR moyen de toute la carte (mm/h)	RR moyen sur les zones de pluie (mm/h)
Carte 1	Care initiale				1.12	3.09
	Sans bruit	0.40   0.01	0.99   -0.08	0.07   0.07	1.13	3.01
	Bruit = 0.5dB	0.95   0.06	2.19   -0.04	0.24   -0.07	1.19	2.86
Carte 2	Care initiale				2.26	3.92
	Sans bruit	0.52   0.01	0.64   -0.24	0.36   0.38	2.28	3.68
	Bruit = 0.5dB	1.86   -0.11	2.90   -0.51	0.43   0.43	2.15	3.40
Carte 3	Care initiale				0.38	1.10
	Sans bruit	0.18   0.02	0.30   -0.16	0.12   0.12	0.41	0.95
	Bruit = 0.5dB	0.35   0.13	0.52   -0.11	0.27   0.27	0.51	0.98

Tableau 5.7. Grandeurs statistiques sur la qualité des champs restitués en utilisant la configuration 3 pour les trois cartes de pluie avec et sans bruit.



L'analyse précédente est confirmée par les boîtes à moustache des résidus calculés sur les trois champs assimilés avec et sans bruit (Figure 5.18).

L'écart type de l'erreur pour la carte 2 est plus large que les deux autres cartes qui sont presque nuls. Les résidus à la résolution de 2 km sont plus faibles que ceux à la résolution 0.5 km. Comme dit précédemment, cela veut dire que les pics qui apparaissent à 0.5 km de résolution sont lissés à 2 km, et comme on a pratiquement la bonne quantité d'eau, la dégradation de la résolution permet de se rapprocher du champ initial.



**Figure 5.18. Boîtes à moustache présentant les résidus à la résolution de 0.5 km et 2 km, avec et sans bruit des 3 cartes utilisées.**

La Figure 5.12 comme la Figure 5.19 représentent la RMSE des pixels de taux précipitant en fonction du nombre de liaisons passant dans le pixel. Alors que précédemment on a étudié les RMSE des pixels en fonction du nombre de liaisons avec la carte 1 en considérant plusieurs configurations (Figure 5.12), on considère maintenant 3 cartes différentes avec une même configuration (Figure 5.19). Les conclusions sont identiques, La RMSE diminue rapidement en

fonction de nombre de liaisons, elle est quasiment nulle sur un pixel s'il est traversé par plus de 2 liaisons.

On constate d'autre part une grande variabilité d'une carte à l'autre pour les pixels ayant 0 liaison, cette variabilité est encore plus importante quand les données sont bruitées. La qualité de l'estimation du taux précipitant dans les zones non couvertes par les liaisons est très variable suivant la configuration mais également suivant la structure de l'évènement de pluie. On constate par contre que pour les pixels traversés par 2 liaisons ou plus, la RMSE reste très faible même en présence de bruit.

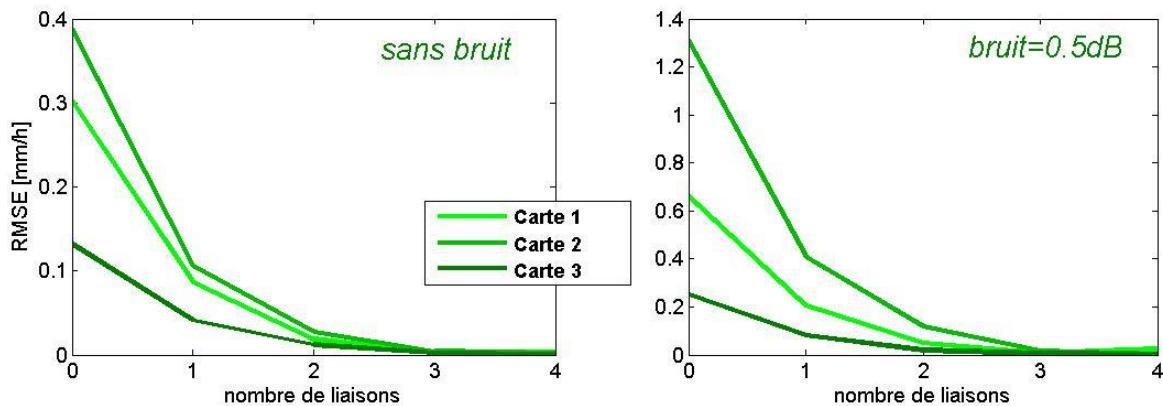


Figure 5.19. RMSE des pixels des trois cartes restituées avec et sans bruit en fonction de nombre de liaisons qui les traversent

## 5.5 Conclusion

L'objectif de ce chapitre était de réaliser une première étude de faisabilité visant à reconstituer des champs de pluie à partir d'observations issues d'affaiblissements d'ondes TV en bande Ku. Cette étude s'est appuyée sur des cartes de précipitation simulées issues du générateur 2D développé au chapitre précédent. Nous avons étendu le modèle développé par Mercier et al. (2015) à des configurations réalistes de zones urbaines.

Pour une zone d'étude de  $30 \times 30 \text{ km}^2$  l'utilisation de quelques dizaines de récepteurs répartis aléatoirement permettait de reconstituer relativement bien les structures des cellules pluvieuses ainsi que la quantité d'eau sur l'ensemble de la zone (erreur moyenne quadratique de l'ordre de  $0.5 \text{ mm/h}$  pour la configuration testée). Une estimation plus précise pouvant être obtenue en

augmentant le nombre de récepteurs ou en optimisant leur position sur la carte. L'ajout d'un bruit de mesure réaliste bien qu'ayant un effet certain n'empêche pas une bonne estimation du champ précipitant. La période d'échantillonnage semble être un paramètre important, un échantillonnage trop lent dégradant fortement les performances.

Cette première étude de faisabilité devra être approfondie pour trouver ce que devrait être une configuration optimale à la fois en termes de nombres de récepteurs et de leur distribution dans l'espace qu'en termes de période d'échantillonnage et de précision de la mesure.

L'apparition de pics sporadiques est bien lisible dans le cas des données bruitées (au-delà de 0.5 dB), surtout dans le cas d'insuffisance d'observations (résolutions temporelles grossières).

Cette étude encore préliminaire, devra être complétée par des simulations supplémentaires (autres configurations) et des expériences de terrain (campagne de mesure prévue en 2016 avec 2 systèmes de mesures hyperfréquences). Les résultats sont en effet très prometteurs, le dispositif ainsi défini permet une restitution en temps réel du champ de pluie avec une résolution spatiale de  $0.5 \times 0.5 \text{ km}^2$  et une répétitivité temporelle d'une seconde. Sous réserve d'optimiser le dispositif, la précision semble pouvoir être comparable ou meilleure que celle des systèmes d'observation existant (Radar, réseau de pluviomètre, satellite). Le système d'observation basé sur l'utilisation opportuniste de sources hyperfréquences disponibles en bande Ku, présente des caractéristiques en termes de résolution, de répétitivité, de coût de maintenance et de robustesse qui peut être très intéressant pour des applications en hydrologie urbaine ou pour les zones du globe actuellement dépourvues de système d'observation efficaces.

## 6 CONCLUSION ET PERSPECTIVES

Le but de ce travail de thèse était l'étude de la variabilité spatiale et temporelle des précipitations avec une attention particulière à la variabilité interne aux événements et au caractère intermittent du processus. Il s'agissait de tirer les conséquences pour la simulation, de travaux antérieurs, relatifs à l'analyse des mesures à partir d'un disdromètre à très fine résolution temporelle (DBS). Cela a permis de renforcer l'hypothèse du caractère « universel » de certaines propriétés d'échelle observées à l'intérieur des événements (thèses antérieures de S. Verrier et de L. de Montera) et d'autre part d'établir le lien entre les propriétés du support et l'existence de différents régimes d'invariance d'échelle.

L'objectif était de pouvoir générer des séries chronologiques ou des cartes de précipitations qui reproduisent les propriétés statistiques à l'échelle des observations sur lesquelles elles reposent, mais également après agrégation à d'autres échelles. Il s'agit donc de générer des taux précipitants ou des lames d'eau dont les distributions statistiques sont identiques à celles observées mais également dont les répartitions dans le temps ou dans l'espace sont réalistes. Pour atteindre cet objectif une attention particulière a été portée à la modélisation et à la simulation de l'intermittence et des relations d'échelles.

Dans cette thèse nous avons donc proposé et implémenté un nouveau modèle de génération de séries temporelles de précipitation à fine échelles pour un climat île de France. La méthodologie développée pour la simulation de séries chronologiques a été ensuite adaptée pour générer des cartes de précipitation 2D. Une méthodologie en trois étapes liées aux hypothèses sur lesquelles reposent les simulations a été développée : simulation du support, simulation multifractale (FIF) des périodes ou zones de pluie, re-normalisation des séries obtenues. Les différentes propriétés des séries ou des cartes synthétiques validées sont les distributions statistiques à différentes échelles, les propriétés multifractales et spectrales.

Concernant le générateur de séries temporelles, plusieurs versions ont été comparées pour la simulation du support : simulation avec une loi GEV ou avec une distribution Pareto ; distinction ou non entre les durées courtes et les durées longues des événements. De même pour re-normaliser les taux de pluie générés par le modèle FIF deux approches ont été envisagées : une calibration quantile à quantile des CDF, une calibration par événement qui repose sur la

modélisation de la relation entre le taux de pluie moyen de l'évènement de pluie et sa durée. Finalement deux générateurs de séries chronologiques sont présentés le premier qui ne fait pas de distinction entre les durées courtes et longues nécessite l'ajustement de seulement 9 paramètres mais ne reproduit pas parfaitement les propriétés spectrales dans toutes les gammes d'échelle. La distinction entre durées courtes et longues introduites dans le second générateur augmente significativement le nombre de paramètres à ajuster mais permet de mieux reproduire les propriétés du processus dans toute la gamme d'échelle considérée.

Si on fait l'hypothèse que les propriétés spécifiques aux durées courtes (inférieures à 5 mn) utilisées dans le second générateur pour la simulation du support et la re-normalisation des séries sont « universelles », en ce sens qu'elles sont identiques en tout point du globe, le nombre de paramètres à ajuster pour s'adapter à la localisation est pratiquement identique pour les deux générateurs. Cependant, pour être testée cette hypothèse nécessiterait d'analyser des observations faites en différents points du globe (ainsi que pour les divers types de temps associés) à une résolution temporelle très inférieure à 5 mn. Nous ne disposons pas actuellement de données de cette nature en quantité suffisante pour qu'une telle étude soit envisageable.

Dans le domaine spatial nous ne disposons pas actuellement d'observations à très fine résolution, l'adaptation du générateur de séries chronologiques à la génération de cartes repose sur l'analyse des cartes de lames d'eau de radars météorologique fournies par Météo France à la résolution du  $\text{km}^2$ . La méthodologie en trois étapes est complétée par une quatrième étape nécessaire pour lisser les bords des zones de pluies et obtenir des transitions pluie/non pluie réaliste. Ce générateur nécessite l'ajustement de seulement 8 paramètres.

Actuellement le caractère « universel » de certaines propriétés, telles que les propriétés multifractales observées à l'intérieur des évènements ou pour l'existence de différents régimes observés dans les spectres, repose sur des études bibliographiques. Ce caractère « universel » laisse supposer que la méthodologie proposée permettra une adaptation aisée à d'autres régions en ajustant simplement les paramètres des différentes lois utilisées.

Seule l'analyse de données recueillies dans d'autres régions climatiques permettra de vérifier le caractère « universel » des propriétés utilisées et donc le caractère générique de la méthodologie envisagée pour la simulation.

Par ailleurs plusieurs études complémentaires seront nécessaires pour affiner les capacités des générateurs proposés ; nous citons par exemple la prise en compte de certaines covariables tel que

le type du temps et la température L'analyse et la prise en compte dans les simulations de la saisonnalité nécessitent d'analyser des séries d'observations plus longues. Dans le cadre de notre modélisation, les évènements sont considérés indépendants, la prise en compte de la corrélation entre les évènements successifs peut améliorer visiblement les distributions de la pluie générés aux résolutions grossières. Pour la simulation des cartes l'étude des propriétés du support à fine échelle nécessite de disposer d'observation à plus fine résolution. Malgré la forte plus-value du radar par rapport à un réseau de pluviomètres, la résolution des données mise à notre disposition ( $1 \text{ km}^2$ ) ne permet pas une distinction fine des contours des zones de pluie.

L'amélioration des connaissances relatives à la variabilité spatiotemporelle des précipitations n'est rendue possible que par l'amélioration des moyens d'observation. Constatant les difficultés rencontrées pour obtenir des séries d'observations suffisamment longues, continues, avec une résolution spatiale fine et couvrant des zones climatiques variées le dernier chapitre de la thèse est dédié à une étude de faisabilité d'un tel système d'observation. Le système hyperfréquence envisagé repose sur l'observation opportuniste des affaiblissements atmosphériques des signaux issus de satellites géostationnaires en bande Ku. L'étude menée à partir de cartes simulées montre qu'il est possible d'obtenir des champ précipitants à la résolution spatiale de  $0,5 \times 0,5 \text{ km}^2$  avec une résolution temporelle inférieure à 1 mn en milieu urbain avec seulement un seul récepteur TV par  $\text{km}^2$  en moyenne

Les perspectives des travaux menées au court de cette thèse sont nombreuses:

A court terme la consolidation des codes informatiques et la mise en ligne des simulateurs développés (Open Sources) permettront leur mise à disposition de la communauté scientifique.

A plus long terme l'extension de la dimension 2 à la dimension 3 pose, entre autres, des problèmes numériques (temps de calcul, espace mémoire). Les cartes 2D sont actuellement générées indépendamment les unes des autres sans prendre en compte les relations temporelles existant éventuellement entre cartes consécutives. Les dépendances temporelles modélisées dans le simulateur de séries chronologiques 1D serviront de point de départ pour l'étendre de la dimension 2 à la dimension 3 (2D spatiale + 1D temporelle).

Le système de mesure hyperfréquence tel qu'il a été envisagé dans les simulations à l'échelle urbaine pourrait théoriquement fournir les observations nécessaires au paramétrage d'un tel

simulateur dans des zones climatiques variées. Les récepteurs TV sont en quantité suffisante en de nombreux points du globe. Des difficultés importantes existent cependant pour obtenir les mesures nécessaires à la reconstruction des champs précipitants. De manière plus réaliste une campagne de mesure sera réalisée prochainement au SIRTA sur le site instrumental de l'IPSL à Palaiseau combinant différents instruments ayant des caractéristiques complémentaires (4 disdromètres, deux récepteurs hyperfréquences en bande Ku mesurant l'affaiblissement sur 8 liaisons satellites, des mesures du radar Météo France) de trappes. Cette campagne devrait fournir les observations adéquates au paramétrage et à la validation d'un simulateur 2D spatial + 1D temporel.

Le travail réalisé a présenté un grand intérêt du fait de l'assimilation de nouvelles compétences dans des domaines variés : analyse et traitement de jeux de données observées et simulées, traitement d'image, traitement de séries chronologiques, modélisation des propriétés statistiques, spectrales, multifractales aussi bien dans le domaine temporel que spatial, implémentation et validation de codes pour la simulation et l'analyse, réflexion méthodologique, assimilation de données dans un modèle numérique simple, sans oublier la rédaction d'article, la préparation de posters et de conférences ainsi que les activités d'enseignement.

## Annexe

### Méthode des moments trace :

Cette Annexe vise à présenter très brièvement la méthode des Moments Trace TM (Schertzer et Lovejoy, 1987) utilisée dans le cadre de notre analyse multifractale. La méthode TM permet de retrouver les paramètres multifractals  $\alpha$  et  $c_1$  en calculant la fonction d'échelle des moments  $K(q)$  (Equation 2.4).

Cette méthode passe par le calcul, pour chaque résolution  $\lambda$  et pour chaque ordre  $q$  ( $0 < q < q_{max}$ ), le moment empirique d'ordre  $q$ .

Ces moments sont déterminés en additionnant les valeurs de  $\varepsilon_\lambda$  sur les pixels de résolution  $\lambda$ , dont on prend ensuite la puissance  $q$ . On calcule ensuite leur somme normalisée.

$$E[(\varepsilon_\lambda)^q] = \frac{\sum_n \varepsilon_\lambda^q}{\lambda^D}$$

Où  $n$  correspond au nombre de pas de temps en 1D ou au nombre de pixels en 2D.

