

UNIVERSITÉ JOSEPH FOURIER - GRENOBLE 1  
SCIENCES ET GÉOGRAPHIE

Thèse pour obtenir le titre de  
Docteur de l'Université Joseph Fourier en  
MÉCANIQUE DES MILIEUX GÉOPHYSIQUES ET ENVIRONNEMENT  
GÉOPHYSIQUE, GÉOCHIMIE, GÉOMÉCANIQUE

## Modélisation statistique des champs de pluie sahéliens.

Application à leur désagrégation  
spatiale et temporelle.

Présentée par  
Gilles Guillot

Soutenue le 6 novembre 1998  
devant le jury composé de :

M. Chevallier,  
M. Lebel, *Directeur de Thèse*  
M. Le Dimet, *Président*  
M. Le Treut, *Rapporteur*  
M. Wackernagel, *Rapporteur.*

*Je tiens à remercier Michel Vauclin de m'avoir accueilli au L.T.H.E, Jean-Dominique Creutin de m'avoir fait une place dans l'équipe d'hydro-météorologie et Thierry Lebel de m'avoir permis de faire un thèse au point de rencontre entre hydrologues, météorologues et statisticiens.*

*Je tiens également à remercier les membres du jury : M. Chevallier pour avoir accepté d'évaluer mon travail sur le plan hydrologique, M. Le Dimet d'avoir accepté de présider ce jury, M. Le Treut d'avoir bien voulu être rapporteur de cette thèse dans un délai très court, et enfin Hans Wackernagel, qui trois ans après mon DEA à Fontainebleau, m'a renouvelé sa confiance en acceptant également d'être rapporteur.*

*Ma pensée va aussi à mes profs de Taupe, Catherine Mignard et Marc Becker à Nice et Luc Abergel à Paris. Leur enseignement exemplaire m'a beaucoup apporté.*

*À Elsa, pour ces années passées ensemble.*

*Lyon, le 17 novembre 1998.*

*C'est le vingt-trois juin mille neuf cent soixante-quinze et il va être huit heures du soir. Assis devant son puzzle, Bartlebooth vient de mourir. Sur le drap de la table, quelque part dans le ciel crépusculaire du quatre cent trente-neuvième puzzle, le trou noir de la seule pièce non encore posée dessine la silhouette presque parfaite d'un X. Mais la pièce que le mort tient entre ses doigts a la forme, depuis longtemps prévisible dans son ironie même d'un W.*

(Georges Perec,  
*La Vie mode d'emploi.*)



# Table des matières

<b>Introduction</b>	<b>9</b>
Position du problème . . . . .	9
Domaine d'étude et données disponibles . . . . .	13
Le cadre général et les différentes composantes du modèle . . . . .	15
<b>1 Le modèle gaussien anamorphosé</b>	<b>19</b>
1.1 Choix d'un cadre statistique . . . . .	19
1.1.1 Fonctions, mesures et échelle d'observation . . . . .	19
1.1.2 L'identification d'un événement pluvieux à la réalisation d'une fonction aléatoire . . . . .	20
1.2 La spécification d'une classe de fonctions aléatoires . . . . .	21
1.2.1 Modèles phénoménologiques, modèles comportementaux . . . . .	21
1.2.2 Caractérisation mathématique . . . . .	21
1.2.3 Caractérisation hydrologique . . . . .	22
1.3 Fonctions aléatoires gaussiennes . . . . .	23
1.3.1 Loi normale à une dimension . . . . .	24
1.3.2 Loi normale vectorielle . . . . .	24
1.3.3 Fonction aléatoire normale . . . . .	25
1.4 Fonctions aléatoires gaussiennes anamorphosées . . . . .	27
1.4.1 Anamorphose d'une variable aléatoire . . . . .	27
1.4.2 Anamorphose d'une fonction aléatoire gaussienne . . . . .	28
1.5 Problème de cohérence interne . . . . .	29
1.5.1 Position du problème . . . . .	29
1.5.2 Un théorème de compatibilité . . . . .	29
1.5.3 Fonction de répartition continue . . . . .	30
1.5.4 Fonction de répartition discontinue . . . . .	30
1.6 Mise en œuvre numérique . . . . .	32
1.6.1 Simulation non conditionnelle . . . . .	32
1.6.2 Conditionnement par une moyenne d'espace . . . . .	33
<b>2 Quelques points de méthode sur la fonction de structure</b>	<b>35</b>
2.1 Définitions . . . . .	35
2.2 Estimateurs . . . . .	36

2.2.1	Mono-réalisation . . . . .	36
2.2.2	Multi-réalisation . . . . .	36
2.3	Fonction de structure en présence de dérive . . . . .	37
2.3.1	Estimation à l'aide d'une seule réalisation . . . . .	38
2.3.2	Estimation à l'aide de plusieurs réalisations . . . . .	38
2.4	Sur le variogramme climatologique . . . . .	38
<b>3</b>	<b>Analyse des données à l'échelle de l'événement</b>	<b>41</b>
3.1	Distribution monovariante . . . . .	41
3.1.1	Introduction . . . . .	41
3.1.2	Choix d'une loi et estimation des paramètres . . . . .	42
3.1.3	Discussion . . . . .	43
3.1.4	Conclusion . . . . .	47
3.2	Fonction de covariance . . . . .	48
3.2.1	Introduction . . . . .	48
3.2.2	Calcul de la covariance empirique . . . . .	48
3.2.3	Analyse . . . . .	48
3.2.4	Synthèse . . . . .	52
<b>4</b>	<b>Propriétés du modèle et comparaison aux données</b>	<b>55</b>
4.1	Procédure de mise en œuvre numérique et exemples graphiques . . . . .	55
4.2	Analyse par groupe . . . . .	59
4.2.1	Position du problème . . . . .	59
4.2.2	Effectifs et statistiques simples . . . . .	60
4.2.3	Distribution monovariante . . . . .	61
4.2.4	Fonction de structure . . . . .	65
4.3	Changement d'échelle . . . . .	67
4.3.1	Position du problème . . . . .	67
4.3.2	Propriétés à l'ordre deux . . . . .	67
4.3.3	Intermittence et fonctions de répartitions . . . . .	68
4.4	Relation entre aire des surfaces sup-seuil et moyenne spatiale . . . . .	72
4.4.1	Définition . . . . .	72
4.4.2	Résultats et analyse . . . . .	72
4.5	Conclusion . . . . .	75
<b>5</b>	<b>Cinématique</b>	<b>77</b>
5.1	Introduction . . . . .	77
5.1.1	Difficultés associées à la notion de mouvement . . . . .	77
5.1.2	Mouvement et temps d'arrivée . . . . .	78
5.2	Analyse des données . . . . .	80
5.2.1	Moyenne et variance . . . . .	80
5.2.2	Covariance . . . . .	85
5.3	Modèle . . . . .	87
5.4	Comparaison aux données . . . . .	89

<b>6</b>	<b>Hyétoqrammes</b>	<b>93</b>
6.1	Introduction, position du problème . . . . .	93
6.2	Étude de la forme des hyétoqrammes . . . . .	96
6.2.1	Représentation simplifiée . . . . .	96
6.2.2	Erreur commise par le modèle de hyétoqramme . . . . .	97
6.2.3	Relation entre la forme du hyétoqramme et le cumul-événement . . . . .	100
6.3	Schéma de désagrégation . . . . .	105
6.4	Éléments de validation . . . . .	106
6.4.1	Distribution des intensités maximales . . . . .	106
6.4.2	Structure spatiale des champs d'intensité maximale . . . . .	107
<b>7</b>	<b>Propriétés spatiales des champs à cinq minutes</b>	<b>109</b>
7.1	Problèmes de méthode . . . . .	109
7.2	Aire moyenne des surfaces sup-seuil . . . . .	110
7.3	Distribution des aire des surfaces sup-seuil . . . . .	111
	<b>Conclusion</b>	<b>117</b>
	Résumé des principaux résultats . . . . .	117
	Limites du modèle . . . . .	118
	Perspectives . . . . .	119
	<b>Index des noms cités</b>	<b>121</b>
	<b>Bibliographie</b>	<b>124</b>



# Introduction

## Position du problème

### La pluie au Sahel : deux niveaux d'enjeux

Le Sahel, vaste région de terre africaine comprise environ entre 10-12 et 17-18 degrés de latitude nord, souffre d'un déficit pluviométrique marqué depuis plus de vingt ans par rapport aux années antérieures. Cette sécheresse a fait l'objet d'une description et d'une analyse volumineuse dans la littérature, on pourra se rapporter par exemple aux références suivantes : [123, 37, 55, 96, 97, 98, 99, 124, 57, 58, 26, 59]. Cette sécheresse dont les manifestations les plus graves et les plus médiatisées ont été celles des années 1972-73 et 1984, met en cause les fragiles équilibres établis entre l'homme et son milieu naturel; toutefois, son ampleur et ses conséquences à plus long terme, en particulier la manière dont elle affecte les ressources en eau de cette région, sont encore mal connues.

Par ailleurs, les précipitations et, plus généralement, les échanges de masse et d'énergie entre le sol et l'atmosphère dans la bande inter-tropicale contribuent de manière significative au bilan synoptique. Elle jouent donc un rôle important dans la formation, la régulation, et les éventuelles modifications du climat de la Terre, y compris sous nos latitudes.

L'étude du cycle hydrologique au Sahel se trouve donc motivée en premier lieu par la nécessité de mieux comprendre les mécanismes en cause dans une sécheresse *locale* préjudiciable à l'activité humaine dans cette région, en second lieu par le besoin d'améliorer nos connaissances sur les fluctuations climatiques au niveau *global*.

### Quelques éléments sur la climatologie sahélienne

La région sahélienne est affectée par une saison des pluies de mai à octobre. Le début de la saison des pluies et son efficacité pluviométrique sont liés au positionnement de la zone de convergence intertropicale (Z.C.I.T). Cette dernière est une zone de forte nébulosité positionnée au voisinage de l'équateur et définie comme l'interface entre le flux de mousson et les masses d'air chaud et sec du Sahara [33].

La variabilité pluviométrique sur la région résulte du positionnement de la Z.C.I.T, de l'intensité de l'African Easterly Jet (A.E.J), qui est un courant des masses d'air créé par le contraste thermique entre le Sahara et l'Atlantique Sud, et de celle du Tropical Easterly Jet (T.E.J) dû au contraste thermique entre l'Océan Indien et le continent asiatique [34]. La température des océans constitue également un facteur majeur, encore mal connu. Il se manifeste notamment par son influence sur le flux d'humidité en basse couche et par son rôle sur les jets.

Ces différents phénomènes donnent naissance aux Systèmes Convectifs de Mésos-échelle (S.C.M), qui sont des amas nuageux prenant la forme de grandes tours convectives pouvant parfois atteindre plusieurs kilomètres de haut et très étendus horizontalement (plusieurs dizaines de milliers de  $\text{km}^2$ ) et se déplaçant d'Est en Ouest à une vitesse d'environ  $50 \text{ km.h}^{-1}$ .

Ces S.C.M sont le siège d'une activité convective alimentée par l'humidité du flux de mousson. Cette activité peut s'organiser, selon des mécanismes encore imparfaitement connus, en un front (désigné par la suite par front convectif) dont la largeur est d'environ 20 km de large [75]. Les endroits du front où l'activité convective est maximale, ou cellules convectives, constitueraient des structures cohérentes de petite taille [113]. La forme la plus simple que peuvent prendre les S.C.M (du point de vue d'une description géométrique extérieure très sommaire), est une organisation en un front quasiment rectiligne se déplaçant à vitesse constante. De tels systèmes sont appelés lignes de grains [110]. Un exemple de ligne de grains observé par un radar est présenté figure 0.3.

## Systèmes hydrologiques et précipitations

La réponse hydrologique au signal pluviométrique présente des aspects contradictoires selon les échelles auxquelles on s'intéresse. La période sèche en cours s'est traduite par une diminution sensible du débit des grands fleuves dans la région : le module moyen annuel du Niger à Niamey est passé de  $1060 \text{ m}^3\text{s}^{-1}$  sur la période 1929-1968 à  $700 \text{ m}^3\text{s}^{-1}$  sur la période 1969-1994, soit une diminution de 34%. Et la moyenne des minimums journaliers en étiage est passée de  $64 \text{ m}^3\text{s}^{-1}$  à  $11 \text{ m}^3\text{s}^{-1}$  pour les mêmes périodes [13]. Parallèlement, le ruissellement sur les petits bassins aurait plutôt eu tendance à augmenter [5, 105]. Plus récemment, on constate une remontée de l'aquifère supérieure du continental terminal à partir de la fin des années 70, alors que la période sèche se prolonge [76].

Concernant la région de Niamey l'hydrologie fonctionne suivant un schéma analysé par Desconnets et al. [31] : le ruissellement de type hortonien provoqué par les pluies intenses est concentré dans des mares, le réseau de drainage de surface étant trop dégradé pour assurer des écoulements exoréiques. Ces mares servent à la fois de ressources en eau de surface durant la période où elles sont pleines et de système de transit pour la recharge des aquifères.

Sans réduire la grande diversité des systèmes hydrologiques sahéliens aux types rencontrés dans la région de Niamey, et malgré la difficulté à définir objectivement les échelles élémentaires auxquelles il convient de représenter les précipitations

en entrée des modèles hydrologiques [38, 102], il est naturel de penser qu'une description correcte du fonctionnement de ces systèmes hydrologiques requiert une prise en compte de la variabilité des précipitations à une échelle très fine : de l'ordre du  $\text{km}^2$  en espace, et de quelques minutes en temps.

## Étude du cycle hydrologique au Sahel et problèmes d'échelle

Les phénomènes en jeu dans la formation des précipitations au Sahel se développent à de grandes échelles spatiales et temporelles (en particulier du fait de la grande inertie thermique des océans). Ces échelles sont représentées à la figure 0.1. Pour des raisons qui sont donc intrinsèques aux phénomènes étudiés, mais aussi pour des raisons de coût de calcul numérique, les modèles météorologiques représentent l'atmosphère sur des grilles dont les mailles sont très grandes au regard des besoins de l'hydrologie. Les Modèles de Circulation Générale Atmosphérique (M.C.G.A) qui réalisent un maillage de toute la planète ont des mailles dont le côté est de 200 à 500 km (dans la bande inter-tropicale).

Concernant l'observation des systèmes précipitants au Sahel, les seules mesures directes disponibles dans un contexte opérationnel sont issues des réseaux nationaux de pluviomètres. Leur densité est inférieure à un poste pour  $6000 \text{ km}^2$  et les données sont fournies au pas de temps journalier. Ces données sont utiles pour réaliser des suivis de long terme de la région. Mais la maille des réseaux étant supérieure aux échelles de la cellule convective (et même dans certaines circonstances à l'extension spatiale des systèmes précipitants), une part très importante de l'information spatiale, utile pour les études hydro-dynamiques, est absente de ces données.

Les estimations dérivées des images satellitales fournissent une alternative ou un complément possible aux données sol. Il s'agit de mesures indirectes dont la précision croît en sens inverse de la résolution souhaitée, et on considère que leur précision est acceptable pourvu qu'on se limite à des pas de temps de l'ordre de 12 heures et des domaines d'au moins  $10000 \text{ km}^2$  [22, 10, 44].

Il apparaît donc que l'information relative aux précipitations telle qu'elle est donnée par les satellites ou les M.C.G.A est fournie à une échelle beaucoup trop grossière pour la modélisation hydrologique.

## Objectifs de l'étude

On se propose dans ce travail de développer une méthode susceptible d'utiliser l'information de grande échelle fournie par un satellite ou un M.C.G.A, afin de produire des champs de pluie aux échelles requises par la modélisation hydrologique.

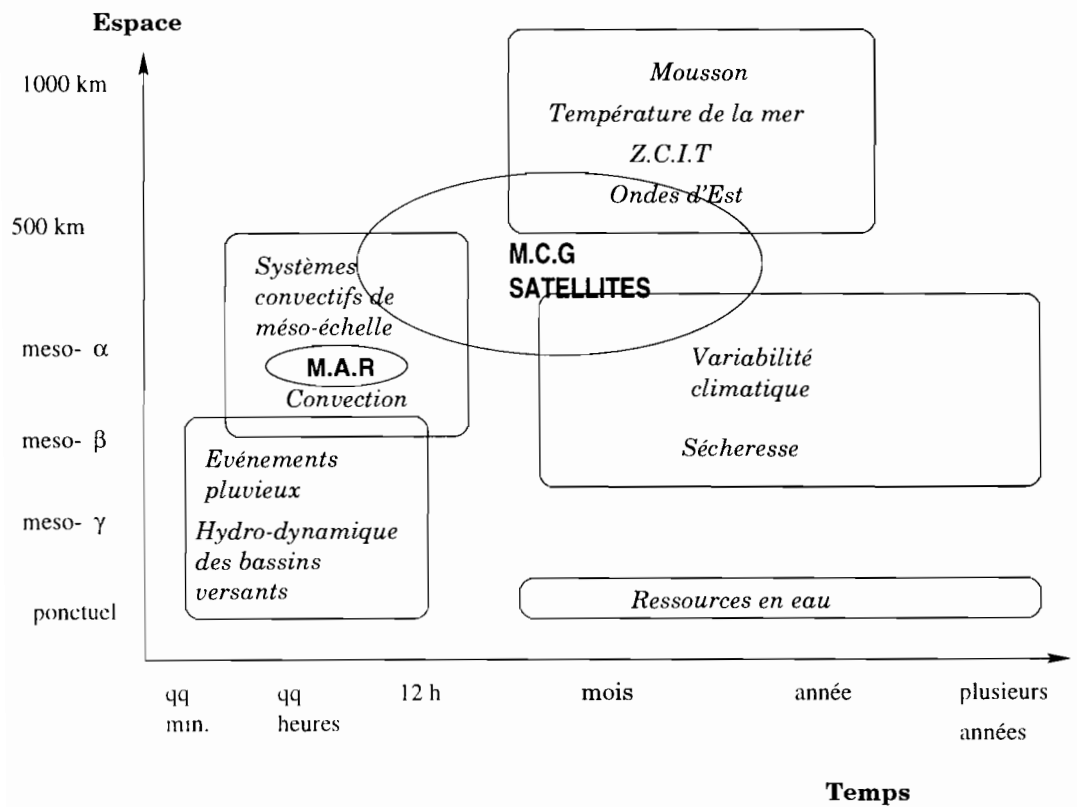


FIG. 0.1: Aperçu sur les échelles caractéristiques à prendre en compte, dans l'étude du cycle hydrologique en Afrique tropicale : il y a discontinuité entre les échelles de la phase continentale et celles de la phase atmosphérique.

De manière générale, les procédures permettant d'utiliser l'information de grande échelle pour produire des champs météorologiques à une plus haute résolution sont désignées dans la littérature par *downscaling models* ou modèles de désagrégation.

On peut trouver l'origine de ces modèles dans le problème qui se pose en hydrologie lorsqu'on dispose d'un volume total écoulé à l'exutoire pour une période donnée et que l'on souhaite reconstituer la série des débits à un pas de temps inférieur à la longueur de cette période [133]. Le problème est donc de retrouver la série  $(Q_1, \dots, Q_n)$  connaissant la somme  $\sum_{i=1}^n Q_i$ . Un problème formellement analogue s'est ensuite posé en hydro-météorologie pour la désagrégation temporelle d'une lame d'eau en un point.

Des modèles très variés ont été proposés pour répondre à ces problèmes [138]. Ils diffèrent les uns des autres par la nature de l'information de grande échelle (spatiale ou temporelle) disponible, et la représentation des phénomènes mise en jeu (statistique, mécanique, thermodynamique, etc).

Les divers modèles décrits dans la littérature abordent principalement le pro-

blème de la désagrégation spatiale, le problème de la désagrégation spatio-temporelle étant abordé via une hypothèse d'invariance temporelle [52]. Certaines de ces procédures ont en outre l'inconvénient de n'être que des algorithmes, sans explicitation des hypothèses utilisées. Un travail de Kumar et Foufoula-Georgiou [64] en est un exemple (très intéressant) dans le formalisme de la morphologie mathématique.

Nous allons présenter un modèle (et expliciter les hypothèses formulées sur les champs de pluie), qui permettra d'utiliser l'information de grande échelle qui est disponible de manière opérationnelle pour la zone sahélienne.

Compte tenu du contexte évoqué à la section précédente, nous considérerons que l'on peut disposer de valeurs de précipitations correspondant à une moyenne sur un domaine d'espace d'environ  $10000 \text{ km}^2$  et sur la durée de l'événement. Cette valeur peut être une valeur mesurée par un satellite, ou une valeur simulée par un M.C.G.A. Dans cette dernière configuration on peut également considérer que l'on dispose d'une information sur le mouvement du système précipitant (un vent moyen par exemple).