Dans le cas de plusieurs échantillons, les grandeurs  $E[(\varepsilon_\lambda)^q]$  sont calculées indépendamment. On détermine ensuite la moyenne globale.

La représentation en échelle logarithmique de la quantité obtenue en fonction de  $\lambda$  permet de déterminer la fonction  $K(q)$ . En effet, si la dépendance est linéaire en échelle log-log, on dit que le champ  $\varepsilon_\lambda$  possède la propriété d'invariance d'échelle et que la pente de la droite obtenue fournit la valeur de  $K(q)$  pour chaque valeur de  $q$ . Cela revient à prendre le log de l'Equation 2.15. Afin de passer d'une résolution  $\lambda$  à une résolution plus grossière, il suffit de dégrader le champ à cette résolution.





## Bibliographie

Adhikari, A., Das, S., Bhattacharya, A., and Maitra, A. (2011). Improving rain attenuation estimation: Modelling of effective path length using Ku-band measurements at a tropical location. *Progress In Electromagnetics Research B*, 34, 173-186.

Akrour, N., Chazottes, A., Verrier, S., Mallet, C., & Barthes, L. (2015). Simulation of yearly rainfall time series at microscale resolution with actual properties: Intermittency, scale invariance, and rainfall distribution. *Water Resources Research*.

Arbesser-Rastburg, B. R., and Brussaard, G. (1993). Propagation research in Europe using the OLYMPUS satellite. *Proceedings of the IEEE*, 81(6), 865-875.

Barnolas, M., T. Rigo, and M. Llasat (2010), Characteristics of 2-d convective structures in catalonia (ne spain): an analysis using radar data and gis, *Hydrology and Earth System Sciences*, 14 (1), 129-139.

Barthès, L., and C. Mallet (2013), Rainfall measurement from the opportunistic use of an earth-space link in the KU band, *Atmospheric Measurement Techniques*, 6 (8), 2181-2193.

Bernardara, P., C. De Michele, and R. Rosso (2007), A simple model of rain in time: An alternating renewal process of wet and dry states with a fractional (non-gaussian) rain intensity, *Atmospheric research*, 84 (4), 291-301.

Bernadara, M. Lang, E. Sauquet, D. Scertzer and I. Tchiguirinskaia, *Analyse multifractale en hydrologie*, editions Quae, 2007]

Bocquet, M. (2005), Introduction aux principes et méthodes de l'assimilation de données en géophysique, Notes de cours de l'ENSTA.

Burton, A., C. Kilsby, 784 H. Fowler, P. Cowpertwait, and P. O'Connell (2008), Rainsim: A spatial-temporal stochastic rainfall modelling system, *Environmental Modelling & Software*, 23 (12), 1356-1369.

Chiles, J.-P., and P. Delfiner (2009), *Geostatistics: modeling spatial uncertainty*, vol. 497, 695pp., John Wiley & Sons.

Cooley, D., D. Nychka, and P. Naveau (2007), Bayesian spatial modeling of extreme precipitation return levels, *Journal of the American Statistical Association*, 102 (479), 824-840.

Daget, N. (2007), *Revue des méthodes d'assimilation*.

De Lima, D. M. (1998). *Multifractal and the temporal structure of rainfall*. Phd thesis

De Michele, C., and M. Ignaccolo (2013), New perspectives on rainfall from a discrete view, *Hydrological Processes*, 27 (16), 2379-2382.

- De Michele, C., and G. Salvadori (2003), A generalized pareto intensity-duration model of storm rainfall exploiting 2-copulas, *Journal of Geophysical Research: Atmospheres* (1984-2012), 108 (D2).
- De Michele, C., and Bernardara, P. (2005). Spectral analysis and modeling of space-time rainfall fields. *Atmospheric Research*, 77(1), 124-136.
- De Montera, L., L. Barthès, C. Mallet, and P. Golé (2009), The effect of rain-no rain intermittency on the estimation of the universal multifractals model parameters, *Journal of Hydrometeorology*, 10 (2), 493-506.
- De Montera, L., S. Verrier, C. Mallet, and L. Barthès (2010), A passive scalar-like model for rain applicable up to storm scale, *Atmospheric Research*, 98 (1), 140-147.
- Deidda, R. (1999), Multifractal analysis and simulation of rainfall fields in space, *Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere*, 24 (1), 73-78.
- Deidda, R. (2000), Rainfall downscaling in a space-time multifractal framework, *Water Resources Research*, 36 (7), 1779-1794.
- Deidda, R., R. Benzi, and F. Siccaldi (1999), Multifractal modeling of anomalous scaling laws in rainfall, *Water resources research*, 35 (6), 1853-1867.
- Delahaye, J.-Y., L. Barthès, P. Golé, J. Lavergnat, and J.-P. Vinson (2006), A dual-beam spectropluviometer concept, *Journal of Hydrology*, 328 (1), 110-120.
- Douglas, E. M., and Barros, A. P. (2003). Probable maximum precipitation estimation using multifractals: application in the Eastern United States. *Journal of Hydrometeorology*, 4(6), 1012-1024.
- Emmanuel, I., E. Leblois, H. Andrieu, and B. Flahaut (2011), Variabilité spatio-temporelle des précipitations aux échelles de l'hydrologie urbaine, *La Houille Blanche*, (4), 31-36.
- Evin, G., and A.-C. Favre (2013), Further developments of a transient poisson-cluster model for rainfall, *Stochastic Environmental Research and Risk Assessment*, 27 (4), 831-847.
- Fabry, F. (1996). On the determination of scale ranges for precipitation fields. *Journal of Geophysical Research: Atmospheres* (1984–2012), 101(D8), 12819-12826.
- Falconer, K. (1990). *Fractal geometry. J. Mathematical Foundations and Applications.*
- Fraedrich, K., and C. Larnder (1993), Scaling regimes of composite rainfall time series, *Tellus A*, 45 (4), 289-298.

- Gaume, E., N. Mouhous, and H. Andrieu (2007), Rainfall stochastic disaggregation models: Calibration and validation of a multiplicative cascade model, *Advances in water resources*, 30 (5), 1301-1319.
- Ghanmi, H., Z. Bargaoui, and C. Mallet (2013), Investigation of the fractal dimension of rainfall occurrence in a semi-arid mediterranean climate, *Hydrological Sciences Journal*, 58 (3), 483-497.
- Gires, A., I. Tchiguirinskaia, D. Schertzer, and S. Lovejoy (2012), Influence of the zero rainfall on the assessment of the multifractal parameters, *Advances in Water Resources*, 45, 13-25.
- Gires, A., I. Tchiguirinskaia, D. Schertzer, and S. Lovejoy (2013), Development and analysis of a simple model to represent the zero rainfall in a universal multifractal framework, *Nonlin. Processes Geophys.*, 20, 343-356.
- Giuli, D., Facheris, L., and Tanelli, S. (1999). Microwave tomographic inversion technique based on stochastic approach for rainfall fields monitoring. *Geoscience and Remote Sensing, IEEE Transactions on*, 37(5), 2536-2555.
- Giuli, D., Toccafondi, A., Gentili, G. B., and Freni, A. (1991). Tomographic reconstruction of rainfall fields through microwave attenuation measurements. *Journal of Applied Meteorology*, 30(9), 1323-1340.
- Gómez-Hernández, J. J., and A. G. Journel (1993), Joint sequential simulation of multigaussian fields, in *Geostatistics Troia92*, pp. 85-94, Springer.
- Groisman, P. Y., and D. R. Easterling (1994), Variability and trends of total precipitation and snowfall over the united states and canada, *Journal of Climate*, 7 (1), 184-205.
- Gupta, V. K., and E. C. Waymire (1979), A stochastic kinematic study of subsynoptic space-time rainfall, *Water Resources Research*, 15 (3), 637-644.
- Gupta, V. K., and E. C. Waymire (1993), A statistical analysis of mesoscale rainfall as a random cascade, *Journal of Applied Meteorology*, 32 (2), 251-267.
- Gupta, V. K., Castro, S. L., and Over, T. M. (1996). On scaling exponents of spatial peak flows from rainfall and river network geometry. *Journal of hydrology*, 187(1), 81-104
- Habib, E., W. F. Krajewski, and A. Kruger (2001), Sampling errors of tipping-bucket rain gauge measurements, *Journal of Hydrologic Engineering*, 6 (2), 159-166.
- Hubert, P., Carbonnel, J. P., and A. Chaouche, (1989). Segmentation des séries hydrométéorologiques—application à des séries de précipitations et de débits de l'Afrique de l'ouest. *Journal of hydrology*, 110(3), 349-367.
- Hubert, P., Y. Tessier, P. Ladoy, S. Lovejoy, D. Schertzer, J.P. Carbonnel, S. Violette, I. Desurosne, et F. Schmitt, 1993. Multifractals and extreme rainfall events. *Geophys. Res. Lett.* 20, 931–934.

Hubert, P., & Carbonnel, J. P. (1993). Segmentation des séries annuelles de débits de grands fleuves africains. *Bulletin de liaison du CIEH*, 92(3).

Hubert, P. (2001), Multifractals as a tool to overcome scale problems in hydrology, *Hydrological sciences journal*, 46 (6), 897-905.

Hubert, P. (2005). La prédétermination des crues. *Comptes Rendus Geoscience*, 337(1), 219-227.

Ignaccolo, M., and C. De Michele (2010), A point based eulerian definition of rain event based on statistical properties of inter drop time intervals: An application to chibolton data, *Advances in Water Resources*, 33 (8), 897-905.

Jameson, A. (1991), A comparison of microwave techniques for measuring rainfall, *Journal of Applied Meteorology*, 30 (1), 32-54.

Kolmogorov, A. N. (1941), The local structure of turbulence in incompressible viscous fluid for very large reynolds numbers, in *Dokl. Akad. Nauk SSSR*, vol. 30, pp. 299-303.

Lavallée, D., S. Lovejoy, D. Schertzer, and P. Ladoy (1993), Nonlinear variability and landscape topography: analysis and simulation, *Fractals in geography*, pp. 158-192.

Lavergnat, J., and P. Golé (1998), A stochastic raindrop time distribution model, *Journal of Applied Meteorology*, 37 (8), 805-818.

Lavergnat, J., and P. Golé (2006), A stochastic model of raindrop release: Application to the simulation of point rain observations, *Journal of Hydrology*, 328 (1), 8-19.

Leblois, E., and J.-D. Creutin (2013), Space-time simulation of intermittent rainfall with prescribed advection field: Adaptation of the turning band method, *Water Resources Research*, 49 (6), 3375-3387.

Legorgeu, C. (2013), Amélioration des estimations quantitatives des précipitations à hautes résolutions: comparaison de deux techniques combinant les observations et application à la vérification spatiale des modèles météorologiques, Ph.D. thesis, Clermont- Ferrand 2.

Leijnse, H., R. Uijlenhoet, and J. Stricker (2008), Microwave link rainfall estimation: Effects of link length and frequency, temporal sampling, power resolution, and wet antenna attenuation, *Advances in Water Resources*, 31 (11), 1481-1493.

Lombardo, F., Volpi, E., Koutsoyiannis, D., and Papalexiou, S. M. (2014). Just two moments! A cautionary note against use of high-order moments in multifractal models in hydrology. *Hydrology and Earth System Sciences*, 18(1), 243-255.

Lovejoy, S., and D. Schertzer (1995), Multifractals and rain. new uncertainty concepts in hydrology and hydrological modelling, ed. aw kundzewicz, 62-103.

Lovejoy, S., and Schertzer, D. (2007). Scale, scaling and multifractals in geophysics: twenty years on. In *Nonlinear dynamics in geosciences* (pp. 311-337). Springer New York.

Lovejoy, S., and D. Schertzer (2008), Turbulence, raindrops and the  $11/2$  number density law, *New Journal of Physics*, 10 (7), 075,017.

Lovejoy, S., and D. Schertzer (2010), On the simulation of continuous in scale universal multifractals, part i: spatially continuous processes, *Computers & Geosciences*, 36 (11), 1393-1403.

Lovejoy, S., D. Schertzer, and V. Allaire (2008), The remarkable wide range spatial scaling of trmm precipitation, *Atmospheric Research*, 90 (1), 10-32.

Macor, J., Schertzer, D., and Lovejoy, S. (2007). Méthodes multifractales appliquées à la prévision de pluie en utilisant des données radar. *La Houille Blanche*, (4), 92-98.

Maitra, A., and Chakravarty, K. (2005). Ku-band rain attenuation observations on an earth-space path in the Indian region. *28th URSI-GA*, 23-29.

Mandelbrot, B. B. (1967). How long is the coast of Britain. *Science*, 156(3775), 636-638.

Mandelbrot, B. B. (1974). Intermittent turbulence in self-similar cascades: divergence of high moments and dimension of the carrier. *J. Fluid Mech.*, 62, 331-350.

Mandelbrot, B. B. (1983). *The fractal geometry of nature/Revised and enlarged edition*. New York, WH Freeman and Co., 1983, 495 p., 1.

Mandelbrot, B. B. (2006). Fractal analysis and synthesis of fracture surface roughness and related forms of complexity and disorder. In *Advances in Fracture Research* (pp. 13-17). Springer Netherlands.

Menabde, M., and M. Sivapalan (2000), Modeling of rainfall time series and extremes using bounded random cascades and levy-stable distributions, *Water Resources Research*, 36 (11), 3293-3300.

Meneveau, C., and Sreenivasan, K. R. (1991). The multifractal nature of turbulent energy dissipation. *Journal of Fluid Mechanics*, 224(429-484), 180

Mercier, F., L. Barthès, and C. Mallet (2015), Estimation of fine-scale rainfall fields using broadcast tv-sat links and 4d-var assimilation method, *Journal of Atmospheric and Oceanic Technology*, (2015).

Molini, A., G. G. Katul, and A. Porporato (2009), Revisiting rainfall clustering and intermittency across different climatic regimes, *Water resources research*, 45 (11).

Morin, E., D. C. Goodrich, R. A. Maddox, X. Gao, H. V. Gupta, and S. Sorooshian (2006), Spatial patterns in thunderstorm rainfall events and their coupling with watershed hydrological response, *Advances in Water Resources*, 29 (6), 843-860.

Northrop, P. (1998), A clustered spatial-temporal model of rainfall, in *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 454, pp. 1875-1888, The Royal Society.

Olsson, J., J. Niemczynowicz, and R. Berndtsson (1993), Fractal Analysis of High-Resolution Rainfall Time Series, *J. Geophys. Res.*, 98(D12), 23,265–23,274, doi:10.1029/93JD02658.

Olsson, J. (1995), Limits and characteristics of the multifractal behavior of a high resolution rainfall time series, *Nonlin. Processes Geophys.*, 2, 23–29

Olsson, J. (1998), Evaluation of a scaling cascade model for temporal rainfall disaggregation, *Hydrol. Earth Syst. Sci.*, 2, 19-30.

Onof, C., R. Chandler, A. Kakou, P. Northrop, H. Wheeler, and V. Isham (2000), Rainfall modelling using poisson-cluster processes: a review of developments, *Stochastic Environmental Research and Risk Assessment*, 14 (6), 384-411.

Over, T. M., and V. K. Gupta (1996), A space-time theory of mesoscale rainfall using random cascades, *Journal of Geophysical Research: Atmospheres* (1984-2012), 101 (D21), 26,319-26,331.

Overeem, A., Leijnse, H., and Uijlenhoet, R. (2013). Country-wide rainfall maps from cellular communication networks. *Proceedings of the National Academy of Sciences*, 110(8), 2741-2745.

Pathirana, A., S. Herath, and T. Yamada (2003), Estimating rainfall distributions at high temporal resolutions using a multifractal model, *Hydrology and Earth System Sciences Discussions*, 7 (5), 668-679.

Pecknold, S., S. Lovejoy, D. Schertzer, C. Hooge, and J. Malouin (1993), The simulation of universal multifractals., in *Cellular Automata*, vol. 1, pp. 228-267.

Polonio, R., and Riva, C. (1998). ITALSAT propagation experiment at 18.7, 39.6 and 49.5 GHz at Spino D'Adda: three years of CPA statistics. *Antennas and Propagation, IEEE Transactions on*, 46(5), 631-635.

Rebora, N., and L. Ferraris (2006), The structure of convective rain cells at mid-latitudes, *Advances in Geosciences*, 7, 31-35.

Richardson, L. F. (1922). *Weather prediction by numerical process* Cambridge University Press. Cambridge  
Richardson Weather prediction by numerical process 1922.