Évidemment, le passage de la valeur agrégée aux champs à haute résolution ne peut se faire qu'en injectant de l'information, ceci sera fait par l'intermédiaire d'un modèle statistique de champs de pluie. L'objet de la prochaine section est de décrire les données ayant permis de le spécifier.

## Domaine d'étude et données disponibles

Les données utilisées seront celles de l'expérience EPSAT-Niger menée entre 1990 et 1995. Cette opération prenait place dans une démarche générale de la communauté scientifique internationale visant à améliorer nos connaissances sur les échanges de masse et d'énergie entre le sol et l'atmosphère à la méso-échelle au Sahel, et organisée au sein de l'opération HAPEX-Sahel (Hydrological Atmospheric Pilot EXperiment). Plus précisément, un des objectifs importants d'EPSAT-Niger était d'améliorer les algorithmes d'estimation des précipitations par satellite [74].

Cette opération se basait sur un réseau de pluviographes disposés sur un domaine d'environ  $16000 \text{ km}^2$  dans la région de Niamey au Niger. Le réseau de mesure est constitué d'un premier groupe de trente pluviographes à peu près uniformément disposés sur le domaine et ayant fonctionné entre 1990 et 1995 (réseau de base, voir figure 0.2). Ce réseau était complété entre 1990 et 1993 par environ 70 autres appareils, déplacés d'une année sur l'autre afin d'obtenir différentes configurations. En particulier, on dispose pour l'année 1993 d'un réseau dont la densité atteint localement une station par  $\text{km}^2$ . Les pluviographes sont à augets basculeurs de 0,5 mm et enregistrent des temps de basculement avec une précision de 1 seconde. Pour obtenir des éléments plus précis sur ce dispositif expérimental, on se rapportera aux rapports de campagne annuels [70, 128, 129, 126, 127].

La zone d'étude ne présente pas de singularités géographiques, elle constitue à

ce titre un lieu privilégié pour l'observation des systèmes précipitants et de leurs interactions avec un milieu naturel relativement homogène.

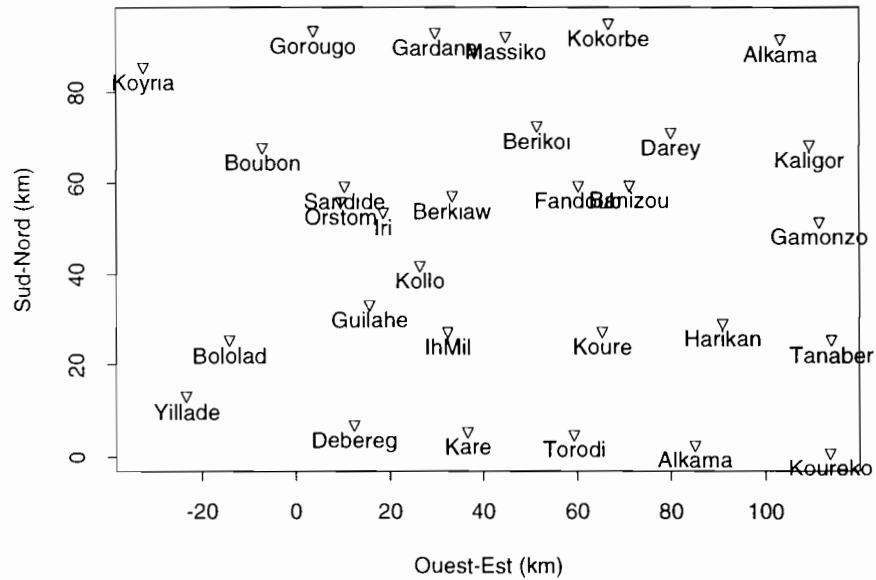


FIG. 0.2: Réseau de base de l'opération EPSAT-Niger.

La saison pluvieuse est rythmée par l'occurrence d'événements associés aux S.C.M bien délimités dans le temps et dans l'espace. Si l'on exclut les événements ayant affecté moins de 30% du réseau<sup>1</sup>, on recense 258 événements pluvieux entre 1990 et 1995. Ces événements principalement d'origine convective seront désignés par la suite par systèmes convectifs de méso-échelle (SCM).

Ces systèmes ont fait l'objet de différents travaux, qui ont permis de se faire une première idée sur quelques grandes caractéristiques des S.C.M d'Afrique de l'Ouest. On sait ainsi que leur forme développée, les Complexes Convectifs de Méso-échelle (M.C.C), est plus fréquente sur le Sahel que sur la région soudanienne [67], que l'orographie peut servir de déclencheur à leur formation [78] et

1. Ces derniers représentant environ 2% du cumul annuel

que les facteurs environnementaux de la circulation générale ont un rôle important dans leur maturation [134]. Ces M.C.C sont repérables dans les sorties de M.C.G.A [104], mais avec des caractéristiques (durée de vie, vitesse de déplacement et pluviosité) différentes de celles qui sont établies à partir des satellites et des données sol. A partir d'une combinaison de différents types de données, LeBarbé et Lebel [69] et Laurent et al [67] ont montré que les fluctuations pluviométriques interannuelles étaient liées avant tout aux fluctuations d'occurrence des SCM, plus qu'à la variation de leur efficacité pluviométrique. On a mis en évidence par ailleurs la grande variabilité intrasaisonnière des pluies sur le Sahel tant à l'échelle synoptique qu'aux échelles moyennes (des gradients de l'ordre de 200 à 300 mm/10 km ont été observés sur des cumuls saisonniers variant entre 300 et 600 mm dans la bande comprise entre 13° et 14° N). Cette variabilité est bien moins connue et documentée que les variabilités interannuelle et décennale. Elle a pourtant un fort impact immédiat sur le cycle hydrologique. Elle se manifeste aussi bien dans le temps (périodes sèches en cours de saison des pluies) que dans l'espace (fortes irrégularités d'un lieu à l'autre).

Les événements pluvieux peuvent se différencier selon différents critères tels que le degré d'organisation spatiale, leur maturité au moment du passage au-dessus du réseau de mesure ou le développement de l'activité convective. Et il est légitime de penser que ces systèmes seront caractérisés par des statistiques spatiales et temporelles différentes.

Dans une perspective d'amélioration des méthodes géostatistiques d'interpolation des champs de pluie, Amani [7, 9] a testé certains critères de classification basés sur des données pluviographiques. Il a été montré que la distinction des événements selon la fraction du réseau non affectée par la pluie au cours de l'événement (intermittence) constituait un bon critère. C'est pourquoi nous garderons à l'esprit l'existence de trois groupes d'événements pluvieux :

**groupe 1 :** intermittence  $\leq 5\%$

**groupe 2 :**  $5\% < \text{intermittence} \leq 25\%$

**groupe 3 :**  $25\% < \text{intermittence} \leq 70\%$

Laurent a montré plus tard [67, 68] que ce critère coïncidait assez bien avec des critères atmosphériques de grande échelle. En conséquence, nous assimilerons les événements du groupe 1 aux *ligne de grain*, la réunion des groupes 1 et 2 aux *complexes convectifs de méso-échelle* et la réunion des trois groupes aux *systèmes convectifs de méso-échelle*.

## Le cadre général et les différentes composantes du modèle

Les simulations que l'on souhaite réaliser étant orientées vers la modélisation hydrologique, leur intérêt sera de rendre compte des propriétés des précipitations

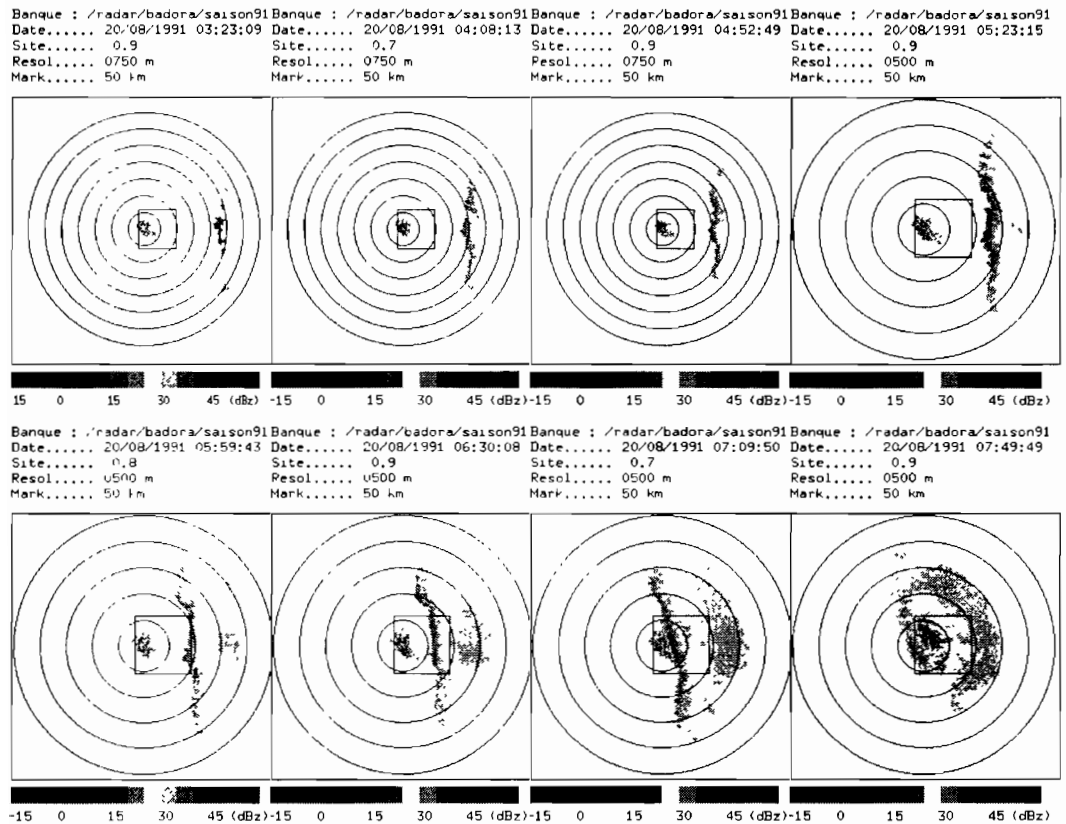


FIG. 0.3: Ligne de grain observée le 20 août 1991 par le radar de N'amey [70, 75]

telles qu'elles sont mesurées au sol. On s'intéressera donc plus aux propriétés bi-dimensionnelles des champs de pluie qu'aux phénomènes atmosphériques qui leur donnent naissance. On cherchera alors à rendre compte à la fois des aspects structurés du phénomène (organisation à des échelles préférentielles, récurrence de certains aspects de la cinématique) et de ses aspects non structurés (ou nous apparaissant tels, compte tenu de nos outils d'analyse et de nos connaissances de ces systèmes). Un cadre statistique est donc bien adapté à cette situation.

Ayant fait le choix d'analyser et de modéliser les précipitations dans un cadre statistique, le problème de la désagrégation s'interprète simplement comme un problème de simulation conditionnelle; toutefois un événement pluvieux apparaît comme un objet statistique relativement complexe. En particulier, comme on peut le voir sur la figure 0.3, les champs de pluie manifestent à la fois une grande hétérogénéité spatiale et une interdépendance temporelle marquée.

Il est donc intéressant de chercher à en simplifier l'étude en essayant, d'une

part de ramener l'étude des champs de pluie à certains de leurs aspects présentant des propriétés d'invariance spatiale, ou temporelle, d'autre part, de découpler les aspects spatiaux des aspects temporels.

À cette fin, ce travail sera structuré, tel que le présente le schéma 0.4, de la manière suivante :

1. Nous étudierons tout d'abord les champs de pluie au pas de temps de l'événement. La dimension temporelle disparaît.
  - On rappelle alors la définition du modèle gaussien anamorphosé.
  - On examine ensuite certains points de méthode liés à son utilisation dans le cadre des champs de pluie.
2. La deuxième partie est consacrée à la mise en œuvre de ce modèle c'est à dire à l'analyse des données pour l'estimation des paramètres du modèle et à une comparaison, essentiellement expérimentale (i.e numérique), du modèle et des données.
3. La troisième partie aborde les propriétés des champs de précipitations aux petits pas de temps.
  - On étudie tout d'abord certaines propriétés cinématiques des systèmes précipitants, afin de simuler les temps d'arrivée de la pluie en chaque point du domaine d'étude.
  - Les distributions temporelles des intensités sont ensuite étudiées à travers les propriétés génériques du hyétogramme ponctuel.
  - Quelques propriétés résultant du couplage des différentes composantes du modèle sont enfin étudiées dans le dernier chapitre.

Signalons enfin que les différentes pistes ayant conduit au modèle dans sa version ultime ne sont pas toutes décrites ici pour ne pas alourdir la lecture. On pourra se rapporter aux publications correspondantes référencées en bibliographie [45, 118, 117, 25, 72, 47, 49, 46, 50, 48].

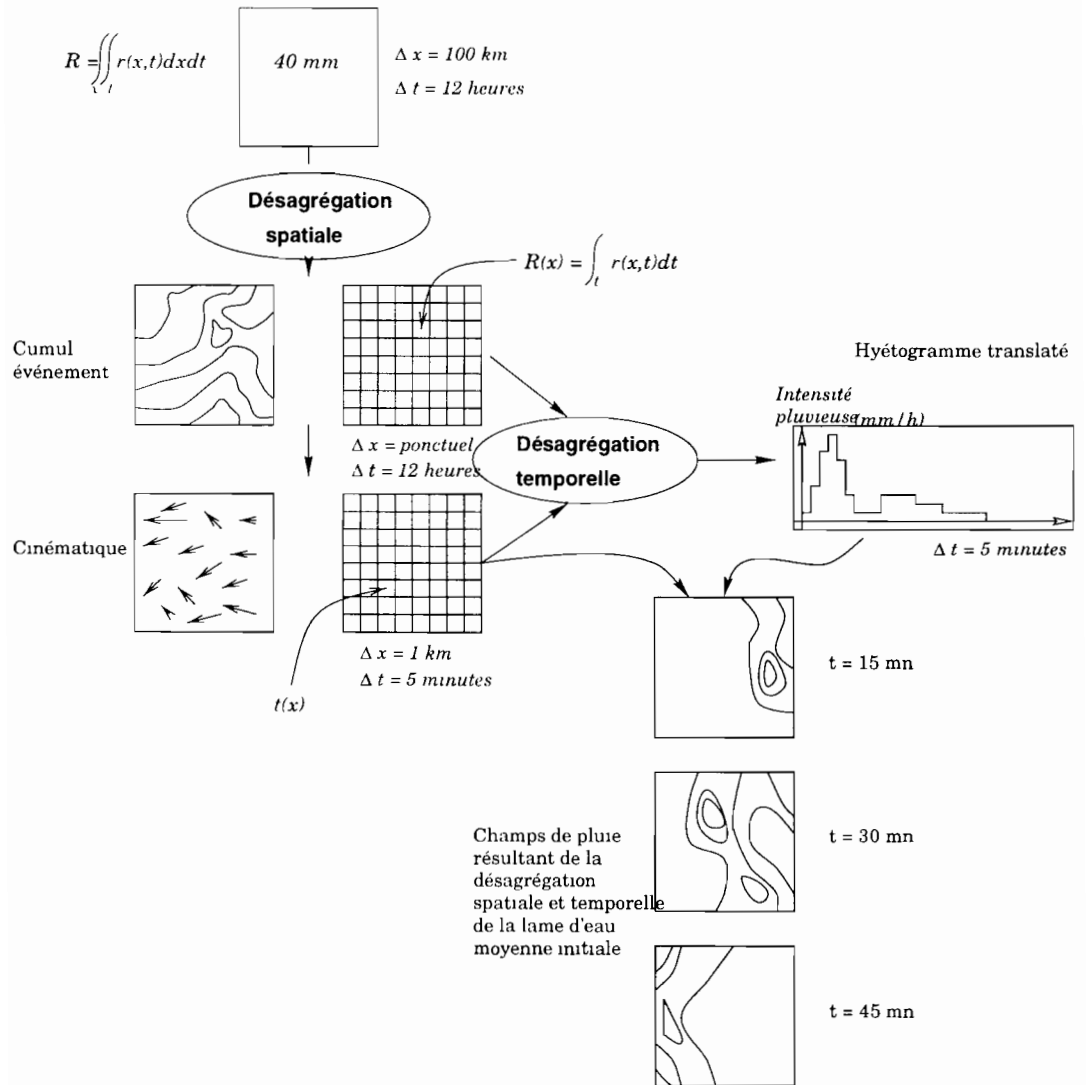


FIG. 0.4: Représentation schématique des différentes composantes du modèle

# Chapitre 1

## Le modèle gaussien anamorphosé

### Présentation

Ce chapitre aborde la modélisation statistique des champs de pluie au pas de temps de l'événement. Il s'agit d'un modèle spatial à temps fixe. On présente le cadre des fonctions aléatoires gaussiennes anamorphosées qui établit un lien entre la loi spatiale gaussienne et une loi de distribution mono-variable donnée. Ce modèle apparaît comme un choix naturel lorsqu'on veut spécifier une covariance (qui contient des informations sur la structure spatiale) et une distribution monovariée (contenant une information sur les fluctuations de la lame d'eau en chaque point). Cette simplicité cache néanmoins des difficultés théoriques au niveau de l'estimation des paramètres. Une solution pragmatique (adaptée aux paramètres que l'on peut s'attendre à trouver dans le cadre de l'estimation des champs de pluie sahéliens à ces échelles) est proposée.

### 1.1 Choix d'un cadre statistique

#### 1.1.1 Fonctions, mesures et échelle d'observation

Toute mesure physique s'applique à une certaine mesure d'espace élémentaire (surface, volume ...), en particulier on ne mesure jamais une lame d'eau en un point. C'est pourquoi il peut sembler judicieux de représenter la pluie par une mesure. C'est à dire une application définie sur une famille de parties de l'espace, satisfaisant un certain nombre d'axiomes [114]. La théorie de la mesure (y compris l'analyse des mesures aléatoires) donne alors un cadre naturel pour décrire un phénomène à différentes échelles, sans avoir à faire référence à des valeurs ponctuelles. Ce formalisme est d'un usage très récent en hydro-météorologie, mais

bien que séduisant, il n'offre actuellement que des modèles assez sommaires [52, 103, 23, 24].

Nous représenterons donc plus classiquement la pluie par une fonction numérique (c'est à dire une application définie en des points de l'espace) de deux variables réelles (les coordonnées d'espace), notée  $r(x)$ , où  $x = (\xi_1, \xi_2)$ . Et on supposera que la variable  $r(x)$  représente la lame d'eau mesurée par un pluviomètre placé au point  $x$ . Ceci nous permettra de disposer d'un cadre mathématique plus ancien et donc mieux connu pour analyser les champs de pluie.

Ce choix revient à privilégier une échelle particulière, celle du pluviographe, qui n'a pas un grand intérêt du point de vue de l'hydrologie de surface. Nous tenterons de le justifier ultérieurement en examinant comment à partir d'un réseau de mesures ponctuelles, on parvient à capturer les propriétés pour différents supports d'intégration spatiale.

Ce choix est évidemment largement influencé par la nature de nos données.

### 1.1.2 L'identification d'un événement pluvieux à la réalisation d'une fonction aléatoire

Nous considérerons que chaque champ de pluie observé est la réalisation d'une fonction aléatoire sur un certain domaine du plan. A partir de là, deux choix sont possibles. Le premier consiste à analyser chaque champ de pluie séparément, en l'interprétant comme une réalisation unique d'une fonction aléatoire propre à l'événement considéré. Chaque événement est donc considéré individuellement. Une autre possibilité est de considérer les différents champs observés comme des réalisations *indépendantes* d'une *même* fonction aléatoire.

Ayant observé un grand nombre d'événements pluvieux et souhaitant disposer d'une représentation la plus synthétique possible, nous avons choisi la deuxième solution.

Nous verrons au cours du développement les imperfections de ce modèle, mais aussi comment il permet de rendre compte de certaines propriétés des champs de pluie. Nous examinerons également les relations existant entre les paramètres que l'on peut estimer sur une réalisation et ceux relatifs à plusieurs réalisations, faisant ainsi un lien entre les deux approches.

On notera  $R(x)$  la valeur de notre fonction aléatoire au point  $x$ , (c'est une variable aléatoire). Nous ferons référence à un événement particulier, repéré par un indice  $k \in [1, \dots, K]$  par  $r_k$ . De manière générale, nous nous efforcerons de faire une distinction entre les données, désignées par une lettre minuscule et les objets aléatoires qui les modélisent, désignées par une lettre majuscule.