Rodríguez, R., M. Casas, and A. Redaño (2013), Multifractal analysis of the rainfall time distribution on the metropolitan area of barcelona (spain), *Meteorology and Atmospheric Physics*, 121 (3-4), 181-187.

Royer, J. F., Biaou, A., Chauvin, F., Schertzer, D., and Lovejoy, S. (2008). Multifractal analysis of the evolution of simulated precipitation over France in a climate scenario. *Comptes Rendus Geoscience*, 340(7), 431-440.

Rupp, D. E., R. F. Keim, M. Ossiander, M. Brugnach, and J. S. Selker (2009), Time scale and intensity dependency in multiplicative cascades for temporal rainfall disaggregation, *Water resources research*, 45 (7).

Rysman, J.-F., S. Verrier, Y. Lemaître, and E. Moreau (2013), Space-time variability of the rainfall over the western mediterranean region: A statistical analysis, *Journal of Geophysical Research: Atmospheres*, 118 (15), 8448-8459.

Schertzer, D., and S. Lovejoy (1987), Physically based rain and cloud modeling by anisotropic, multiplicative turbulent cascades, *J. Geophys. Res.*, 92 (D8), 9693-9714.

Schertzer, D., and S. Lovejoy (1991), Nonlinear geodynamical variability: multiple singularities, universality and observables, in *Non-Linear Variability in Geophysics*, pp. 41-82, Springer.

Schertzer, D. S. Lovejoy, 1997: Universal Multifractals do Exist!: Comments on "A statistical analysis of mesoscale rainfall as a random cascade", *J. Appl. Meteor.* 36, 1296-1303.

Schertzer, D., and S. Lovejoy (2011), Multifractals, generalized scale invariance and complexity in geophysics, *International Journal of Bifurcation and Chaos*, 21 (12), 3417-3456.

Schertzer, D., S. Lovejoy, and D. Lavallée 1993: Generic multifractal phase transitions and self-organized criticality. *Cellular Automata: Prospects in astrophysical applications*, Eds. J.M. Perchang, A. Lejeune, World Scientific, p 216-227.

Schertzer, D., S. Lovejoy, and P. Hubert (2002), An introduction to stochastic multifractal fields, in *ISFMA Symposium on Environmental Science and Engineering with related Mathematical Problems*, pp. 106-179, Higher Education Press, Beijing.

Schleiss, M., A. Berne, and R. Uijlenhoet (2009), Geostatistical simulation of 2d fields of raindrop size distributions at the meso-scale, *Water Resources Research*, 45 (W07415; 2008WR007545).

Schleiss, M., S. Chamoun, and A. Berne (2014), Stochastic simulation of intermittent rainfall using the concept of "dry drift", *Water Resources Research*, 50 (3), 2329-2349.

Schmitt, F. (2014), Continuous multifractal models with zero values: a continuous  $\beta$ -multifractal model, *Journal of Statistical Mechanics: Theory and Experiment*, 2014 (2), P02,008.



Schmitt, F., S. Vannitsem, and A. Barbosa (1998), Modeling of rainfall time series using two-state renewal processes and multifractals, *Journal of Geophysical Research: Atmospheres* (1984-2012), 103 (D18), 23,181-23,193.

Serinaldi, F. (2010), Multifractality, imperfect scaling and hydrological properties of rainfall time series simulated by continuous universal multifractal and discrete random cascade models, *Nonlinear Processes in Geophysics*, 17 (6), 697-714.

Smolarkiewicz, P. K. (1983), A simple positive definite advection scheme with small implicit diffusion, *Monthly Weather Review*, 111 (3), 479-486.

Soubeyroux, J.-M. (2011), Climatologie des épisodes pluvieux en France, *Magazine Météo* No 14, Décembre 2011.

Tessier, Y., S. Lovejoy, and D. Schertzer (1993), Universal multifractals: Theory and observations for rain and clouds, *Journal of Applied Meteorology*, 32 (2), 223-250.

Tessier, Y., S. Lovejoy, P. Hubert, D. Schertzer, and S. Pecknold (1996), Multifractal analysis and modeling of rainfall and river flows and scaling, causal transfer functions, *Journal of Geophysical Research: Atmospheres* (1984-2012), 101 (D21), 26,427-26,440.

Thiria, S., F. Badran, and C. Sorrer (2006), Yao: Un logiciel pour les modèles numériques et l'assimilation de données, Rapport technique, LOCEAN, France.

Tobin, C., L. Nicotina, M. B. Parlange, A. Berne, and A. Rinaldo (2011), Improved interpolation of meteorological forcings for hydrologic applications in a swiss alpine region, *Journal of Hydrology*, 401 (1), 77-89.

Tyralis, H., and D. Koutsoyiannis (2011), Simultaneous estimation of the parameters of the hurst-kolmogorov stochastic process, *Stochastic Environmental Research and Risk Assessment*, 25 (1), 21-33.

Veneziano, D., and V. Iacobellis (2002), Multiscaling pulse representation of temporal rainfall, *Water Resources Research*, 38 (8), 13-1.

Veneziano, D., Furcolo, P., and Iacobellis, V. (2006). Imperfect scaling of time and space-time rainfall. *Journal of Hydrology*, 322(1), 105-119.

Veneziano, D., and P. Furcolo (2009), Improved moment scaling estimation for multifractal signals, *Nonlinear Processes Geophys.*, 16, 641-653, doi:10.5194/npg-16-641-2009.

Veneziano, D., and C. Lepore (2012), The scaling of temporal rainfall, *Water Resources Research*, 48 (8).

Veneziano, D., R. L. Bras, and J. D. Niemann (1996), Nonlinearity and self-similarity of rainfall in time and a stochastic model, *Journal of Geophysical Research: Atmospheres* (1984-2012), 101 (D21), 26,371-26,392.

Veneziano, D., P. Furcolo, and V. Iacobellis (2006), Imperfect scaling of time and space-time rainfall, *Journal of Hydrology*, 322 (1), 105-119.

Verrier, S. (2011), *Modélisation de la variabilité spatiale et temporelle des précipitations à la sub-mésoéchelle par une approche multifractale*, Ph.D. thesis, Université de Versailles-Saint Quentin en Yvelines.

Verrier, S., L. De Montera, L. Barthès, and C. Mallet (2010), Multifractal analysis of African monsoon rain fields, taking into account the zero rain-rate problem, *Journal of hydrology*, 389 (1), 111-120.

Verrier, S., C. Mallet, and L. Barthès (2011), Multiscaling properties of rain in the time domain, taking into account rain support biases, *Journal of Geophysical Research: Atmospheres* (1984-2012), 116 (D20).

Villarini, G., Lang, J. B., Lombardo, F., Napolitano, F., Russo, F., and Krajewski, W. F. (2007). Impact of different regression frameworks on the estimation of the scaling properties of radar rainfall. *Atmospheric Research*, 86(3), 340-349.

Von Hardenberg, J., L. Ferraris, and A. Provenzale (2003), The shape of convective rain cells, *Geophysical research letters*, 30 (24).

Wang, Z., M. Schleiss, J. Jaffrain, A. Berne, and J. Rieckermann (2012), Using markov switching models to infer dry and rainy periods from telecommunication microwave link signals, *Atmospheric Measurement Techniques*, 5 (7), 1847-1859.

Xie, P., and P. A. Arkin (1997), Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs, *Bulletin of the American Meteorological Society*, 78 (11), 2539-2558.

Yaglom, A. (1966), The influence of fluctuations in energy dissipation on the shape of turbulence characteristics in the inertial interval, in *Soviet Physics Doklady*, vol. 11, p. 26.

Zinevich, A., Alpert, P., and Messer, H. (2008). Estimation of rainfall fields using commercial microwave communication networks of variable density. *Advances in water resources*, 31(11), 1470-1480.

Zinevich, A., H. Messer, and P. Alpert (2009), Frontal rainfall observation by a commercial microwave communication network, *Journal of Applied Meteorology and climatology*, 48 (7), 1317-1334



## Communication

**“Simulation of yearly rainfall time series at micro-scale resolution with actual statistical properties: intermittency, scale invariance, rainfall distribution”.**

Akrour N., Aymeric C., Verrier S., Mallet C., Barthes L.

Water Resources Research 2015 (Article journal)

**“Rainfall at Urban Hydrological Scales : Opportunist use of TV-SAT Receivers Network”**

Akrour N., Mercier F., Mallet c., Barthès L. ,

RainGain conference. Juin 2015, Ecole des Ponts ParisTech, France. (poster)

**“Generation of 2D rain maps with realistic properties: methodology and results”.**

Akrour N., Mallet C., Barthès L. Chazottes A.

RainGain conference. Juin 2015, Ecole des Ponts ParisTech, France (présentation orale)

**“A 2D rainfall simulator based on multifractal generator”.**

Akrour N., Mallet C., Barthès L.

European Geosciences Union, Vienne, Autriche. Avril 2015 (poster)

**“The dual beam spectropluviometer”**

Akrour N., Barthès L., Chazottes A., Delahaye J.-Y., Golé P., Lavergnat J., Vinson J.-P.

Journée Scientifique SIRTA, Jun 2014, Palaiseau, France. (poster)

**“Generation of mono-site Rainfall time series from micro to large scale”**

Akrour N., Chazottes A., Verrier S., Mallet C., Barthès L.

Workshop on Stochastic Weather Generators, Avignon, France. Sep 2014, (présentation orale)

**“Simulation of rain time series with zero rainfall and actual statistical distribution in a universal Multifractal framework”**

Akrour N., Chazottes A., Barthès L., Mallet C.

1st Megha Tropiques Ground Validation Workshop on Rainfall Products Validation and Hydrological Applications in the Tropics, Toulouse : France. 2013 (présentation orale)

**“A Multi-fractal rainfall time series generator”**

Akrour N., Chazottes A., Verrier S., Mallet C., Barthes L. journée Scientifique du SIRTA Palaiseau, France. juin2013. (poster)

**“Calibrating synthetic multifractal times series with observed data”**

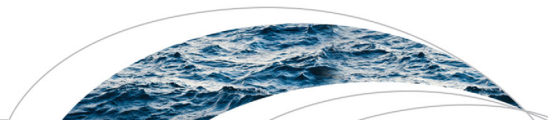
Akrour N., Chazottes A., Barthès L., Mallet C.

11th International Precipitation Conference 2013, Ede-Wageningen Pays-Bas (2013) (poster)



**Article**





## RESEARCH ARTICLE

10.1002/2014WR016357

### Key Points:

- Rainfall generator at fine temporal resolution
- Rain support modeling
- Rain events intensity/duration relationship

### Correspondence to:

N. Akrou, [Nawal.akrou@Latmos.ipsl.fr](mailto:Nawal.akrou@Latmos.ipsl.fr)

### Citation:

Akrou, N., A. Chazottes, S. Verrier, C. Mallet, and L. Barthes (2015), Simulation of yearly rainfall time series at microscale resolution with actual properties: Intermittency, scale invariance, and rainfall distribution, *Water Resour. Res.*, 51, doi:10.1002/2014WR016357.

Received 4 SEP 2014

Accepted 28 JUL 2015

Accepted article online 4 AUG 2015

# Simulation of yearly rainfall time series at microscale resolution with actual properties: Intermittency, scale invariance, and rainfall distribution

Nawal Akrou<sup>1</sup>, Aymeric Chazottes<sup>1</sup>, Sébastien Verrier<sup>2</sup>, Cécile Mallet<sup>1</sup>, and Laurent Barthes<sup>1</sup>

<sup>1</sup>Laboratoire Atmosphère, Milieux, Observation Spatiale, Guyancourt, France, <sup>2</sup>Laboratoire d'Océanographie et du climat, Expérimentation et Approches Numériques, Paris, France

**Abstract** Rainfall is a physical phenomenon resulting from the combination of numerous physical processes involving a wide range of scales, from microphysical processes to the general circulation of the atmosphere. Moreover, unlike other geophysical variables such as water vapor concentration, rainfall is characterized by a relaxation behavior that leads to an alternation of wet and dry periods. It follows that rainfall is a complex process which is highly variable both in time and space. Precipitation is thus characterized by the following features: rain/no-rain intermittency, multiple scaling regimes, and extreme events. All these properties are difficult to model simultaneously, especially when a large time and/or space scale domain is required. The aim of this paper is to develop a simulator capable of generating high-resolution rain-rate time series (15 s), the main statistical properties of which are close to an observed rain-rate time series. We also attempt to develop a model having consistent properties even when the fine-resolution-simulated time series are aggregated to a coarser resolution. In order to break the simulation problem down into subcomponents, the authors have focused their attention on several key properties of rainfall. The simulator is based on a sequential approach in which, first, the simulation of rain/no-rain durations permits the retrieval of fractal properties of the rain support. Then, the generation of rain rates through the use of a multifractal, Fractionally Integrated Flux (FIF), model enables the restitution of the rainfall's multifractal properties. This second step includes a denormalization process that was added in order to generate realistic rain-rate distributions.

## 1. Introduction

Rainfall is a highly complex, naturally occurring phenomenon, which is extremely variable in both time and space. This variability is the consequence of several characteristics, namely intermittency (rain/no-rain), rain and drought extremes, rain-rate variability, and multiple scaling regimes. One of the challenges of rainfall modeling is that of taking all meteorological scales into account, from the synoptic or mesoscales to the microscale. In recent decades, rainfall modeling has been the focus of an intensive research. However, the modeling of rainfall variability still remains an open topic, especially at small scales. In order to correctly describe natural phenomena, in hydrology in particular, nonlinear dynamic models are needed. In this perspective, it is crucial to improve the fine-scale modeling of rainfall events, in order to take this fine-scale information correctly into account.

In this context, several stochastic models have been proposed. The modeler must initially decide whether rainfall will be viewed as a continuous or discrete process; in this context a more thorough review of rain models can be found, for example, in *De Michele and Ignaccolo* [2013]. Once this choice has been made, various approaches can be used. For example *Cooley et al.* [2007] proposed a Bayesian hierarchical model to characterize extreme precipitations events produced by a regional climate model. *Wang et al.* [2012] chose to infer dry and wet periods using Markov switching models. A large number of rainfall generators are based on a pulse/point Poisson process, among others [*Onof et al.*, 2000; *Burton et al.*, 2008; *Evin and Favre*, 2013; *Bernardara et al.*, 2007]. Geostatistical techniques such as kriging have been used to model rainfall [*Schleiss et al.*, 2009, 2014; *Leblois and Creutin*, 2013], and other types of model, such as that of *Tyralis and Koutsoyiannis* [2011], based on a Hurst-Kolmogorov model, have also been used. Among these numerous models, multifractal approaches, especially the Universal Multifractal [UM] one [*Schertzer and Lovejoy*, 1987,



1991; Deidda *et al.*, 1999; Deidda, 2000] have received a great deal of attention in the geophysical literature. Based on an observed scale invariance property called multifractality, these models use the concept of either discrete or continuous multiplicative cascades [e.g., Schertzer and Lovejoy, 1987; Lovejoy and Schertzer, 1995; Lovejoy *et al.*, 2008; Lovejoy and Schertzer, 2010; Serinaldi, 2010; Gaume *et al.*, 2007].

Among the various approaches used to reproduce different scale invariance regimes through the concept of multiplicative cascades, two of these methods are generally applied. The first of these is the “bottom-up” approach, in which the simulation process focuses on the finest scale in an attempt to converge toward large-scale fractal behavior (i.e., from finer to coarser scales). The second is the “top-down” approach, in which the simulation process goes from coarser resolutions to finer resolutions, through the use of fractal cascades to achieve successive downscaling steps. In this case, the rain support could be obtained simply by applying a very small threshold. In practice, rain intermittency is not adequately handled by basic Universal Multifractal (UM) models. As it will be explained in section 3.4, these models are not designed to generate zero values. Various approaches have been investigated, to solve the problem of simulating zero values [Olsson, 1998; Gires *et al.*, 2013; Schmitt *et al.*, 1998]. Several researchers have introduced beta components, i.e., a simple on/off cascade model, in the generation process (see Schertzer *et al.* [2002] for further details). The beta-log-stable process generalizes most of the multifractal scaling models proposed for rainfall simulation. These are also referred to as beta-lognormal or beta-log-Levy models, depending on whether the stability parameter  $\alpha$  is equal to, or is strictly smaller than 2 [e.g., Over and Gupta, 1996; Serinaldi, 2010; Veneziano *et al.*, 2006; Veneziano and Lepore, 2012]. Although this model allows zero values to be generated, it does not provide realistic rain/no-rain durations [Schmitt *et al.*, 1998]. The generalization of this process to continuous cascades, as proposed by Schmitt [2014], provides a more realistic modeling, but it is not able to correct the model’s intrinsic inability to simulate realistically both dry and wet periods.

Although some of the proposed models are able to simulate the multiple scaling regimes (and coherent multifractal properties) of rain observations [e.g., Hubert, 2001; Lovejoy and Schertzer, 2010; Schertzer and Lovejoy, 2011; Serinaldi, 2010], the simulation must reproduce the scaling properties of the three meteorological scales (i.e., synoptic scale, mesoscale, and microscale). These models thus fail to provide simulations that are fully representative of measured rain data, especially to represent correctly all of these scaling regimes (as well as the transitions in between). Furthermore, the statistical distribution of the simulated rain-rate series is not consistent with those directly estimated from rain gauge measurements [e.g., Hubert, 2001].

Another important aspect is the nature of the data set used to build the model. Since large rainfall databases have been created throughout the world, most rain models were developed from rain gauge data with a daily or hourly temporal resolution. Although, as pointed out by Rupp *et al.* [2009], an interesting way to generate fine-scale rainfall on the basis of coarser scale historical records or long-term forecasts is to stochastically disaggregate rainfall from coarser to finer scales, this approach is not considered here. Indeed, in section 4, it is shown that with a (sufficiently long) high-resolution data set, the rain support and average rain rate of the time series used in the present study (see section 2) require specific modeling for periods of time shorter than 5 min. As explained by Veneziano and Lepore [2012] and Gires *et al.* [2013], the simulation of the rain support is crucial, when rainfall time series need to be simulated with scaling properties. These observations suggest that the straightforward disaggregation of rain gauge time series may not be a well-adapted approach.

The simulator developed in the present study is based on the sequential simulation of rain support and intrarain/within-rain variability, using an approach similar to that of some previously cited studies [Schleiss *et al.*, 2014; Leblois and Creutin, 2013]. The rain support is considered as a sequence of alternating, independent periods of rain/no-rain. These generated using two different Pareto laws, proposed in the study of Lavergnat and Golé [1998, 2006] who studied rainfall by analyzing rain droplets. When the rain support is considered to be the result of a two-state renewal process, a simulation of intrarain/within-rain events is required.

In order to generate realistic rainfall time series over a wide range of scales, namely from 15 s to 1 year, we have developed an approach based on a continuous, multiplicative cascade model. Following the concepts proposed by De Montera *et al.* [2009] and Verrier *et al.* [2011], the scaling properties of the rain rates are simulated, through the use of parameters estimated from fine-scale observations. For this, we adopted the

multifractal Fractionally Integrated Flux (FIF) model, which simulates a log-stable stochastic process, based on the Universal Multifractal (UM) model developed by *Schertzer and Lovejoy* [1987].

The aim of this study is to address (and solve) some of the main drawbacks of the model proposed by *De Montera et al.* [2009]. The latter uses a thresholded multifractal FIF model, in which intermittency is introduced by applying an appropriate threshold to the time series. The goal is to assess the main properties of a measured time series, given a limited number of parameters. The improvements proposed here are designed to provide a more realistic simulation, reproducing various important statistical properties of rainfall, in particular its variability, intermittency, and power spectrum, while maintaining a consistent rain-rate distribution at different time scales. As could be expected, these improvements are achieved at the cost of increased complexity in the modeling of realistic rain events.

This paper is organized in six parts. Following the introduction, section 2 presents the data set and its specificities. In section 3, the hypotheses on which the simulator is based are described in detail. For each hypothesis presented in section 3, section 4 provides the corresponding formulation and parameters. Each subsection highlights the coherence between the measured and simulated data series. At the end of section 4, the complete simulation process is summarized. Section 5 discusses the simulator's accuracy, together with various results in support of our initial hypotheses. We thus show that by aggregating the simulated fine-resolution time series, a coarser time series can be retrieved, which is consistent with the aggregated, measured time series. We also demonstrate the ability of the simulation process to reproduce various scaling regimes over the full time series. Our conclusions, together with possible perspectives, are discussed in the last section.

## 2. The Data Set

This study relies on the use of rain-rate (hereafter RR) time series obtained using a disdrometer (more specifically a Dual-Beam Spectropluviometer, hereafter DBS), described in detail by *Delahaye et al.* [2006]. This instrument allows the incoming droplets passing through its capture area to be monitored in terms of their arrival time, diameter, and fall velocity. Each drop is time stamped with an accuracy of 1 ms, thus ensuring good temporal resolution for rain support retrieval. As the capture area of this sensor is a 100 cm<sup>2</sup> rectangle, observations of the rainfall process are extremely localized. The collected drops are then analyzed to estimate the corresponding rain rate, with an integration time of 15 s. Rain studies are highly dependent on the sensor used for rain-rate acquisition, and all data-based models rely on its given resolution, accuracy, and detection threshold. As a consequence, there are almost as many rain models (and in particular, as many parameter sets) as there are types of sensors. Even if they are different, these various sets of parameters may be consistent, such that the simulations obtained with them can have similar properties. Hence, they may be used to reproduce rain-rate series observed at various resolutions, or measured with various accuracies.

Following *De Montera et al.* [2009], the 15 s time resolution was determined from the quality of the scaling restitution. The limits of the power spectrum scaling range are studied, because they indicate the maximum and minimum scales over which it is meaningful to perform multifractal analysis. In the case of the present measurements (not shown here), the spectrum clearly levels off at frequencies above 1/30 Hz, such that the time series must be averaged over 15 s time lags in order to retain the scaling component of the power spectrum only.

In the present study, we have used a data set made up from two time series. The first of these, used as training data, was recorded over a 2 year period (July 2008 to July 2010), and the second (used only in section 5 for validation) was recorded over a period of 2 years and 7 months (from October 2010 to June 2013) at the "Site Instrumental de Recherche par Télédétection Atmosphérique" (SIRTA) in Palaiseau, France. Since the DBS detects droplet velocities, it can easily distinguish between rain and snow, due to considerably lower velocity of snowflakes. Snowy episodes, as well as outliers, were thus removed from the data during the pre-processing of the rainfall time series. As previously described, the properties of the observed rain-rate time series, depend on the one hand, on the sensor's temporal resolution, and on the other hand, on its rain-rate detection threshold. A high detection threshold may overlook weak rain events, and thus overestimate the length of dry periods (hereafter, referred to as no-rain durations). Conversely, a low threshold may not discard small isolated "false detection" raindrops resulting from dust or insects, leading to an underestimation

of the duration of dry periods. In the case of the DBS series, the sensing threshold was set to  $0.05 \text{ mm h}^{-1}$  (for a 15 s integration time), thus ensuring that both the estimated rain support and rain-rate measurements are accurate. Finally, the training data series (resp. validation data series) is composed of  $5.2 \times 10^6$  samples (resp.  $7.6 \times 10^6$ ) recorded with a 15 s resolution (i.e., rain rates integrated over a 15 s time span). In this data set, 4.6% (resp. 5.3%) of the samples are rainy samples, corresponding to more than 1000 h of rain (resp. 1670 h), distributed over 12, 195 (resp. 18, 840) continuous wet periods. Some of the rain events included in the training time series had been used in previous studies [Verrier *et al.*, 2011], to analyze the scaling properties of rain at a fine temporal scale, by implementing a slightly different preprocessing technique.

### 3. From Modeling Toward Simulation

#### 3.1. Introduction

Different studies have shown that rain-rate time series have multiple scaling regimes. In Verrier *et al.* [2011], multifractal analysis of high-resolution data showed that two multiscaling regimes (with different UM parameters) could be distinguished, i.e., from 3 days to 32 min and from 16 min to 15 s. The former is likely to represent interevent variability, whereas the latter is likely to represent the internal variability of a single event (hereafter, referred to as within-rain variability). The position of the break between the two scaling regimes is directly related to the characteristics of the support. Thus, as described by De Montera *et al.* [2009], Veneziano and Lepore [2012], and Gires *et al.* [2012], multiscaling parameters of entire rain time series are found to strongly depend on the rain support characteristics. An event within-rain variability could be perceived as being governed by the laws of atmospheric fluid mechanics [De Montera *et al.*, 2010; Lovejoy and Schertzer, 2008]. Thus, as explained in Verrier *et al.* [2010] and De Montera *et al.* [2009, 2010], consistent values of the UM parameters (resulting from an event within-rain variability) can be expected, in both the spatial and time domains anywhere in the world. The classical (meteorological) definition of rain events includes small values, with relatively short zero periods inside events. These notions are inevitably conditioned by the characteristics of the sensor. Indeed, conventional rain gauges measure integrated quantities and do not allow the transition between rain and no-rain periods to be clearly distinguished. As DBS measurements accurately measure the arrival time of each drop, they make it possible to determine whether or not it is raining within time periods 15 s each. The approach presented in this study is based on the use of DBS measurements, and thus calls for a clear definition of the entangled notions of rain and no-rain events. Here we do not use the classical definition of rain events, but rather the notion of a “continuity” of rain and no-rain periods. To avoid confusion, the term “period” is used in the following. In this context, a “no-rain period” is a series of consecutive, null 15 s rain rates, characterized by their duration  $d_{nr}$ . Similarly, a “rain period” will correspond to a series of consecutive, nonnull, 15 s long rain interval. Each “rain period” is characterized by its duration  $d_r$ , that is, the number of consecutive 15 s non zero rain-rate intervals and its corresponding average rain rate  $\langle RR \rangle$ , which is the average value of the continuous rain rates  $RR$  (integrated over 15 s) belonging to the same “rain period.”

#### 3.2. Main Assumptions

Our simulator was designed on the basis of three main assumptions. The first of these is related to the rain support, whereas the other two are related to the internal properties of rain. The validity of these assumptions with respect to our data sets is discussed in the following (sections 4 and 5).

**Assumption 1:** Adopting the findings of Lavergnat and Golé [2006] and Schmitt *et al.* [1998], we consider the rain support to be a sequence of alternating (and independent) rain  $d_r$  and no-rain  $d_{nr}$  durations.

It is important to note that, in the meteorological sense, whether it be convective or stratiform, a rain event is disaggregated into several rain periods that are likely to be governed by the same dynamics. In accordance with the notions proposed by De Montera *et al.* [2009] and Gires *et al.* [2013], in the present study, we use the same mean multifractal parameters for all rain periods. Furthermore, we assume that these periods differ mainly in terms of their average rain rates  $\langle RR \rangle$ .

**Assumption 2:** The rain’s multifractal behavior occurs within-rain periods, and the FIF model is thus used to generate rain periods, one rain period at a time. In addition to this hypothesis, we consider the rain’s scaling properties to be stationary and independent of the rain period.

Thus, for each rain period, the FIF model makes it possible to generate normalized rain rates  $RR_N$ . A denormalization step is then needed, in order to simulate actual rain rate  $RR$  time series.

**Assumption 3:** To simulate actual rainfall, we assume that, for each rain period, there is a relationship between the rain duration  $d_r$  (resulting from the support simulation) and the corresponding average rain rate  $\langle RR \rangle_{d_r}$ .

Since the duration of a rain period,  $d_r$ , is known when its average rain rate  $\langle RR \rangle$  is drawn, we can also use the notation  $\langle RR \rangle_{d_r}$  to emphasize the rain rate's dependence on the assumed value of rain duration  $d_r$  (assumption 3). Thus, the  $RR$ s are obtained by multiplying the normalized  $RR_N$  values with respect to the average rain rate  $\langle RR \rangle_{d_r}$ .

### 3.3. Rain Support Model (Assumption 1)

When the rain rates are thresholded at a given resolution, the support can be considered to behave as a Boolean process. Here since we are dealing with a one dimensional problem, and it is assumed that the rain duration  $d_r$  and the no-rain duration  $d_{nr}$  are two independent random variables, we can simply draw consecutive (and alternating) rain  $d_r$  and no-rain  $d_{nr}$  durations. Using the definition of rain/no-rain durations given at the end of section 3.1, our definition for rain support is based on the presence or absence of a rain droplet over a time span of 15 s.

Using a disdrometer, allowing accurate time stamping of raindrop times of arrival, *Lavergnat and Golé* [1998] were able to separate interdrop arrival times into two classes: less than 5 min, and more than 5 min. As a consequence, no-rain durations  $d_{nr}$  could also have two different behaviors: one prevailing over periods of less than 5 min, and the other characterized by periods longer than 5 min. Similarly, a rain duration  $d_r$  can have two prevailing behaviors, i.e., less than and more than 5 min [see *Lavergnat and Golé*, 2006, Appendix A]. According to *Lavergnat and Golé* [2006], the distribution of rain/no-rain durations is very close to a Pareto distribution.

### 3.4. Modeling Within a Single Rain Period (Assumption 2)

The core of the rain simulator is a time series generator, based on the implementation of multifractal theory [*Schertzer and Lovejoy*, 1987], which is characterized by scale invariance properties complying with those observed on continuous rain-rate time series. It is commonly accepted that rain can be described in a multifractal framework. When, introduced into the turbulence theory, multiplicative cascades [*Novikov and Stiuart*, 1964; *Yaglom*, 1966] were the conceptual tool used to model stochastic processes and fields having a multifractal behavior. They are defined by hierarchical and iterative methods, which allow the appropriate quantity of energy to be determined at a finer resolution  $\lambda_n$ , on the basis of the energy known at a given initial (coarser) resolution  $\lambda_1$ . Such a cascade is constructed from a large scale, down to smaller scales. After  $n - 1$  steps, the transition from resolution  $\lambda_{n-1}$  to resolution  $\lambda_n$  is obtained from equation (1):

$$\Phi_{(n,i)} = \Phi_{(n-1,j)} \mu\phi_{(n,i)} \quad (1)$$

where the positive field at resolution  $\lambda_{n-1}$  is denoted by  $\Phi_{(n-1,j)}$ ,  $j = 1, \dots, 2^{n-1}$  is the time index of the cascade process, and  $i$  is the time index at the next resolution step. For a given value of  $j$ ,  $i$  has two possible values ( $i=2j-1$  and  $i=2j$ ).

Multiplicative cascades  $\Phi$  are composed of a series of random variables  $\mu\phi_i$  referred to as "multiplicative increments," which are generally assumed to be independent and identically distributed. A discussion of this multifractal model hypothesis is provided in *Veneziano et al.* [2006] and *Serinaldi* [2010]. As the fields generated by such models are scale-invariant multiplicative cascades are widely used to build multifractal fields [*Schertzer et al.*, 2002]. It is generally assumed that the mean value of the process is statistically preserved, regardless of the resolution  $\lambda$ :  $\forall \lambda, \langle \Phi_\lambda \rangle = M$  (where  $M = 1$  in the present case). Whatever the resolution, the statistical properties of the cascade are characterized by a moment scaling function  $K(q)$ :

$$\forall q, \langle \Phi_\lambda^q \rangle \propto \lambda^{K(q)} \quad (2)$$

where  $\langle \cdot \rangle$  is the averaging operator and  $q$  is the order of the moments. In the context of the Universal Multifractal (UM) models proposed by *Schertzer and Lovejoy* [1987], the moment scaling function is defined by two parameters only,  $\alpha$  and  $C_1$ :

$$\begin{cases} K(q) = \frac{C_1}{\alpha-1} (q^\alpha - q) & \text{if } \alpha \neq 1 \\ K(q) = C_1 \cdot q \cdot \ln(q) & \text{if } \alpha = 1 \end{cases} \quad (3)$$

where  $\alpha$  represents the multifractality parameter and can have values between 0 (monofractality) and 2 (log-normality).  $C_1$  provides a measure of the inhomogeneity of the mean level of the field, and for a  $D$ -dimensional process, it lies in the interval  $[0, D]$ .

Contrary to UM fields, most geophysical data are nonconservative. From a conservative UM field  $\Phi_\lambda$ , the Fractionally Integrated Flux (FIF) model [Schertzer and Lovejoy, 1991; Schertzer et al., 2002; Pecknold et al., 1993] computes a nonconservative field noted  $\Phi_\lambda \Delta t^H$ , through a fractional integration of order  $H$ . The fractional integration parameter  $H$  is referred to as the nonconservative parameter. This additional parameter is equivalent to the Hurst exponent and leads to the following statistics for the nonconservative field increments:

$$\langle (\Phi_\lambda \Delta t^H)^q \rangle \sim \lambda^{-\zeta(q)} \quad (4)$$

where  $\Phi_\lambda$  is the conservative field at resolution  $\lambda$ ,  $\Delta t$  is the time increment ( $\Delta t = \frac{T}{\lambda}$ ),  $T$  is the length of the time series,  $\langle \cdot \rangle$  denotes the averaging operator, and  $\zeta(q)$  is the exponent of the FIF structure function of order  $q$  which is given by Vainshtein et al. [1994]:

$$\zeta(q) = qH - K(q) \quad (5)$$

The structure function can be used to determine the parameter  $H$ . For more details concerning the properties of UM and FIF models, see Schertzer et al. [2002] and Schertzer and Lovejoy [2011].