## 1.2 La spécification d'une classe de fonctions aléatoires

### 1.2.1 Modèles phénoménologiques, modèles comportementaux

On peut distinguer deux démarches de modélisation stochastique. Une première démarche, que l'on pourrait qualifier de phénoménologique [136], consiste à inclure dans le modèle une information *a priori* sur le phénomène étudié. La nature de cette information peut être très diverse : mécanismes permettant d'écrire des équations de conservation, ordres de grandeur concernant certaines échelles caractéristiques, analogie avec un autre phénomène physique dont on connaît les équations. On déduit alors de cette connaissance une forme de modèle admissible. A l'opposé, se trouve une démarche qu'on pourrait qualifier de comportementale et qui consiste à reproduire un comportement observé, à partir d'une classe assez générale de fonctions aléatoires, sans connaissance sur le processus ayant engendré les données. C'est dans cette démarche que s'inscrit cette partie de notre travail : on a dépouillé le phénomène de la « physique élémentaire » (le mouvement des systèmes précipitants) et on essaye de rendre compte de ce qui subsiste de structuré sans connaissance sur la genèse des champs de pluie.

### 1.2.2 Caractérisation mathématique

Si  $X$  est une variable aléatoire, la connaissance de sa fonction de répartition permet de calculer les probabilités de n'importe quel événement relatif à  $X$ . Autrement dit, toute l'information relative à  $X$  est contenue dans une fonction de  $\mathbb{R}$  dans  $\mathbb{R}$ . S'agissant d'un vecteur aléatoire  $X_1, \dots, X_n$ , la situation est analogue puisque la fonction de répartition multivariable

$$F : x_1, \dots, x_n \longrightarrow P[X_1 \leq x_1, \dots, X_n \leq x_n]$$

permet de calculer la probabilité de n'importe quel événement attaché au vecteur  $(X_1, \dots, X_n)$ .

Dans le cas d'une fonction aléatoire  $X(t)$  définie sur  $\mathbb{R}^n$ , la connaissance des probabilités attachées aux événements faisant intervenir un nombre fini de points requiert la donnée de la loi spatiale, à savoir, la donnée de toutes les fonctions du type :

$$F_{t_1, \dots, t_n} : \begin{array}{ll} \mathbb{R}^n & \longrightarrow \mathbb{R} \\ (x_1, \dots, x_n) & \longrightarrow P[X(t_1) \leq x_1, \dots, X(t_n) \leq x_n] \end{array}$$

ceci pour tout n-uplet  $(t_1, \dots, t_n)$  d'éléments de  $\mathbb{R}$ , pour tout entier  $n$ .

Mais la loi spatiale n'est en général pas suffisante pour calculer toutes les propriétés relatives à  $X(t)$ , en particulier pour connaître la régularité des réalisations [16, 42]. On conçoit alors que le contrôle de tous les aspects d'un modèle de fonction aléatoire est impossible.

Il conviendra donc de distinguer deux types de propriétés : un petit nombre de propriétés qui nous paraissent importantes pour l'hydrologie, un grand nombre

de propriétés (assez fortes) qu'il est nécessaire de spécifier pour être en mesure de réaliser des simulations.

### 1.2.3 Caractérisation hydrologique

Dans la modélisation pour la synthèse vocale, on cherche à représenter la parole par une fonction. On considère qu'une fonction représente convenablement un signal de parole si le signal simulé est reconnaissable par un être humain. Des travaux récents [77] ont montré, que la donnée de la valeur d'un signal en quelques points, et de sa régularité hölderienne<sup>1</sup> suffisaient pour que ce signal soit reconnu par un être humain. Le fait remarquable est que les deux propriétés utilisées pour représenter une fonction (régularité et valeur en quelques points) sont mathématiquement très faibles<sup>2</sup>, mais l'information contenue dans ces deux propriétés est suffisante *pour une application donnée*: la synthèse d'un signal et sa reconnaissance par un être humain.

Il s'agit là d'une circonstance assez remarquable puisque les conditions expérimentales qui prévalent permettent de tester immédiatement l'adéquation du modèle à une application donnée. Or, selon la formule de G. Matheron [89]: *«C'est toujours à nos risques et périls que nous choisissons un modèle. Ce modèle, moyennant des calculs plus ou moins rapides mais faisables, nous fournit une solution. Maintenant que vaut cette solution? Exactement autant que notre modèle. Si ce modèle est grossièrement faux, la solution correspondante risque d'être illusoire. L'expérience et la pratique en décident: les bons modèles sont justement ceux qui fournissent des solutions que la pratique ultérieure confirme»*.

Malheureusement, les conditions expérimentales qui prévalent en hydrologie font qu'il est très difficile de déterminer empiriquement quelles sont les propriétés des champs de pluie qu'il convient de prendre en compte. Car d'une part, les conditions expérimentales ne sont pas reproductibles, on en est (en général) réduit à observer, mais aussi parce que les précipitations ne sont pas étudiées pour elles-mêmes mais en liaison avec un système hydrologique particulier, qui est lui-même très imparfaitement connu.

Certains auteurs [137, 102, 38, 120] ont examiné à quelles échelles il convient de représenter la pluie en entrée des modèles hydrologiques. Leurs conclusions sont très partielles et concernent seulement les modèles et non pas les systèmes qu'ils représentent.

On ne dispose donc pas d'éléments pour définir quelles sont les propriétés pertinentes pour l'hydrologie. Nous nous laisserons guider par les considérations sui-

---

1. La régularité hölderienne est une régularité locale utilisée notamment pour l'analyse des fonctions fractales [51, 94].

2. Car l'espace des fonctions admettant une régularité hölderienne donnée est dense dans l'espace des fonctions continues.

vantes :

- Un certain usage a consacré l'étude des fonctions de covariance des champs de pluie, en particulier parce que leur étude est nécessaire pour mettre en oeuvre des méthodes d'estimation comme le krigeage. Par ailleurs, on sait que dans le cas gaussien (voir section 1.3.3.2), toute l'information relative à la loi spatiale de la fonction aléatoire est contenue dans la covariance. Nous étudierons donc les corrélations spatiales à travers la fonction de covariance, ce qui permettra également de rendre compte de la variance des champs de pluie à différentes échelles.
- Nous verrons que la distribution gaussienne est inadaptée aux distributions ponctuelles des champs de pluie, en particulier à cause de la présence de vastes zones non touchées par la pluie à l'intérieur des zones affectées par la pluie (intermittence). Nous essayerons de prendre en compte ce type de propriétés par une transformation adéquate de la loi gaussienne ponctuelle.
- Afin de spécifier complètement notre modèle, le choix d'une covariance et d'une distribution monovariante n'est pas suffisant. Il est nécessaire de faire des hypothèses sur la loi spatiale. Ceci sera fait dans le cadre du modèle gaussien anamorphosé. Ce modèle permet de restituer dans un cadre cohérent les propriétés de corrélation spatiale et de la distribution marginale ; il est en outre relativement simple à mettre en oeuvre. Le rappel de sa définition est l'objet des sections suivantes.

### 1.3 Fonctions aléatoires gaussiennes

L'intérêt de la loi gaussienne en statistique est multiple. Tout d'abord c'est une loi simple, dépendant de deux paramètres concrets (ses deux premiers moments). Elle présente ensuite la particularité d'apparaître naturellement dans des contextes très divers à cause du théorème central limite, qui affirme qu'une somme de variables aléatoires indépendantes de même loi, pas trop dispersées est approximativement gaussienne. Enfin c'est une famille qui est stable pour les opérations linéaires (ce qui peut paraître faible mais qui est en fait une circonstance très rare en Probabilités). Ceci se traduit par des propriétés algébriques très utiles pour le calcul et les simulations.

Les vecteurs gaussiens généralisent la notion de loi gaussienne mono-variable, au sens où les propriétés énoncées ci-dessus sont conservées. Les fonctions aléatoires gaussiennes généralisent la notion de vecteur aléatoire gaussien à des familles de v.a indexées par un ensemble infini (éventuellement non dénombrable). La présente section est consacrée à un rappel rapide de la définition de ces objets, avec une attention particulière pour les propriétés en rapport avec les problèmes d'estimation des paramètres.<sup>3</sup>

---

3. Concernant la bibliographie, il serait souhaitable de pouvoir se rapporter à une unique

### 1.3.1 Loi normale à une dimension

#### 1.3.1.1 Définition, propriétés fondamentales

**Définition 1** Une variable aléatoire  $Y$  suit une loi normale centrée réduite si sa densité est de la forme

$$g(y) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}y^2\right)$$

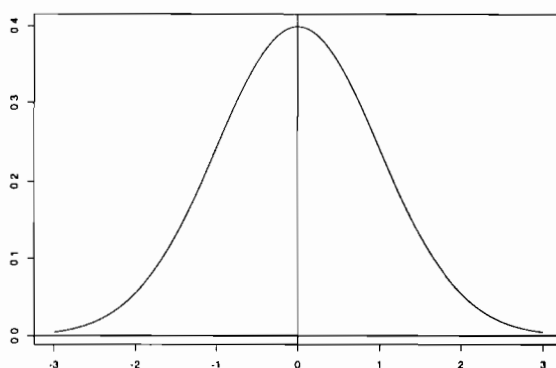


FIG. 1.1: Densité gaussienne centrée réduite

Une variable normale  $X$  d'espérance  $m$  et de variance  $\sigma^2$  est de la forme  $X = m + \sigma Y$  où  $Y$  est gaussienne centrée réduite et admet alors pour densité

$$g_{m,\sigma}(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2\right)$$

### 1.3.2 Loi normale vectorielle

**Définition 2** Un vecteur aléatoire  $Y$  est dit gaussien ou normal si toute combinaison linéaire de ses composantes est gaussienne.

référence. Malheureusement, on trouve d'une part de très nombreux ouvrages de probabilité élémentaire, parfois très bien faits, mais qui se limitent à une présentation sommaire des processus stochastiques. D'autre part il existe des ouvrages entièrement consacrés aux processus. Certains abordent le sujet dans un cadre (excessivement) général et nécessairement assez abstrait [41], tandis que d'autres traitent d'un aspect particulier du sujet de manière très détaillée (représentation spectrale, géométrie intégrale, inférence ...). La conséquence de cette situation est qu'il est nécessaire de se référer à des ouvrages généraux pour les définitions et propriétés élémentaires et à des publications plus spécialisées pour les nombreuses petites questions que l'on peut se poser lorsqu'on fait de la modélisation d'un phénomène géophysique.

En notant  $m$  son espérance et  $\Sigma$  sa matrice de variance-covariance, la fonction caractéristique de  $Y$  est

$$\Phi(u) = \exp\left(iu^t m - \frac{1}{2}u^t \Sigma u\right)$$

Ceci exprime en particulier que  $m$  et  $\Sigma$  suffisent à caractériser la loi du vecteur.