Although a FIF model implies long-range dependence, given assumption 2, this dependence is valid during rain periods only. The FIF model has just one scaling regime and is thus ideal for the modeling of stochastic processes with scaling properties, provided no zero values are involved. Thus, as in De Montera et al. [2009], and as discussed in Veneziano and Lepore [2012], the multifractal parameters were estimated on a per rain period basis (since we work with rain periods rather than events). By working with consecutive, nonnull 15 s rain rates, during rain periods, we ensure that we are working “within rain,” and that the properties and parameters estimated using the rain rates (integrated over 15 s) are not biased by the rain support. The multifractal parameters therefore correctly reflect the variability occurring within a given rain period and not that corresponding to the rain support. As assumed by Verrier et al. [2011], and mentioned by Gires et al. [2013] and Veneziano and Lepore [2012], we believe that the many parameter sets reported in the literature can be largely explained by the fact that they result from a mixture of “within rain” and “rain support” variabilities. Just as it is consistent to estimate a set of FIF model parameters corresponding to each separate rain period, it should also be consistent to simulate the rain-rate time series in the same manner.

### 3.5. Multifractal Denormalization Within a Rain Period (Assumption 3)

The FIF model is basically a multiplicative cascade, based on the use of ratios. It allows the simulation of time series with accurate scale properties but not with actual rain-rate time series, since the UM model works with the relative values  $RR_N$  (introduced previously) such as  $\langle RR_N \rangle = 1$ . Here, the denormalization is carried out, rain period by rain period, at the smaller scale. We thus denormalize each simulated rain period separately.

Once the rain support (and therefore the rain durations) is known, we can “denormalize” each rain period relative values  $RR_N$  to obtain correct rain rates  $RR$ . More precisely, each rain period is denormalized by multiplying the  $RR_N$  by an average rain rate  $\langle RR \rangle_{d_r}$  drawn according to its duration  $d_r$ . By considering the duration  $d_r$  only, when retrieving  $\langle RR \rangle_{d_r}$ , we discard a certain degree of seasonal variability. Since rain events can be viewed as a succession of quiescent and active phases [Ignaccolo and De Michele, 2010], the disaggregation of events into rain periods emphasizes the heterogeneity of rain and therefore weakens the seasonal information contained in the original data. This preference is discussed in section 4.3 in which the corresponding value of  $\langle RR \rangle_{d_r}$  is characterized by a random variable.

## 4. Parameterization of the Simulator

As previously explained, the primary aim of this study is to develop a methodology allowing realistic, synthetic rain-rate time series to be generated over a wide range of scales, i.e., to conserve the statistical

**Table 1.** Percentage of Short Durations Used to Simulate the Rain Support (Column 2) and the Fitted Pareto Parameters (Columns 3–8)

	Probability to Have Short Durations: $Pr(d < 5 \text{ min})$	Pareto Parameters for Short Durations			Pareto Parameters for Long Durations		
		$K$	$\sigma$	$\theta$	$K$	$\sigma$	$\theta$
No-rain durations	0.78	1.56	0.32	0.25	1.88	14.35	4.75
Rain durations	0.87	1.79	0.31	0.25	0.74	7.77	4.75

properties of within-rain periods, while correctly taking the existence of rain and no-rain durations into account. In order to implement this methodology, the various laws and parameters of the proposed simulator (Pareto law for the rain support, multifractal parameters for the within-rain period, and a law further defined for within-rain period denormalization) are estimated based on the training time series described in section 2.

#### 4.1. Rain Support Simulation Parameters (Assumption 1)

The rain support modeling described here is derived from a previous study of *Lavergnat and Golé* [1998, 2006]. These authors observed that the distribution of interdrop time intervals “suggests that the arrival of raindrops may be governed by a point renewal process,” and found that the correlation coefficient between two consecutive arrivals is 0.08. From the data set used, in the present study, we are able to confirm some of these properties. Visual analysis of the joint distribution between two consecutive rain no-rain durations (not shown) shows that these two variables are indeed independent. This is confirmed by the very low correlation coefficient, of 0.003. This result makes it possible to simulate the rain support as a succession of independently drawn, alternating periods of rain and no-rain (assumption 1). For each rain  $d_r$  and no-rain  $d_{nr}$  duration, we first need to randomly draw whether it is a short or a long duration (cf. second column of Table 1). The simulated short/long durations,  $d_r$  and  $d_{nr}$ , are then drawn from a Generalized Pareto law (characterized by equation (6)) using parameters estimated from measured values of  $d_r$  and  $d_{nr}$ .

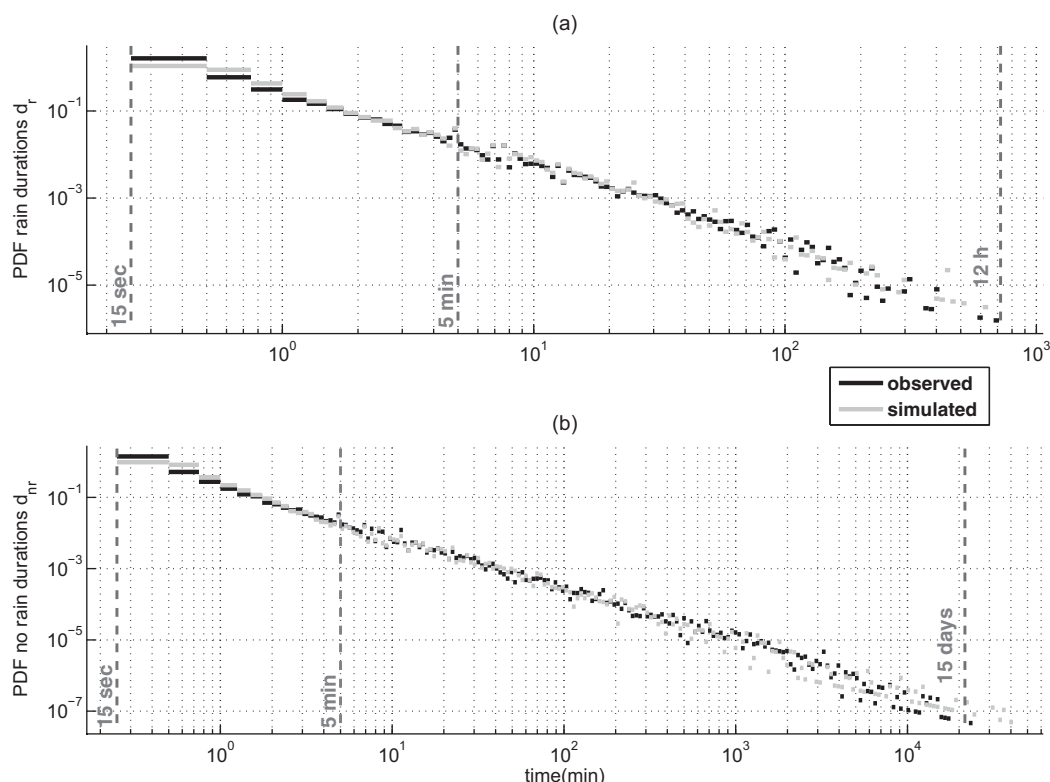
$$f(d/k, \sigma, \theta) = \left(\frac{1}{\sigma}\right) \left(1 + k \frac{(d-\theta)}{\sigma}\right)^{-1-\frac{1}{k}} \quad (6)$$

where  $d$  denotes the duration, i.e., short or long  $d_r$  or  $d_{nr}$ .  $k$ ,  $\sigma$ , and  $\theta$  are the shape, scale, and location parameters, respectively, of the generalized Pareto probability density function. Since the rain intensities are estimated over a 15 s duration, the durations have discrete values. As described in section 2, their distribution is discretized at 15 s intervals, corresponding to the integration time chosen for the calculation of the rain rate. Accordingly, the same discretization is used when simulating the durations. Table 1 shows the Generalized Pareto-distribution parameters estimated from measured, short/long rain/no-rain durations. The location parameter,  $\theta$ , for long durations is set to 4.75 min, since the minimum value of a long duration has to be greater than or equal to 5 min and 0.25 min, i.e., 15 s, for short durations.

The normalized histograms of the simulated rain/no-rain durations are shown in Figure 1. According to our hypothesis, the good agreement between measured and simulated durations leads to a correctly simulated rain support. For each distribution, this is confirmed by a two-sample Kolmogorov-Smirnov test, which does not reject the null hypothesis, with a 99% level of confidence. To highlight this aspect, in section 5.1, it is shown that the scaling properties of the measured series are relatively well simulated.

#### 4.2. Within-Rain Period $RR_N$ Simulation Parameters (Assumption 2)

The FIF model used to simulate the normalized values of a rain period  $RR_N$  is described in section 3.4. This generator was derived from the simulator of *Lovejoy and Schertzer* [2010] which is freely available on the Internet ([www.physics.mcgill.ca/~gang/software/](http://www.physics.mcgill.ca/~gang/software/)). The generator requires three previously defined parameters  $H$ ,  $\alpha$ , and  $C_1$ . We recall that a rain period does not correspond to a real rain event, but rather to consecutive rain-rate values exceeding a certain threshold (see the end of section 3.1). The development of an appropriate methodology for the determination of the multifractal, time-domain properties of rain has been discussed in many studies [e.g., *Pathirana et al.*, 2003; *Molini et al.*, 2009; *De Montera et al.*, 2009]. In particular, *Verrier et al.* [2011] who use a portion of the same data set as the one used here, investigate the scaling properties of a high-resolution data set within the framework of Universal Multifractals. In this study, different approaches (continuous events and continuous series) are compared, the effect of the rain support



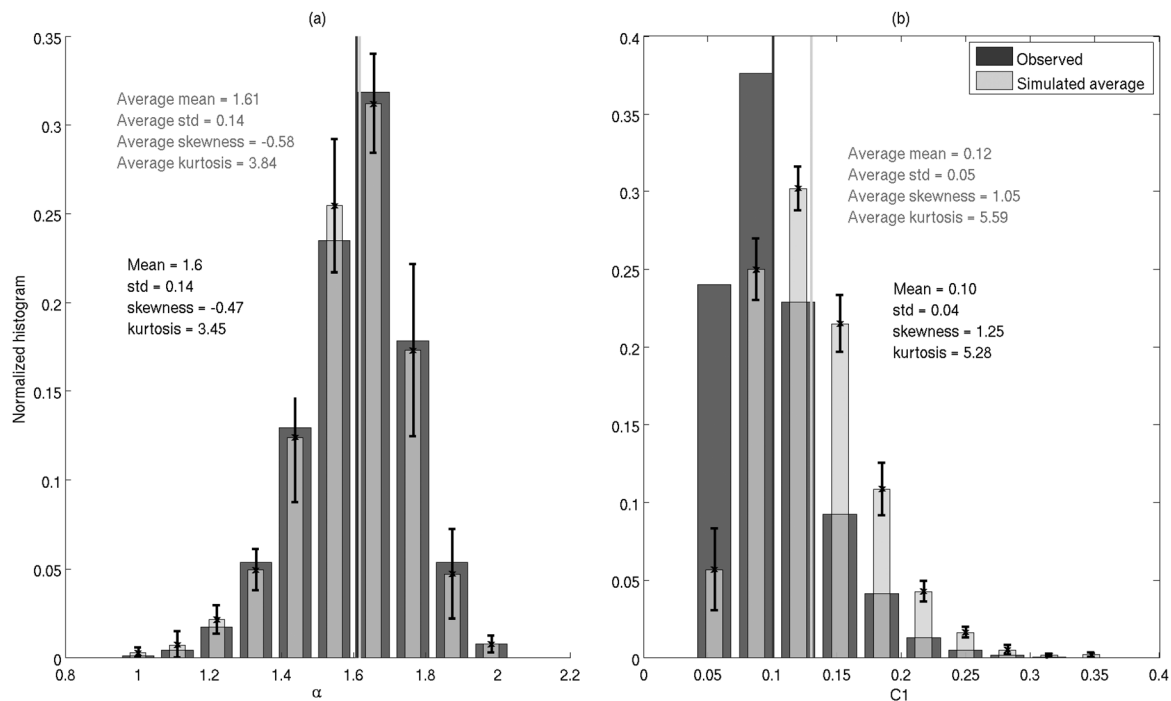
**Figure 1.** Estimated probability density function for (top) rain durations  $d_r$  and (bottom) no-rain durations  $d_{nr}$  for the training measured time series (black) and the simulated time series (grey). Both time series have the same length equal to 2.5 years.

is discussed, and semitheoretical formulae are proposed for the bias introduced by the support zeros. The aim of the present study is not to develop the analysis methodology itself but, rather to demonstrate the simulator's ability to reproduce realistic rain properties. We use the same methodology as that of Verrier *et al.* [2011], which is implemented rain period by rain period, in order to estimate the corresponding parameters.

According to Tessier *et al.* [1996], the fractional integration parameter  $H$  can be estimated by fitting the first-order structure function. From the latter study, it could be conjectured that the rain process is nonconservative at small scales, between 1 min and about 30 min [see Verrier *et al.*, 2011, Figure 8]. The study of the first-order structure function was used to infer the parameter  $H$  corresponding to our data set. A break of the first-order structure function is observed around 32 min. It allows the retrieving of  $H \sim 0.4$  for scales that are shorter than 32 min. To check the consistency of the model, the mean value of the parameter  $H$  was estimated over 32-min long simulated rain periods. The corresponding value was found to be 0.44, which was considered sufficiently close to the value 0.4 retrieved from observed rain-period data.

In order to estimate the parameters  $\alpha$  and  $C_1$ , a classical procedure for the investigation of multifractality (the Trace Moments method) was applied to the differentiated series ( $|\Delta RR_N|$ ) [Lavallée *et al.*, 1993]. For each rain period, the moment scaling function  $K(q)$  was estimated for several given orders  $q$  (16 linearly spaced values going from 0 to 1.5). For each order  $q$ , the value of  $K(q)$  was computed thanks to a power law fitting of equation (2) (as a function of the resolution  $\lambda$ ). Given this set of points, the  $\alpha$  and  $C_1$  parameters were then inferred by a mean square fit of the function  $K(q)$  given in equation (3) (see Tessier *et al.* [1993] for details). This estimation was restricted to rain periods lasting strictly more than 12 min. Since some rain periods did not display scaling, we limited the  $\alpha$  and  $C_1$  estimation to rain periods satisfactorily in agreement with equation (2). This was done by keeping only the rain periods having an average  $R^2$  (of the power law fittings) greater than 0.8. Hence, 9% of the remaining rain periods were discarded.

Although the variation of the estimated parameters  $\alpha$  and  $C_1$  was studied as a function of the rain duration  $d_r$ , this did not produce any significant results. As a result of the lack of long rain-duration data in our time



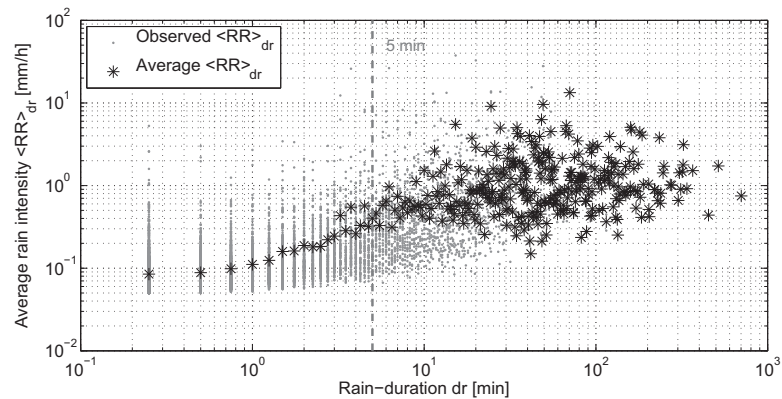
**Figure 2.** Fluctuations of (left)  $\alpha$  and (right)  $C_1$ , estimated for all rain periods extracted from the training measured time series (dark grey), and fluctuations of  $\alpha$  and  $C_1$  (light grey), estimated for a thousand simulated time series of the same length, where  $\alpha=1.6$  and  $C_1=0.1$ . For each bin of the histograms, the error bar corresponding to a standard deviation is also displayed.

series, it is difficult to determine whether there is a relationship between these parameters. The histograms of the parameters estimated from the observed data (the training 2 year long measured time series presented in section 2) are shown in Figure 2. The values of the Fractionally Integrated Flux (FIF) parameters,  $\alpha$  and  $C_1$ , were obtained from our data set as the average of the estimated values, taken over all of the rain periods, that is, 1.6 for the parameter  $\alpha$  and 0.1 for the parameter  $C_1$ . These values are consistent with the parameters estimated in the literature, using similar methodologies. Using this approach, *Pathirana et al.* [2003] estimated the values of  $\alpha$  to lie between 1.09 and 1.92, whereas the value of  $C_1$  was found to lie in the range between 0.26 and 0.42, for several time series. *De Montera et al.* [2009] found  $\alpha=1.69$  and  $C_1=0.13$ , and *Rodríguez et al.* [2013] estimated  $\alpha=1.23$  and  $C_1=0.14$ . There is currently no consensus as to the characteristic parameters of multifractality, and it is not possible to know whether the differences in the values found by various authors are due to the methodologies they adopted, the sensor's characteristics or the natural variability of rainfall.

The estimated mean values are used to implement the simulations, i.e., the multifractal parameters are considered to be constant values. The average histograms of the parameters  $\alpha$  and  $C_1$ , estimated over all the rain periods extracted from a 1002 year long simulated time series, are also shown in Figure 2. The shapes of the  $\alpha$  and  $C_1$  distributions estimated from simulated rain periods are quite close to those estimated from real measurements of rain. The average (mean, standard deviation, skewness, and kurtosis) values for the parameter  $\alpha$  are found to be close to those estimated from the measured data. However, the average distribution of the parameter  $C_1$  for simulated series has a higher kurtosis, suggesting a heavier tail distribution. Similarly, the mean value of  $C_1$  estimated from simulations (i.e., average mean) is greater (0.12) than that estimated from observations (0.1). Since the distribution of the parameters estimated from data simulated with fixed parameters is similar to that obtained from observed data, we can assume that the observed spread over these constant values may be a consequence of estimation errors (stained among others by the limited size of the samples/rain periods). This justifies the choice of constant parameters.

Assumption 2 is validated, since a multifractal behavior is observed in the data corresponding to within-rain periods. However, as there is no consensus on the estimation of parameters characterizing multifractality, a certain degree of uncertainty remains concerning the exact values of the parameters to be retained.





**Figure 3.** Average rain rate of rain periods  $\langle RR \rangle_{d_r}$  as a function of rain duration  $d_r$ , over all the rain periods of the training measured time series (2 years DBS time series).

### 4.3. Rain Period by Rain-Period Denormalizing Parameters (Assumption 3)

At this stage, we are able to generate normalized rain periods corresponding to continuous series of normalized rain rates ( $RR_N$ ) of known duration  $d_r$ . In the following, we examine the usefulness of the rain duration  $d_r$  for the retrieval of the corresponding average rain rate (knowing the duration  $d_r$ ),  $\langle RR \rangle_{d_r}$ . In Figure 3, two behaviors can be observed. When  $d_r$  is shorter than 5 min, a relationship between the average rain rate  $\langle RR \rangle_{d_r}$  and the duration could be distinguished, whereas when  $d_r$  is greater than 5 min, no significant trend is observed. An approach similar to the particularly attractive method proposed by *De Michele and Salvadori* [2003] could be used, in which, an intensity-duration model using a 2-Copula, is combined with a Generalized Pareto. As our data set did not reveal any clear relationship for durations in excess of 5 min, we chose not to use this approach.

As described by *Menabde and Sivapalan* [2000], the values of  $\langle RR \rangle_{d_r}$ , corresponding to  $d_r$ , greater than 5 min, are drawn in accordance with an alpha-stable distribution (determined by its characteristic function—see equation (7)), while assuming there is no dependence between  $\langle RR \rangle_{d_r}$  and the corresponding value of  $d_r$  (given  $d_r > 5$  min). This hypothesis is consistent with our data set and does not go any further. In the present case, it was found that  $\alpha_r \neq 1$  (see Table 2). Since  $\alpha_r$  is not equal to unity, in the following we provide the expression for the corresponding alpha-stable characteristic function only:

$$\varphi(\langle RR \rangle_{d_r} | \alpha_r, \beta, \gamma, \delta) = \exp \left[ -\gamma^{2\alpha_r} |\langle RR \rangle_{d_r}|^{2\alpha_r} \left( 1 - i\beta \tan \left( \frac{\pi\alpha_r}{2} \right) \right) + i\delta \langle RR \rangle_{d_r} \right] \text{ for } \alpha_r \neq 1 \quad (7)$$

where  $\alpha_r \in ]0, 2]$  is the stability parameter and describes the tail of the distribution,  $\beta \in [-1, 1]$  is the skewness parameter,  $\gamma \in ]0, +\infty[$  is the scale parameter, and  $\delta \in ]-\infty, +\infty[$  is the location parameter.

In the case where  $d_r$  is less than 5 min, taking the relationship  $\langle RR \rangle_{d_r} = f(d_r)$  into account does not demonstrate its relevance. Then,  $\langle RR \rangle_{d_r}$  is also drawn using an alpha-stable distribution assuming there is no dependence between  $\langle RR \rangle_{d_r}$  and the corresponding value  $d_r$  (given  $d_r < 5$  min). Finally, we draw  $\langle RR \rangle_{d_r}$  while considering two different sets of parameters, depending on whether  $d_r$  is greater or less than 5 min. As the  $\langle RR \rangle_{d_r}$  distribution is highly asymmetric, the parameter  $\beta$  is set to 1, and the value of  $\delta$  is set to 0, since the average rain intensities are strictly positive values (and hence greater than 0). The only remaining parameters to be determined are  $\alpha_r$  and  $\gamma$ . The values of the estimated parameters are shown in Table 2.

**Table 2.** Parameters of the Conditional (Average Rain Rate  $\langle RR \rangle_{d_r}$ ) Alpha-Stable Distribution Knowing the Rain Duration  $d_r$ .

Rain Duration Lasting Less Than 5 min				Rain Duration Lasting 5 min or More			
$\alpha_r$	$\beta$	$\gamma$	$\delta$	$\alpha_r$	$\beta$	$\gamma$	$\delta$
0.90	1	0.01	0	0.77	1	0.16	0

One of the specificities of this study is its use of a finely resolved data set (15 s resolution). Our definition of rain duration (a set of consecutive, 15 s nonzero rain rates) has various implications. When compared to other sensors, we may have an increased number of short rain periods, and therefore a lower number of medium/long rain periods. In practice, our

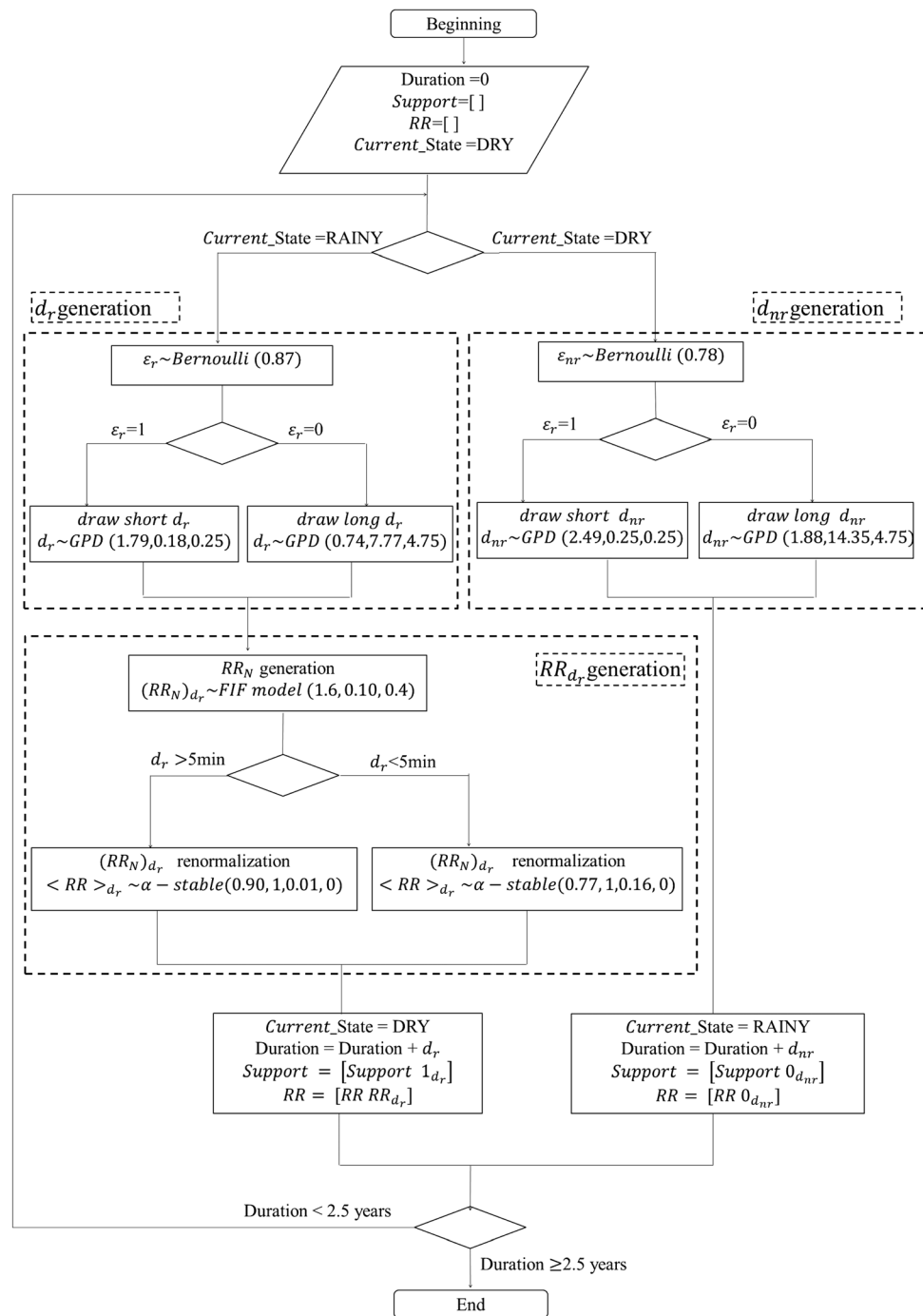


Figure 4. Scheme of the simulation process.

definition of a rain period generally corresponds to portion of what is usually referred to as a rain event. Thus, in a certain sense, we split up single rain events and shuffle the resulting parts before establishing a relationship between  $\langle RR \rangle_{d_r}$  and  $d_r$ . This process is inherent to the fine-resolution modeling of rain and allows rain rates resulting from a short burst of raindrops to be simulated.

As further discussed in section 5.1, Figure 7c compares the water height Empirical Survival Functions (ESF), obtained for 100 simulated time series at the finest resolution, with the two observed ones. This allows us to validate the third assumption.

#### 4.4. The Simulator

In this section, we provide a detailed description of the simulator (also presented in the scheme in Figure 4). The simulation procedure consists in the following two main steps:

**Step 1:** Simulation of the rain support.

As previously mentioned, no-rain durations ( $d_{nr}$ ) and rain durations ( $d_r$ ) are independent. The rain support is thus obtained by alternating rain and no-rain durations until the simulated time series reaches the desired length. The short (<5 min) or long (>5 min) values of  $d_{nr}$  and  $d_r$  are drawn randomly, in accordance with their expected occurrences, such that each rain or no-rain duration is simulated according to the following procedure:

- Draw a random Bernoulli variable  $\epsilon$  simulating the probability of short  $d_{nr}/d_r$  (see Table 1, column 2).
- If  $\epsilon = 1$ , draw a short  $d_{nr}$  or  $d_r$  according to a Generalized Pareto law (equation (6)), using the parameters provided in Table 1 (columns 3–5).
- If  $\epsilon = 0$ , draw a long  $d_{nr}$  or  $d_r$  according to a Generalized Pareto law (equation (6)), using the parameters provided in Table 1 (columns 6–8).

The  $d_{nr}$  and  $d_r$  values are rounded to the closest multiple of 15 s.

**Step 2:** Simulation of the rain periods.

For each rain period of duration  $d_r$ , we sequentially implement the FIF model, in order to generate normalized  $RR_N$  values and then, by multiplying the  $RR_N$  by the randomly drawn  $\langle RR \rangle_{d_r}$ , apply a denormalization step to obtain the corresponding rain rates  $RR$  ( $RR = RR_N \times \langle RR \rangle_{d_r}$ ).

For each rain duration  $d_r$ :

- Using the FIF model, generate series of length  $d_r$  of  $RR_N$  normalized values with the corresponding parameters ( $\alpha=1.6$ ,  $C_1=0.1$ , and  $H=0.4$ ).
- Draw a stable random value  $\langle RR \rangle_{d_r}$ , using the appropriate parameters provided in Table 2, to denormalize  $RR_N$  and obtain the corresponding  $RR$  (and thus the simulated rain period).

The process defined above allows rain-rate time series consistent with our data set to be simulated but is representative of a specific (medium latitude) meteorological region only. However, this process can also be adapted to other regions with different specific properties. The multifractal parameters can be considered to be universal, i.e., characterizing a mean physical relationship within an event. The remaining parameters characterizing the support and the rain duration/average rain-rate relationship can be estimated or drawn based on a physical knowledge (or hypotheses), in order to simulate rain-rate time series in other locations governed by different climatic conditions.

## 5. Characteristics of the Simulated Times Series

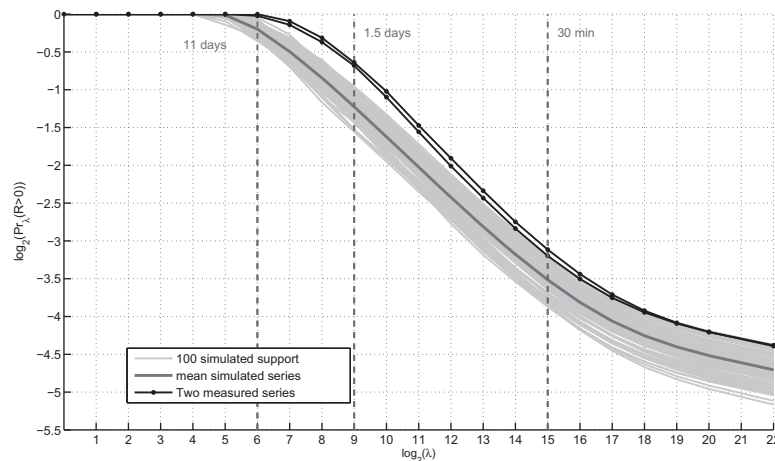
At this stage, the simulator has been partially validated by systematically controlling each hypothesis of the model. The next step involves additional controls, through the investigation of various properties of the entire rain-rate time series. This also allows us to check the model's ability to restore rain properties that are not directly used in the model. In this section, in order to assess the correct behavior of the simulator, we used 100 simulated time series having a 2.5 year duration and 15 s resolution. These were then compared with the two original time series (see section 2).

### 5.1. Scaling Properties of Rain Series

To illustrate the simulator's ability to generate realistic rainfall time series, we studied the fractal properties of the 100 simulated rain supports. In various studies, the rain support is assumed to be fractal. This is one of the reasons for which *Over and Gupta* [1996], *Schmitt et al.* [1998], *Serinaldi* [2010], *Veneziano et al.* [2006], and *Veneziano and Lepore* [2012] (among others) used the beta model to generate rain supports. This property implies that the rain support behaves according to the following relationship:

$$Pr(RR_\lambda > 0) \propto \lambda^{-C_f} \quad (8)$$

where  $\lambda$  denotes the resolution,  $C_f$  denotes the codimension, and  $RR_\lambda$  is the rain rate at scale  $\lambda$ . To validate the rain support simulation, it is also necessary to verify that the fractal nature of the resulting support is



**Figure 5.** Probability of rain occurrence at a given resolution  $Pr(RR_\lambda > 0)$  for the two measured time series (black solid + circle) and for a 100 simulated 2 year rain supports (grey lines) as a function of resolution  $\lambda$ . The average of the 100 simulated  $Pr(RR_\lambda > 0)$  curves is also shown (dark grey).