### 1.3.2.1 Existence d'un vecteur aléatoire gaussien admettant une covariance donnée

**Théorème 1** [18] *Soit  $m \in \mathbb{R}^n$  et  $\Sigma$  une matrice  $n \times n$  symétrique définie positive (i.e  $u^t \Sigma u > 0, \forall u \neq 0 \in \mathbb{R}^n$ ), il existe un vecteur aléatoire gaussien d'espérance  $m$  et de matrice de covariance  $\Sigma$ .*

### 1.3.2.2 Densité d'un vecteur aléatoire gaussien

**Théorème 2** [18] *Si  $Y$  est un vecteur gaussien à  $n$  dimensions de matrice de covariance  $\Sigma$  régulière,  $Y$  admet sur  $\mathbb{R}^n$  la densité :*

$$g(y_1, \dots, y_n) = \frac{1}{(2\pi)^{\frac{n}{2}} \sqrt{\det \Sigma}} \exp\left[-\frac{1}{2}(y - m)^t \Sigma^{-1} (y - m)\right]$$

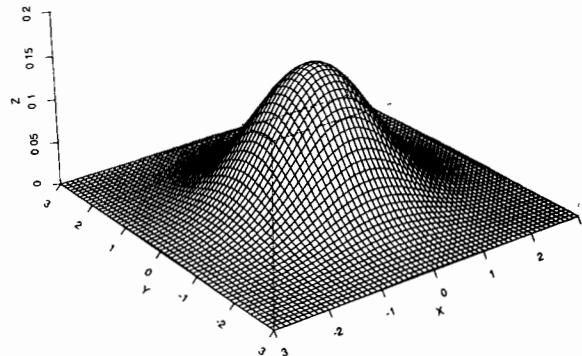


FIG. 1.2: Densité d'un couple bi-gaussien de coefficient de corrélation nul (indépendance)

### 1.3.3 Fonction aléatoire normale

**Définition 3** *Une fonction aléatoire  $Y$  sur  $\mathbb{R}^2$  est dite normale ou gaussienne si pour tout  $n$ -uplet  $(x_1, \dots, x_n)$  de points de  $\mathbb{R}^2$  le vecteur aléatoire  $(Y(x_1), \dots, Y(x_n))$  est normal au sens de la définition 2, ceci pour tout  $n \in \mathbb{N}$ .*

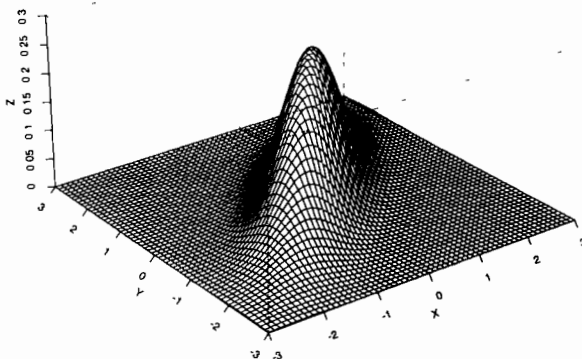


FIG. 1.3: Densité d'un couple bi-gaussien de coefficient de corrélation  $\rho = 0,8$

Cette définition n'est pas très parlante au sens où elle caractérise une classe de fonctions aléatoires de manière non constructive. On peut se représenter d'une manière un peu différente les fonctions aléatoires gaussiennes grâce au théorème central limite pour les fonctions aléatoires selon lequel

**Théorème 3** [66] *si  $(Y_n)_{n \in \mathbb{N}}$  est une suite de fonctions aléatoires indépendantes, identiquement distribuées, centrées réduites d'ordre deux (mais de loi quelconque), alors  $(Y_1 + \dots + Y_n)/\sqrt{n}$  converge en loi vers une fonction aléatoire gaussienne, lorsque  $n$  tend vers  $+\infty$ .*

On voit donc que les fonctions aléatoires gaussiennes apparaissent dès que l'on combine de manière additive des fonctions aléatoires de loi quelconque pas trop dispersées. Il se trouve toutefois qu'en dépit de la manière synthétique dont on peut les définir et les caractériser, les f.a. gaussiennes peuvent prendre des formes assez diverses [2, 66]. Elles constituent donc une classe assez riche de modèles dépendant d'un petit nombre de paramètres fonctionnels, ce qui justifie leur usage dans des contextes aussi variés que le génie minier [86], l'océanographie [21], la théorie de la percolation [56], la finance [79, 80] ou encore la physique théorique [1].

### 1.3.3.1 Existence d'une fonction aléatoire gaussienne admettant une covariance donnée

**Théorème 4** [18] *Une application  $C : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}$  est la covariance d'une fonction aléatoire sur  $\mathbb{R}^2$  si  $C$  vérifie :  $\forall n \in \mathbb{N}, \forall (\lambda_1, \dots, \lambda_n), \lambda_i \in \mathbb{R}, \forall (x_1, \dots, x_n), x_i \in \mathbb{R}^2,$*

$$\sum_{i=1}^n \sum_{j=1}^n \lambda_i C(x_i, x_j) \lambda_j \geq 0$$

**Théorème 5** [18] *Soit  $C$  une fonction de covariance, il existe une fonction aléatoire gaussienne admettant la covariance  $C$*

### 1.3.3.2 Caractérisation

La loi spatiale d'une fonction aléatoire gaussienne  $Y$  est entièrement spécifiée par la donnée de son espérance et de sa covariance, respectivement :

$$m_Y(x) = E[Y(x)] \quad (1.1)$$

$$C_Y(x, x+h) = E[(Y(x) - m_Y(x))(Y(x+h) - m_Y(x+h))] \quad (1.2)$$

## 1.4 Fonctions aléatoires gaussiennes anamorphosées

Une propriété incontournable de la distribution gaussienne est d'être symétrique par rapport à sa moyenne. Cette propriété devient évidemment gênante dès que l'on s'intéresse à des phénomènes possédant des distributions très dissymétriques. Plus généralement la représentation d'un phénomène physique par une fonction aléatoire gaussienne est incorrecte dès que la distribution monovariée n'est plus approximativement gaussienne. Pour remédier à ce type de problèmes, on peut se doter d'une classe un peu plus générale de modèles de f.a, en considérant les images de f.a gaussiennes par une transformation de  $\mathbb{R}$  dans  $\mathbb{R}$ . Nous allons voir que ceci permet de régler le problème de la distribution monovariée mais que cela complique la procédure d'estimation des paramètres. On peut également noter que ce type de transformations ne permettra pas en général de régler à la fois les problèmes relatifs à la loi monovariée et ceux relatifs à la loi bivariable. Pour disposer d'une fonction aléatoire admettant une loi monovariée gaussienne et une loi bivariable qui ne soit pas du type bi-gaussien anamorphosé, il faudrait se tourner vers une autre famille de modèles, par exemple le *modèle Gamma* [54], ce qui dépasse largement le cadre de ce travail.

### 1.4.1 Anamorphose d'une variable aléatoire

**Définition 4** *Soit  $Y$  une variable aléatoire réelle admettant une fonction de répartition  $G$  continue (i.e la loi de  $G$  est sans atome) et  $F$  une fonction de répartition quelconque.*

*On note  $\phi : \mathbb{R} \rightarrow \mathbb{R}$  l'application définie par  $\phi(u) = \text{Inf}_{x \in \mathbb{R}} \{x : F(x) \geq G(u)\}$ . (voir figure 1.4.1) Notons que  $\phi$  est non décroissante et monotone et  $\forall x \in \mathbb{R} \text{Pr}[\phi(Y) \leq x] = F(x)$ . De plus si  $F$  est continue (pas d'atome),  $\phi$  est inversible et  $\phi^{-1} = G^{-1} \circ F$ .*

*On désignera par anamorphose toute fonction non décroissante monotone transportant une loi continue vers une autre loi (continue ou non).*

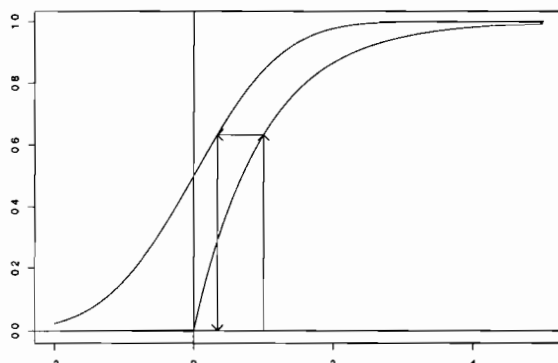


FIG. 1.4: Anamorphose : transformation d'une loi gamma en loi gaussienne

#### 1.4.2 Anamorphose d'une fonction aléatoire gaussienne

**Définition 5** Une fonction aléatoire  $Z$  sur  $\mathbb{R}^2$  sera dite gaussienne anamorphosée s'il existe une fonction non décroissante monotone  $\phi: \mathbb{R} \rightarrow \mathbb{R}$  et une fonction aléatoire gaussienne  $Y$  sur  $\mathbb{R}^2$  telles que  $Z \stackrel{L}{=} \phi(Y)$ . Cette dernière égalité en loi peut être remplacée par une égalité presque sûre, ce qui sera équivalent en pratique puisque  $Y$  sera toujours un objet fictif que l'on introduit pour fixer les idées.

Remarque: Une fonction aléatoire  $Z$  est dite bi-gaussienne anamorphosée si on a  $Z \stackrel{L}{=} \phi(Y)$  pour une f.a  $Y$  dont le caractère gaussien n'est assuré que pour les lois bivariées.

##### 1.4.2.1 Caractérisation dans le cas stationnaire

**Définition 6** Une fonction aléatoire  $Z$  est dite stationnaire au sens fort si sa loi spatiale est invariante par translation, i.e  $\forall h \in \mathbb{R}^2$

$$Pr[Z(x_1) \leq z_1, \dots, Z(x_n) \leq z_n] = Pr[Z(x_1 + h) \leq z_1, \dots, Z(x_n + h) \leq z_n]$$

Elle est dite stationnaire d'ordre deux si l'invariance par translation ne s'applique qu'à la fonction de covariance.

Si  $Z$  est une fonction aléatoire gaussienne anamorphosée stationnaire d'ordre deux, elle l'est aussi au sens fort. On peut supposer qu'elle est de la forme  $Z = \phi(Y)$  où  $Y$  est stationnaire centrée réduite. On peut donc paramétrer la loi de  $Z$  par sa fonction de répartition univariée et sa covariance (qui sont stationnaires).

## 1.5 Problème de cohérence interne

### 1.5.1 Position du problème

Nous avons vu au paragraphe 1.4.2.1 que l'on pouvait paramétrer une fonction aléatoire gaussienne anamorphosée stationnaire par sa loi monovariée et sa covariance. Cette paramétrisation est intéressante parce qu'elle sépare le problème selon deux aspects apparemment tout à fait distincts, celui de la distribution des valeurs ponctuelles et celui de la structure spatiale. Il se trouve en fait que ces deux aspects sont liés, c'est à dire que *le choix d'un des deux paramètres implique certaines conditions pour le deuxième*.

On montre par exemple [92] que dans  $\mathbb{R}$ , il n'existe pas de F.A admettant une loi bi-variable log-bigaussienne et une covariance de la forme  $A \cos(\alpha h)$ . Ceci en dépit du fait que la fonction  $h \rightarrow A \cos(\alpha h)$  soit de type positif. En effet, la log-binormalité implique que la corrélation demeure strictement supérieure à  $-1$ , alors qu'avec la covariance cosinus, l'égalité est atteinte.

Un autre exemple, plus intéressant (dans la mesure où il ne fait pas intervenir la loi bi-variable) est fourni par les ensembles aléatoires (ou fonctions aléatoires admettant une loi monovariée de Bernoulli, en 0 ou 1). Il s'agit de fonctions aléatoires non dérivables en moyenne quadratique [87, 91]. La régularité de leur covariance à l'origine est au plus du type  $A|h|^\lambda$  avec  $\lambda \leq 1$ . Les covariances des ensembles aléatoires doivent donc vérifier une condition strictement plus forte que la condition générale du théorème 4. Mais on ne sait pas à quelle condition une fonction de type positif est la covariance d'un ensemble aléatoire<sup>4</sup>.

Dès que l'on spécifie séparément différentes propriétés d'un modèle, la question se pose donc de savoir s'il existe effectivement des fonctions aléatoires vérifiant simultanément ces propriétés. Le problème est général et difficile. Certaines solutions ont été données dans des cas particuliers, nous allons voir qu'elles sont inadaptées à la modélisation des champs de pluie.

### 1.5.2 Un théorème de compatibilité

On a les résultats suivants :

**Théorème 6** [85, 92] *Les lois monovariées indéfiniment divisibles<sup>5</sup> sont compatibles avec les covariances qui sont combinaison linéaire finie à coefficients positifs de covariogrammes géométriques<sup>6</sup>.*

4. Matheron [93] a exhibé une propriété nécessaire conjecturée suffisante.

5. Une probabilité  $\mu$  est indéfiniment divisible si  $\forall n$ , on peut exhiber une loi  $\mu_n$  telle que  $\mu$  se représente par le produit de convolution  $\mu = \mu_n * \dots * \mu_n$ . [63]

6. Un covariogramme géométrique est une fonction de la forme  $g(h) = \int_{\mathbb{R}^2} \mathbb{I}_S(x) \mathbb{I}_{S_h}(x) dx$ , où  $S$  (resp.  $S_h$ ) est un domaine du plan (resp. son translaté d'un vecteur  $h$ ) et  $\mathbb{I}_S$  (resp.  $\mathbb{I}_{S_h}$ ) l'indicatrice de  $S$  (resp. de  $S_h$ ) [119].

La classe des lois indéfiniment divisibles est très vaste : la loi de Poisson, les lois gamma et les lois stables [63], la loi log-normale [17] sont indéfiniment divisibles. Par ailleurs, la covariance sphérique est un covariogramme géométrique dans  $\mathbb{R}^2$ . On dispose donc d'un catalogue assez vaste de couples  $(F, C)$ , où  $F$  est une loi monovariante et  $C$  une covariance, qui sont compatibles. Mais le théorème ne dit pas si de tels couples sont compatibles avec *le modèle gaussien anamorphosé*.

On peut voir que cela n'a rien d'automatique, par exemple en ce qui concerne la loi log-normale, très utilisée en hydro-météorologie, en effet :

**Théorème 7** [92] *Dans  $\mathbb{R}^n$ , il existe un réel  $\alpha_0$  tel qu'il n'existe pas de f.a log-bigaussienne admettant une covariance sphérique et un coefficient de variation supérieur à  $\alpha_0$ .*

### 1.5.3 Fonction de répartition continue

Lorsqu'on a un jeu de données  $\{z(x_1), \dots, z(x_n)\}$ , et qu'on a ajusté une loi monovariante  $F_Z$  et une covariance  $C$  on n'est donc jamais certain qu'elles soient compatibles avec le modèle gaussien anamorphosé. Mais dans le cas où la loi ajustée est sans atome, on peut toujours revenir à la gaussienne sous-jacente par l'anamorphose inverse  $\phi^{-1} = G \circ F_Z^{-1}$ . On peut alors ajuster une covariance  $\hat{C}_Y$  à ce nouveau jeu de données sans contraintes particulières, puisque ces données sont supposées issues d'une fonction aléatoire gaussienne. On revient ensuite aux données initiales, par la relation

$$C_Z = \sum_{k=1}^{k=\infty} \frac{\phi_k^2}{k!} C_Y^k \quad (1.3)$$

où les  $\phi_k$  sont les coefficients du développement de  $\phi$  en polynômes d'Hermite [40, 112].

Cette méthode est résumée par le diagramme suivant :

$$\begin{array}{ccc} z_n(x_p) & \xrightarrow{1} & y_n(x_p) \\ & & \downarrow 2 \\ \hat{C}_Z & \xleftarrow{3} & \hat{C}_Y \end{array}$$

### 1.5.4 Fonction de répartition discontinue

Lorsque dans un jeu de données, une valeur apparaît un grand nombre de fois et qu'on a des raisons de penser qu'elle joue un rôle particulier, on peut lui attribuer une probabilité d'occurrence non nulle. C'est le cas des valeurs nulles

dans les champs de pluie. Nous devons donc examiner le cas où la loi ajustée est discontinue.

Dans ce cas, l'anamorphose n'est pas injective, on ne peut pas revenir à la gaussienne sous-jacente. En revanche, on peut observer que l'application

$$\begin{aligned} \zeta : [0, 1] &\longrightarrow \mathbb{R} \\ u &\longrightarrow \sum_{k \geq 1} \frac{\phi_k^2}{k!} u^k \end{aligned} \quad (1.4)$$

qui relie  $C_Y$  et  $C_Z$  dans l'équation 1.3 est toujours inversible, en tant que fonction continue strictement croissante sur  $[0, 1]$ .

En l'absence de corrélations négatives, on peut donc procéder aux calculs suivants :

1. ajuster une covariance aux données brutes, que l'on notera  $K_Z$
2. calculer l'image  $K_Y$  de  $K_Z$  par  $\zeta^{-1}$
3. ajuster une covariance à  $K_Y$  que l'on notera  $\widehat{C}_Y$
4. revenir aux données brutes en calculant l'image  $\widehat{C}_Z$  de  $\widehat{C}_Y$  par  $\zeta$

$$\begin{array}{ccc} z_n(x_p) & & \\ \downarrow 1 & & \\ K_Z & \xrightarrow{2} & K_Y \\ & & \downarrow 3 \\ \widehat{C}_Z & \xleftarrow{4} & \widehat{C}_Y \end{array}$$

Il convient de noter que l'étape 4 est nécessaire car  $K_Y$  n'est pas nécessairement une fonction de type positif. En procédant de cette manière, on est alors assuré d'obtenir un couple  $(F_Z, \widehat{C}_Z)$  qui est compatible avec le modèle gaussien anamorphosé. En revanche, on ne disposera pas en général d'une expression analytique pour  $\widehat{C}_Z$ .

## 1.6 Mise en œuvre numérique

### 1.6.1 Simulation non conditionnelle

Le problème de la simulation de fonctions aléatoires gaussiennes anamorphosées se réduit à la simulation de fonctions aléatoires gaussiennes, dans la mesure où on déduit l'anamorphosée de la gaussienne par une transformation uniforme sur tout le domaine.

Il existe un très grand nombre de méthodes de simulations de fonctions aléatoires gaussiennes. On peut distinguer principalement trois types :

**les méthodes matricielles** basées sur la décomposition  $LU$  de la matrice de covariance [4] et leurs différentes variantes séquentielles [15]. Notons toutefois que la méthode générale n'est pas à proprement parler une méthode de simulation de fonctions aléatoires sur  $\mathbb{R}^n$  puisque l'algorithme ne tire pas partie de la nature spatiale du problème.

**Les méthodes d'épandage** incluant les méthodes spectrales [121, 122] (basées sur la représentation de Cramér de la fonction aléatoire [132]), et la méthode des bandes tournantes [88] ;

**Les méthodes géométriques** moins générales que les précédentes, telles que la méthode de dilution poissonnienne [60] mais parfois plus proches du phénomène qu'une représentation abstraite comme celle de Cramér.

Parmi celles-ci, la méthode des bandes tournantes est notoirement la plus rapide dans le cas non-conditionnel.

L'algorithme revient à simuler des fonctions aléatoires indépendantes identiquement distribuées sur un faisceau de  $n$  droites dont les directions sont régulièrement réparties sur  $[0, 2\pi]$ .

La relation entre la covariance des f.a sur les lignes et celle de la f.a à simuler est une équation intégrale d'Abel, pour laquelle on n'a pas de solution analytique en général. Une approche détournée (utilisée au cours du présent travail) consiste à relier non pas les covariances des deux f.a mais leurs densités spectrales [84]. Le problème est alors de simuler une f.a sur  $\mathbb{R}$  admettant une densité spectrale donnée. En général (et dans le cadre de ce travail), la densité spectrale de la f.a sur les lignes est continue et à support non borné. Dans ce cas, on ne dispose pas de méthodes numériques exactes de simulation au sens où il est nécessaire de discrétiser le spectre et de le tronquer [121].

Les conséquences de ces deux approximations numériques sont respectivement une régularisation de la covariance (convolution) et une périodisation de la f.a. La fréquence à laquelle il faut tronquer le spectre dépend évidemment des paramètres de la covariance (plus la portée est courte, plus les hautes fréquences sont porteuses d'énergie).

Il est donc impossible de donner une valeur générale pour le choix de ces paramètres. Mais une manière de s'affranchir (partiellement) de ce type de problème

est de réaliser une discrétisation non uniforme du spectre, c'est à dire de prendre beaucoup d'harmoniques dans les régions du spectre où la densité spectrale varie beaucoup (en général dans les basses fréquences) et peu d'harmoniques ailleurs (en particulier dans les hautes fréquences qui peuvent contenir beaucoup d'énergie).