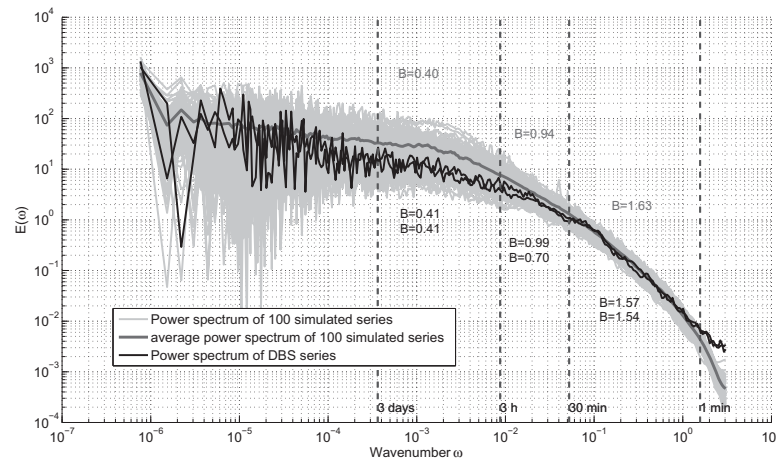
preserved. Figure 5 shows the probability of rain occurrence ( $Pr(RR_\lambda > 0)$ ), for a given resolution  $\lambda$ . This relationship is shown on a log-log scale. The rain occurrence of the two measured rain supports (black solid curve), as well as the occurrences of the 100 simulated supports (grey curves), is characterized by fractal behavior between the scales  $\lambda = 2^{15}$  and  $\lambda = 2^9$  (corresponding to time scales ranging between 30 min and 1.5 days). The linear behavior of this log-log plot corresponds to fractal behavior, between 30 min and 1.5 days, for both the measured and simulated supports. The slopes of the log-log fits to the two measured time series correspond to estimated codimensions  $C_f$  equal to 0.42 and 0.41, respectively (see Table 3), whereas the 100 simulated supports lead to a mean value of 0.38. The simulated and measured codimensions are thus very close.

The 100 simulated supports (at 15 s) have a mean probability of rain occurrence equal to 3.85% as opposed to the values of 4.6% and 5.31% observed for the two measured rain supports. Although the slopes of the linear portion are rather similar, there is nevertheless a noticeable offset between the two plots. At the finer resolution ( $\lambda = 2^{22}$ , corresponding to an integration time of 15 s), the two (black) measured curves lie above the 100 (grey) simulated curves. The simulated  $Pr(RR_\lambda > 0)$  is lower than expected. This implies a lack of rain support (or an excess of zeros/no-rain periods). For fine scales (from  $\lambda = 2^{22}$  to  $\lambda = 2^{12}$ ), apart from an offset, all of the curves are similar in shape.

The offset between simulations and measurements can be explained by an excess of longer no-rain  $d_{nr}$  durations. As it can be seen (see Figure 1), there is a tendency to slightly over sample the simulated long no-rain  $d_{nr}$  durations. In Figure 5, for  $\lambda = 2^6$  (corresponding to approximately 11 days), whereas the two observed curves reach the value 0 (probability of rain occurrence = 1, at resolution  $\lambda$ ) the dark grey curve representing the average values of the simulated curves does not reach this value. This outcome means that the model is able to simulate no-rain durations  $d_{nr}$  that can be longer (or slightly more frequent) than those measured during real observations. It explains the offset found at the resolution  $\lambda = 2^6$  as well as at finer scales (until resolution  $\lambda = 2^{22}$ ). This comparison between the properties of simulated/measured rain supports provides a posteriori validation of the relevance of assumption 1. Furthermore, it is

	Observed DBS Series		100 Simulated Series			
			Mean	$Q_1$	$Q_2$	$Q_3$
Codimension	0.42	0.41	0.38	0.39	0.38	0.37
Percentage of rain	4.6	5.31	3.85	3.51	3.85	4.15

<sup>a</sup>Mean, first, second, and third quartiles (columns 4–7) of the codimension and the percentage of rain (at the finer resolution), estimated for the corresponding 100 simulated time series.



**Figure 6.** Power spectra of the two measured time series (thin, black) and average power spectrum (thick, dark grey), computed from a 100 simulated time series spectra (also indicated by light grey).

important to notice that for measurements with a resolution finer than 30 min, the scaling behavior of the support is broken. In this case, a disaggregation model based on a monofractal cascade would not function correctly. Here since there is no scaling of the support under 30 min, a beta-log-stable UM model (coarsely a multifractal process generated over a fractal support) would not provide a correct distribution of rain occurrences for resolutions finer than 30 min. For this kind of UM model, the rain support can influence the multifractal parameter estimation (through an inappropriate mixing of rain support and “within-rain” properties).

Figure 6 provides a log-log plot of the power spectra of the two measured time series, and of the 100 simulated time series. The average spectrum is also shown (wide, dark grey line). While reproducing a certain variability (light grey area), the mean spectrum of the 100 simulated time series is consistent with measurements. We can only hope that the measurements behave as a sample among the others, which they do. Indeed, for the full time span covered by these time series, the corresponding spectra have similar scaling factors, over coherent time scales. As reported in previous studies [e.g., *Fraedrich and Larnder, 1993; Tessier et al., 1996; De Montera et al., 2009; Schertzer and Lovejoy, 2011; Verrier et al., 2011*], the scaling properties between 1 min and 3 days are analyzed by partitioning the full range into a set of different scaling regimes. The scaling regime corresponding to more than 3 days is strictly due to the rain support, and that equal to 30 min characterizes within-rain periods, the central part of which corresponds to a transition stage. These scaling regimes are characterized by the slope value  $B$ , which is used rather than the traditional  $\beta$ , to avoid confusion with  $\beta$  used here to indicate the stable characteristic function of equation (7).

The “within-rain” slopes (1–30 min) of the simulated power spectra have a value  $>1$ , indicating a nonconservative field; this is thus consistent with the parameter  $H$ , which is equal to 0.4. The slope  $B$  of the mean spectra is found to be equal to 1.63 (see Table 4), which is very close to the two values estimated from the measured DBS series (1.57 and 1.54, respectively).

The scaling properties introduced by the simulation of rain periods (generally  $<30$  min) clearly explain the first slope (for time scales shorter than 30 min) resulting from the use of the FIF model. The consistency of the scaling properties obtained for the entire time span (Table 4) necessarily results from the rain support properties (for time scales greater than 30 min). In other words, from a spectral point of view, the good agreement between the measurements and the simulations validates the rain support simulation (independent rain/no-rain durations: assumption 1), as well as the within-rain-period simulation (rain-rate series simulated with the FIF model: assumption 2).

In this subsection, it has been shown that the rain support influences the power spectrum in the case of low frequencies (periods longer than 30 min), and that care should be taken when estimating/explaining scaling properties over such periods of time.

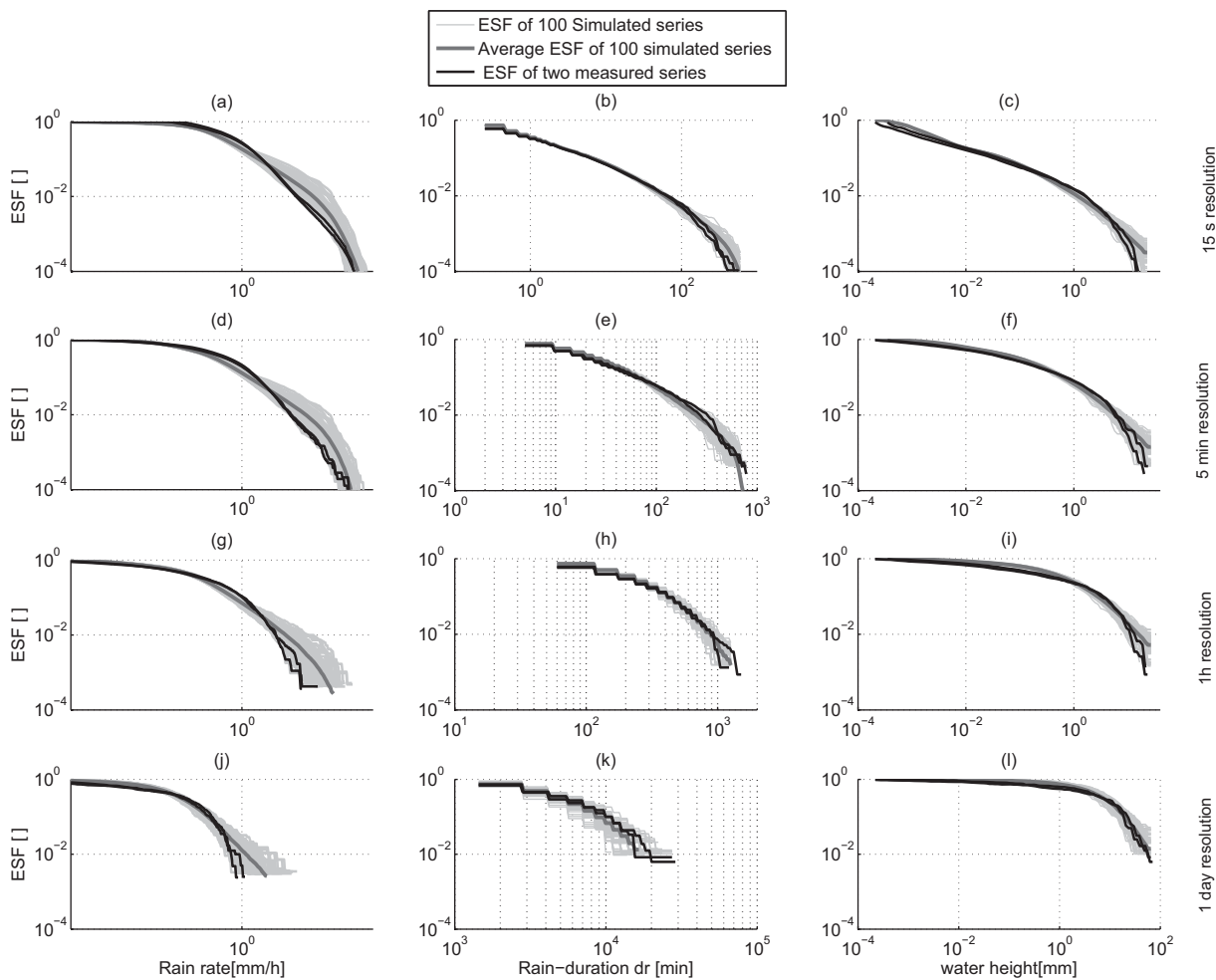
**Table 4.** Slope of the Power Spectrum Obtained With the Two Measured Time Series for the Three Scaling Regimes (Line 2)<sup>a</sup>

	B in Scaling Regime 1 min → 30 min				B in Scaling Regime 30 min → 3 h				B in Scaling Regime 3 h → 3 days			
	Mean	Q <sub>1</sub>	Q <sub>2</sub>	Q <sub>3</sub>	Mean	Q <sub>1</sub>	Q <sub>2</sub>	Q <sub>3</sub>	Mean	Q <sub>1</sub>	Q <sub>2</sub>	Q <sub>3</sub>
DBS series		1.57	1.54		0.99	0.70			0.41	0.41		
100 simulated series	1.63	1.57	1.62	1.66	0.94	0.74	0.93	1.10	0.40	0.22	0.39	0.48

<sup>a</sup>Mean, first, second, and third quartiles of the slope estimated for the 100 simulated spectra for three different scaling regimes.

**5.2. More Classical Properties of Rain**

Beyond the multifractal properties of rain, it is important to verify that the simulated time series are consistent with the behavior that could normally be expected with rainfall. It is commonly studied through the Empirical Survival Function (ESF hereafter). Since we are dealing with rain periods rather than rain events, the ESF is computed for three characteristic quantities: rain rates  $RR$ , rain duration  $d_r$ , and rain-duration water heights (referred to as water height in the following). It is important to verify that the scaling properties do in practice lead to simulated time series that are consistent with measurements made at coarser resolutions. This can be assessed by aggregating all the 15 s time series, to obtain the ESF at four commonly used time resolutions (5 min, 1 h, and 1 day). As shown in the previous section, we use the two measured time series, as well as the 100 simulated time series (spanning 2.5 years), with a resolution of 15 s.



**Figure 7.** (left to right) Empirical Survival Functions of rain rates  $RR$ , rain durations  $d_r$  (i.e., length of consecutive nonzero rain rates), and the (corresponding) water height. (top to bottom) The ESFs are shown for the high-resolution series (15 s), and for the 5 min, 1 h, and 1 day aggregated series.

For each plot shown in Figure 7, we have drawn the ESF curves corresponding to the measured time series (the two black curves), and the curves representing the 100 simulated time series (light grey curves), together with the mean value of the latter (dark grey curve). Each row of plots in Figure 7 corresponds to the same value of resolution, shown in the right of the figure, and each column corresponds to a rain property indicated at the bottom of the figure. The plots shown in the first row thus correspond to the finest, 15 s resolution (i.e., that used for the measurements and the simulations). The plots in the other three rows show the aggregated ESFs of the simulated or measured 15 s time series, corresponding to coarser temporal resolutions (5 min, 1 h, and 1 day).

This figure shows that the ESF computed from simulated data still remains in agreement with those obtained from measured data, which means that temporal scale relations and rain intermittency are well represented in our model. From that perspective, it is reasonable to consider the measured time series as a representative realization of the simulation process. The good behavior of the aggregated series (at all different resolutions) validates the simulator's ability to generate coherent time series, for all values of temporal resolution. As expected, larger values are characterized by a stronger variability, due to their lower frequency of occurrence. The occurrence of extreme values is thus in agreement with the measurements and hypotheses used to build the model.

The rain-duration plots shown in the second column (Figures 7b, 7e, 7h, and 7k) confirm the information revealed in Figure 5. This provides a different way of viewing the rain support, and validates assumption 1. Despite its tendency to slightly overestimate the short durations, the simulated rain-duration ESFs are very close to the two measured ESFs. For larger values, a difference can be noted between the behaviors of the rain-duration ESFs. A slight bend can also be seen in one of the measured curves in Figure 7e—as previously discussed, this type of effect can be accounted for by the natural variability of the simulation.

The ESFs shown as a function of rain-duration water height in the third column (Figures 7c, 7f, 7i, and 7l) are proportional to the product of the rain duration and the average rain rate:  $\langle RR \rangle_{d_r} \cdot d_r$ . Since the consistence of the simulated rain support has been demonstrated, this can provide an additional method for validation of the relationship between  $\langle RR \rangle_{d_r}$  and  $d_r$  (assumption 3). Although the 15 s ESFs corresponding to the mean of the 100 simulations are not as smooth as the original ESFs, the simulations show that these are consistent with the measurements. When 2 year time series and their aggregated counterparts are considered, there is no apparent consequence in term of the dependencies that may have been suppressed by the simulation process between other rain characteristics.

The four rain-rate plots shown in the first column (Figures 7a, 7d, 7g, and 7j) represent rain-rate ESFs. These result from the combined use of our three assumptions, and it can be seen that there is also a good correspondence between the measured and simulated ESFs at all resolutions.

As a result, Figure 7 shows that the simulated time series provide a very good representation of the information contained in the originally measured data (two series of 2 year periods), at several commonly used resolutions. A more thorough investigation accounting for seasonal trends and long-range dependences will be investigated when these data become available. Nevertheless, the simulator could be run over longer time spans, in order to study the occurrence of potentially extreme values, and their corresponding return periods.

## 6. Conclusion and Discussion

Guided by previous studies [Lavergnat and Golé, 1998, 2006; De Montera et al., 2009, 2010; Verrier et al., 2010, 2011], the main objective of this research has been to simulate rain time series with statistical properties (distribution and scaling invariance) representative of a particular climatic region. Our methodology is based on two main steps, namely the simulation of rain support, and of realistic rain rates. These steps are based on several important hypotheses, relying mainly on the high resolution of the data set. These supporting hypotheses are thoroughly validated. The originality of this study is that a very good agreement is established between measurements and simulations, for a significant number of properties that coherently characterize rain. When necessary, the validation was systematically run over the full time series, rather than for selected events only. The resulting simulated fine-resolution time series incorporates information that is coherent with the various time scales of the original measurements. When aggregated over a duration of

up to 1 day, the simulated time series preserves its coherent features (rain-rate distributions in particular) at various resolutions.

This study shares similarities with other studies, such as that of *Veneziano and Lepore* [2012], dealing with the simulation of time series with scaling properties. The difference in resolution between our measurements (15 s resolution) and those used in the aforementioned study (1 h resolution) explains some of the apparent discrepancies. As previously discussed and demonstrated by *Veneziano and Lepore* [2012], we emphasize the importance of rain support and explain how this can introduce a bias into the estimated multifractal parameters. As described in section 3.2, the fine resolution of the time series allows the influence of the rain's intermittency (zero of the rain support) on the rain-rate values to be cancelled. This corroborates assumption 2, namely the independence of rain support properties with respect to the multifractal/scaling of within-rain parameters. Our model successfully retrieves a simulated power spectrum (Figure 6) showing scaling breaks consistent with the original data set. This is not commonly achieved with standard UM models. Although our study does not focus on the retrieval of multifractal parameters, we can try to discuss this. As a consequence of their coarse time resolution, rain gauges do not allow within-event rain-rate scaling properties to be correctly discriminated from intermittency. Since numerous multifractal models are based on parameters estimated through the processing of rain gauge time series, the extensive set of parameters (some of which are presented at the end of section 4.2) can be explained by this mixing. *De Montera et al.* [2009] showed that coherent multifractal parameters are found in three different climatic regions. We believe that this set of parameters could be universal or at least considerably reduced through the use of fine-resolution sensors. In other words, as suggested by *Gires et al.* [2013], we believe that these parameters vary from one geographical location to another through the zero values of the various rain supports. Among others, *Veneziano and Lepore* [2012] determined a value of 2 for parameter  $\alpha$ , whereas our study led to a value of 1.6. Since their measurements are integrated over 1 h, they can be considered to smooth out the rain rate. Indeed, the 1 h measurements result from mixing fine-resolution rain periods with fine-resolution zeros (corresponding to the no-rain durations). This interpretation would explain the greater value of  $\alpha$  characterizing less sparse time series (less small values and less spikes).

Although rain event modeling is considered to be universal, the parameters describing rain support and rain duration/average rain-rate relationship can fluctuate from one geographical area to another. The use of parametric methods, associated with reasonable hypotheses and expert knowledge, can be very beneficial. Finally, it is important to note that there is a relatively high cost associated with the design and development of a simulator with characteristics (and especially those of rain support) close to those of the "real world." Indeed 25 parameters were represented: 14 to model the rain support, 3 to represent the within-rain variability, and 8 for the denormalization process. However, among these parameters, some have fixed values and some others are not expected to change significantly. Indeed, if it is assumed that short durations and within-rain parameters result from the laws of physics governing rain, the parameter requirement could be simplified by the following: eight parameters for rain support modeling and three FIF parameters could be considered as "global" or "universal." Four of the rain duration/average rain rate relationship parameters are fixed, whatever the climatic region (due mainly to the positivity of the rain rate). As the four remaining parameters related to the duration/average rain-rate relationship were not studied with another data set, it is currently difficult to gain a clear understanding of their variability. Finally, at maximum of 10 parameters need to be evaluated, in order to adapt the simulator to a different climatic region. Ideally, a time series with a 15 s resolution would be required to correctly estimate these. However, they could, in practice, be estimated from a series with coarser resolution. Typically, a 1 or 5 min time series should suffice, since the within-rain properties are not expected to change significantly from one place to another.

A question can be raised on the specific behavior for the rain support below 5 min identified in this work. Could this behavior be common to all rain events, whatever the location?

In principle, this simulation/validation approach could be applied to other fine-resolution time series. As rain modeling is considered to vary from one location to another, some of our hypotheses and parameters could be expected to vary. Thus, to improve the modeling of rainfall at different geographic locations, it would be helpful to have an improved understanding of their differences and/or similarities.

In view of the limited time span covered by our time series (2 years), it is not possible to draw special attention to the seasonality of the results. Furthermore, we were able to use a coarse resolution relationship only,



between the average rain rate and the rain duration. To improve these aspects of our model, it would be necessary we need to study a longer time series. As we do not currently have longer fine-resolution time series, it would be useful to compare our results with those obtained using other sensors for which longer series are available. An intercomparison of sensors (typically the DBS and a conventional tipping bucket rain gauge) would allow us to gain a better understanding of the influence of the sensor's resolution and sensitivity on the accuracy of rain property retrieval. At a resolution coarser than 1 h, although both the rain gauge and the DBS disdrometer give consistent water volumes, discrepancies appear when finer resolutions are considered. A better understanding of this should help us improve the simulator and/or test its accuracy over longer periods of time. This would lead to longer simulated series, and thus give access to statistics over longer time spans (providing information, on less frequent extreme events). The study of series over longer periods of time may contribute toward emphasizing any seasonal trend in the rain series. It could also invalidate the independence of no-rain and rain durations, and therefore lead to enhanced modeling of the rain support.

### Acknowledgments

The authors thank the two reviewers who contributed their judicious and helpful remarks. This study was partially supported by the French "Programme National de Télédétection Spatiale" (PNTS, <http://www.insu.cnrs.fr/actions-sur-projets/pnts-programme-national-de-teledetection-spatiale>), grant PNTS-2013-01. The data set used in this study is available, on demand, from the authors.

### References

- Bernardara, P., C. De Michele, and R. Rosso (2007), A simple model of rain in time: An alternating renewal process of wet and dry states with a fractional (non-Gaussian) rain intensity, *Atmos. Res.*, *84*(4), 291–301.
- Burton, A., C. Kilsby, H. Fowler, P. Cowpertwait, and P. O'Connell (2008), RainSim: A spatial-temporal stochastic rainfall modelling system, *Environ. Modell. Software*, *23*(12), 1356–1369.
- Cooley, D., D. Nychka, and P. Naveau (2007), Bayesian spatial modeling of extreme precipitation return levels, *J. Am. Stat. Assoc.*, *102*(479), 824–840.
- De Michele, C., and M. Ignaccolo (2013), New perspectives on rainfall from a discrete view, *Hydrol. Processes*, *27*(16), 2379–2382.
- De Michele, C., and G. Salvadori (2003), A generalized pareto intensity-duration model of storm rainfall exploiting 2-copulas, *J. Geophys. Res.*, *108*(D2), 4067, doi:10.1029/2002JD002534.
- De Montera, L., L. Barthès, C. Mallet, and P. Golé (2009), The effect of rain-no rain intermittency on the estimation of the universal multifractals model parameters, *J. Hydrometeorol.*, *10*(2), 493–506.
- De Montera, L., S. Verrier, C. Mallet, and L. Barthès (2010), A passive scalar-like model for rain applicable up to storm scale, *Atmos. Res.*, *98*(1), 140–147.
- Deidda, R. (2000), Rainfall downscaling in a space-time multifractal framework, *Water Resour. Res.*, *36*(7), 1779–1794.
- Deidda, R., R. Benzi, and F. Siccardi (1999), Multifractal modeling of anomalous scaling laws in rainfall, *Water Resour. Res.*, *35*(6), 1853–1867.
- Delahaye, J.-Y., L. Barthès, P. Golé, J. Lavergnat, and J.-P. Vinson (2006), A dual-beam spectropluviometer concept, *J. Hydrol.*, *328*(1), 110–120.
- Evin, G., and A.-C. Favre (2013), Further developments of a transient Poisson-cluster model for rainfall, *Stochastic Environ. Res. Risk Assess.*, *27*(4), 831–847.
- Fraedrich, K., and C. Larnder (1993), Scaling regimes of composite rainfall time series, *Tellus, Ser. A*, *45*(4), 289–298.
- Gaume, E., N. Mouhous, and H. Andrieu (2007), Rainfall stochastic disaggregation models: Calibration and validation of a multiplicative cascade model, *Adv. Water Resour.*, *30*(5), 1301–1319.
- Gires, A., I. Tchiguirinskaia, D. Schertzer, and S. Lovejoy (2012), Influence of the zero-rainfall on the assessment of the multifractal parameters, *Adv. Water Resour.*, *45*, 13–25.
- Gires, A., I. Tchiguirinskaia, D. Schertzer, and S. Lovejoy (2013), Development and analysis of a simple model to represent the zero rainfall in a universal multifractal framework, *Nonlinear Processes Geophys.*, *20*, 343–356, doi:10.5194/npg-20-343-2013.
- Hubert, P. (2001), Multifractals as a tool to overcome scale problems in hydrology, *Hydrol. Sci. J.*, *46*(6), 897–905.
- Ignaccolo, M., and C. De Michele (2010), A point based Eulerian definition of rain event based on statistical properties of inter drop time intervals: An application to Chibolton data, *Adv. Water Resour.*, *33*(8), 897–905.
- Lavallée, D., S. Lovejoy, and D. Schertzer (1993), Nonlinear variability and landscape topography: Analysis and simulation, in *Fractals in Geography*, edited by L. De Cola and N. Lam, pp. 158–192, Prentice Hall, Englewood Cliffs.
- Lavergnat, J., and P. Golé (1998), A stochastic raindrop time distribution model, *J. Appl. Meteorol.*, *37*(8), 805–818.
- Lavergnat, J., and P. Golé (2006), A stochastic model of raindrop release: Application to the simulation of point rain observations, *J. Hydrol.*, *328*(1), 8–19.
- Leblois, E., and J.-D. Creutin (2013), Space-time simulation of intermittent rainfall with prescribed advection field: Adaptation of the turning band method, *Water Resour. Res.*, *49*, 3375–3387, doi:10.1002/wrcr.20190.
- Lovejoy, S., and D. Schertzer (1995), Multifractals and rain, in *New Uncertainty Concepts in Hydrology and Hydrological Modelling*, edited by Z. W. Kundzewicz, pp. 62–103, Cambridge Univ. Press, Cambridge.
- Lovejoy, S., and D. Schertzer (2008), Turbulence, raindrops and the 11/2 number density law, *New J. Phys.*, *10*(7), 075017.
- Lovejoy, S., and D. Schertzer (2010), On the simulation of continuous in scale universal multifractals, Part I: Spatially continuous processes, *Comput. Geosci.*, *36*(11), 1393–1403.
- Lovejoy, S., D. Schertzer, and V. Allaire (2008), The remarkable wide range spatial scaling of TRMM precipitation, *Atmos. Res.*, *90*(1), 10–32.
- Menabde, M., and M. Sivapalan (2000), Modeling of rainfall time series and extremes using bounded random cascades and levy-stable distributions, *Water Resour. Res.*, *36*(11), 3293–3300.
- Molini, A., G. G. Katul, and A. Porporato (2009), Revisiting rainfall clustering and intermittency across different climatic regimes, *Water Resour. Res.*, *45*, W11403, doi:10.1029/2008WR007352.
- Novikov, E., and R. Stiuart (1964), The intermittency of turbulence and the spectrum of energy dissipation, *Izv. Geophys. Ser.*, *3*, 408–413.
- Olsson, J. (1998), Evaluation of a scaling cascade model for temporal rain-fall disaggregation, *Hydrol. Earth Syst. Sci.*, *2*, 19–30.
- Onof, C., R. Chandler, A. Kakou, P. Northrop, H. Wheeler, and V. Isham (2000), Rainfall modelling using Poisson-cluster processes: A review of developments, *Stochastic Environ. Res. Risk Assess.*, *14*(6), 384–411.
- Over, T. M., and V. K. Gupta (1996), A space-time theory of mesoscale rainfall using random cascades, *J. Geophys. Res.*, *101*(D21), 26,319–26,331.

- Pathirana, A., S. Herath, T. Yamada, et al. (2003), Estimating rainfall distributions at high temporal resolutions using a multifractal model, *Hydrol. Earth Syst. Sci. Discuss.*, 7(5), 668–679.
- Pecknold, S., S. Lovejoy, D. Schertzer, C. Hooge, and J. Malouin (1993), The simulation of universal multifractals, in *Cellular Automata: Prospects in Astronomy and Astrophysics*, edited by J. M. Perdang and A. Lejeune, vol. 1, pp. 228–267, World Sci., Singapore.
- Rodríguez, R., M. Casas, and A. Redaño (2013), Multifractal analysis of the rainfall time distribution on the metropolitan area of Barcelona (Spain), *Meteorol. Atmos. Phys.*, 121(3–4), 181–187.
- Rupp, D. E., R. F. Keim, M. Ossiander, M. Brugnach, and J. S. Selker (2009), Time scale and intensity dependency in multiplicative cascades for temporal rainfall disaggregation, *Water Resour. Res.*, 45, W07409, doi:10.1029/2008WR007321.
- Schertzer, D., and S. Lovejoy (1987), Physically based rain and cloud modeling by anisotropic, multiplicative turbulent cascades, *J. Geophys. Res.*, 92(D8), 9693–9714.
- Schertzer, D., and S. Lovejoy (1991), Nonlinear geodynamical variability: Multiple singularities, universality and observables, in *Non-Linear Variability in Geophysics*, pp. 41–82, Springer, Netherlands.
- Schertzer, D., and S. Lovejoy (2011), Multifractals, generalized scale invariance and complexity in geophysics, *Int. J. Bifurcation Chaos*, 21(12), 3417–3456.
- Schertzer, D., S. Lovejoy, and P. Hubert (2002), An introduction to stochastic multifractal fields, in *ISFMA Symposium on Environmental Science and Engineering With Related Mathematical Problems*, pp. 106–179, Higher Educ. Press, Beijing.
- Schleiss, M., A. Berne, and R. Uijlenhoet (2009), Geostatistical simulation of 2d fields of raindrop size distributions at the meso-scale, *Water Resour. Res.*, 45, W07415, doi:10.1029/2008WR007545.
- Schleiss, M., S. Chamoun, and A. Berne (2014), Stochastic simulation of intermittent rainfall using the concept of “dry drift,” *Water Resour. Res.*, 50, 2329–2349, doi:10.1002/2013WR014641.
- Schmitt, F. (2014), Continuous multifractal models with zero values: A continuous  $\beta$ -multifractal model, *J. Stat. Mech. Theory Exp.*, 2014(2), P02008.
- Schmitt, F., S. Vannitsem, and A. Barbosa (1998), Modeling of rainfall time series using two-state renewal processes and multifractals, *J. Geophys. Res.*, 103(D18), 23,181–23,193.
- Serinaldi, F. (2010), Multifractality, imperfect scaling and hydrological properties of rainfall time series simulated by continuous universal multifractal and discrete random cascade models, *Nonlinear Processes Geophys.*, 17, 697–714, doi:10.5194/npg-17-697-2010.
- Tessier, Y., S. Lovejoy, and D. Schertzer (1993), Universal multifractals: Theory and observations for rain and clouds, *J. Appl. Meteorol.*, 32(2), 223–250.
- Tessier, Y., S. Lovejoy, P. Hubert, D. Schertzer, and S. Pecknold (1996), Multifractal analysis and modeling of rainfall and river flows and scaling, causal transfer functions, *J. Geophys. Res.*, 101(D21), 26,427–26,440.
- Tyrallis, H., and D. Koutsoyiannis (2011), Simultaneous estimation of the parameters of the Hurst–Kolmogorov stochastic process, *Stochastic Environ. Res. Risk Assess.*, 25(1), 21–33.
- Vainshtein, S. I., K. R. Sreenivasan, R. T. Pierrehumbert, V. Kashyap, and A. Juneja (1994), Scaling exponents for turbulence and other random processes and their relationships with multifractal structure, *Phys. Rev. E*, 50(3), 1823–1835.
- Veneziano, D., and C. Lepore (2012), The scaling of temporal rainfall, *Water Resour. Res.*, 48, W08516, doi:10.1029/2012WR012105.
- Veneziano, D., P. Furcolo, and V. Iacobellis (2006), Imperfect scaling of time and space–time rainfall, *J. Hydrol.*, 322(1), 105–119.
- Verrier, S., L. De Montera, L. Barthès, and C. Mallet (2010), Multifractal analysis of African monsoon rain fields, taking into account the zero rain-rate problem, *J. Hydrol.*, 389(1), 111–120.
- Verrier, S., C. Mallet, and L. Barthès (2011), Multiscaling properties of rain in the time domain, taking into account rain support biases, *J. Geophys. Res.*, 116, D20119, doi:10.1029/2011JD015719.
- Wang, Z., M. Schleiss, J. Jaffrain, A. Berne, and J. Rieckermann (2012), Using Markov switching models to infer dry and rainy periods from telecommunication microwave link signals, *Atmos. Meas. Tech.*, 5(7), 1847–1859.
- Yaglom, A. (1966), The influence of fluctuations in energy dissipation on the shape of turbulence characteristics in the inertial interval, in *Soviet Physics Doklady*, vol. 11, p. 26–29.