Notons toutefois que Dietrich a récemment explicité le lien entre la covariance de la f.a sur les lignes et celle de la f.a en dimension 2 pour une famille assez générale de fonctions de covariance [35, 43, 36].

### 1.6.2 Conditionnement par une moyenne d'espace

La désagrégation spatiale revient à simuler une fonction aléatoire - gaussienne anamorphosée dans le cas présent - sous contrainte d'une moyenne d'espace sur le domaine de simulation.

Le problème du conditionnement en général a reçu une grande attention dans la littérature. Le krigeage conditionnant permet de simuler une fonction aléatoire gaussienne sous contrainte de valeurs ponctuelles [88], mais cette méthode ne se généralise pas immédiatement au cas gaussien anamorphosé lorsque la fonction de répartition est discontinue.

En dehors du conditionnement par des valeurs ponctuelles, on connaît certains algorithmes pour le conditionnement par des contraintes non-linéaires, mais ces méthodes concernent le cas gaussien.

De manière générale, on a des solutions à ces problèmes dans le cas gaussien parce que c'est une classe de fonctions aléatoires stables par conditionnement par des contraintes linéaires.

On peut envisager différents algorithmes (au moins du point de vue théorique) pour conditionner une f.a gaussienne par une moyenne d'espace. Par exemple on peut tirer parti du fait que, si  $Y$  est une f.a gaussienne sur  $\mathbb{R}^2$  et  $D$  un domaine du plan,  $Y(x) \mid \int_D Y(x)dx$  est encore gaussienne.

Le calcul de la loi de  $Y(x)$  conditionnellement à l'intégrale ne requiert que le calcul de l'espérance et de la variance conditionnelles.

Malheureusement, la moyenne d'espace conditionnante dans notre problème (la sortie de M.C.G.A ou l'estimation satellitale) est relative à la f.a de l'espace direct, et pas à la gaussienne sous-jacente. Et comme l'anamorphose est une fonction *non linéaire*, on ne peut pas transposer la contrainte portant sur la f.a de l'espace direct en une contrainte sur la f.a gaussienne sous-jacente. Il est donc difficile d'adapter les méthodes valables pour le cas gaussien au cas gaussien anamorphosé.

#### 1.6.2.1 Méthode directe

La méthode la plus directe à laquelle on peut songer consiste à tirer un vecteur aléatoire uniforme dans l'hypercube  $[0, 1]^n$  où  $n$  est le nombre de points de la grille à simuler. On calcule ensuite l'image de ce vecteur par l'anamorphose qui transporte la loi  $\mathcal{U}([0, 1]^n)$  sur la loi de  $Z \mid \bar{Z}$ , où  $\bar{Z}$  est la moyenne d'espace conditionnante [32].

Cette méthode a été mise en œuvre récemment [125] dans un cadre non paramétrique sur des courtes séries de données hydrologiques. Elle est excessivement lourde à mettre en œuvre puisqu'elle nécessite le calcul (numérique) de la loi conditionnelle dans un espace de très grande dimension.

#### **1.6.2.2 Méthode de rejet**

Il n'est pas concevable d'utiliser la méthode précédente dans le cadre de ce travail qui, comme on le verra à la section 4 concernera des grilles de dix milles points.

Quand cela a été nécessaire, le conditionnement a été réalisé en simulant un grand nombre de réalisations et en choisissant les réalisations satisfaisant la contrainte, à un degré de précision près.

## Chapitre 2

# Quelques points de méthode sur la fonction de structure

### Présentation

On a montré dans le précédent chapitre que les fonctions aléatoires gaussiennes anamorphosées pouvaient être paramétrées par leur loi monovariante et leur covariance.

Le présent chapitre se place sur un plan plus pratique en abordant le problème de l'utilisation de l'information disponible (les lames d'eau de l'événement aux points d'un réseau irrégulier) pour calculer un estimateur de la covariance d'une fonction aléatoire. Les résultats exposés sont classiques mais leur rappel permet de clarifier l'usage des règles de calcul héritées de la géostatistique minière en hydro-météorologie. En particulier nous verrons que certaines limitations qui prévalent lorsque l'on dispose d'une seule réalisation, disparaissent dès que l'on dispose de plusieurs réalisations.

## 2.1 Définitions

Étant donné une fonction aléatoire  $Z$  définie sur  $\mathbb{R}^2$ , dont les accroissements sont de variance finie, i.e vérifiant :

$$\text{Var}[Z(x) - Z(y)] < +\infty \quad \forall x, y \in \mathbb{R}^2$$

on définit le variogramme  $\gamma$  de  $Z$  par

$$\gamma(x, y) = \frac{1}{2} \text{Var}[Z(x) - Z(y)] \quad (2.1)$$

Dans le cas où  $Z$  est elle-même de variance finie, on notera

$$\mu(x) = E[Z(x)] \quad (2.2)$$

$$\sigma^2(x) = \text{Var}[Z(x)] \quad (2.3)$$

et on peut définir la covariance  $C$  par :

$$C(x, y) = E[(Z(x) - \mu(x))(Z(y) - \mu(y))]$$

On a alors la relation :

$$\gamma(x, y) = \frac{1}{2} [\sigma^2(x) + \sigma^2(y)] - C(x, y) \quad (2.4)$$

## 2.2 Estimateurs

Si la définition de la fonction de structure est sans ambiguïté, la manière dont on l'estime en pratique dépend de la nature de l'information disponible. On peut disposer d'un seul champ, comme dans le domaine minier, ou au contraire de plusieurs champs considérés comme des réalisations indépendantes d'une même fonction aléatoire. Dans ce dernier cas, les mesures peuvent avoir été faites aux mêmes points pour chaque réalisation (homotopie) ou en des points variant d'une réalisation à l'autre (hétérotopie).

Dans un contexte mono-réalisation on notera  $\{z(x_j)\}_{1 \leq j \leq p}$  les mesures obtenues aux points  $x_1, \dots, x_p$ .

Dans un contexte multi-réalisation, on fera référence aux  $n \times p$  données par  $\{z_i(x_j)\}_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}$

### 2.2.1 Mono-réalisation

Dans ce cas on ne peut pas estimer le variogramme exactement aux points informés, on peut seulement calculer une approximation du variogramme dans un voisinage des points informés.

On estime  $\gamma(x, y)$  par

$$\begin{aligned} \gamma^*(x, y) &= \frac{1}{2|v_{x,y}|} \sum_{x_i, y_i \in v_{x,y}} [z(x_i) - z(y_i)]^2 \\ &\quad - \frac{1}{2|v_{x,y}|^2} \left[ \sum_{x_i, y_i \in v_{x,y}} z(x_i) - z(y_i) \right]^2 \end{aligned} \quad (2.5)$$

où  $v_{x,y}$  désigne l'ensemble des couples de points informés  $(x_i, y_i)$  (où  $x_i$  est proche de  $x$  et  $y_i$  proche de  $y$ ), et  $|v_{x,y}|$  son cardinal.

### 2.2.2 Multi-réalisation

En cas d'homotopie, on dispose de plusieurs réalisations de  $Z$  aux points informés, on peut donc évaluer directement  $\gamma$  pour tous les couples  $(x_{j_1}, x_{j_2})$ . On estime  $\gamma(x_{j_1}, x_{j_2})$  par

$$\hat{\gamma}(x_{j_1}, x_{j_2}) = \frac{1}{2} \left[ \frac{1}{N} \sum_{i=1}^N [z_i(x_{j_1}) - z_i(x_{j_2})]^2 - \frac{1}{N^2} \left[ \sum_{i=1}^N z_i(x_{j_1}) - z_i(x_{j_2}) \right]^2 \right] \quad (2.6)$$

De même, on estime  $C(x_{j_1}, x_{j_2})$  par,

$$\hat{C}(x_{j_1}, x_{j_2}) = \frac{1}{N} \sum_{i=1}^N z_i(x_{j_1}) z_i(x_{j_2}) - \frac{1}{N^2} \sum_{i=1}^N z_i(x_{j_1}) \sum_{i=1}^N z_i(x_{j_2}) \quad (2.7)$$

De plus, en notant

$$\hat{\sigma}^2(x_j) = \frac{1}{N} \sum_{i=1}^N z_i(x_j)^2 - \left( \frac{1}{N} \sum_{i=1}^N z_i(x_j) \right)^2 \quad (2.8)$$

Les valeurs estimées de  $\hat{C}(x_{j_1}, x_{j_2})$  et  $\hat{\gamma}(x_{j_1}, x_{j_2})$  sont liées par

$$\hat{\gamma}(x_{j_1}, x_{j_2}) = \frac{1}{2} (\hat{\sigma}^2(x_{j_1}) + \hat{\sigma}^2(x_{j_2})) - \hat{C}(x_{j_1}, x_{j_2}) \quad (2.9)$$

Ainsi, à partir du moment où l'hypothèse d'existence des moments d'ordre deux est correcte, on peut étudier indifféremment le variogramme ou la covariance et passer du variogramme à la covariance (estimés dans un contexte multi-réalisation) via les variances empiriques.

### 2.3 Fonction de structure en présence de dérive

Considérons une fonction aléatoire  $Z$ , éventuellement non stationnaire, et  $d$  une fonction non aléatoire définies sur  $\mathbb{R}^2$ .

Notons  $\gamma_Z$  et  $\gamma_{Z+d}$  les variogrammes respectifs de  $Z$  et  $Z+d$ . Comme la variance n'est pas modifiée par l'addition d'une valeur non aléatoire, il est clair que

$$\gamma_Z = \gamma_{Z+d} \quad (2.10)$$

Autrement dit, le variogramme n'est pas modifié par l'addition d'une dérive, pas plus que la covariance d'après la relation 2.4, soit :

$$C_Z = C_{Z+d} \quad (2.11)$$

### 2.3.1 Estimation à l'aide d'une seule réalisation

A l'aide de l'estimateur défini précédemment, on obtient

$$\begin{aligned}
 \gamma_{Z+d}^*(x, y) &= \gamma_Z^*(x, y) \\
 &+ \frac{1}{2|v_{x,y}|} \sum_{v_{x,y}} (d(x_i) - d(y_i))^2 - \frac{1}{2|v_{x,y}|^2} \left( \sum_{v_{x,y}} d(x_i) - d(y_i) \right)^2 \\
 &+ \frac{1}{|v_{x,y}|} \sum_{v_{x,y}} (z(x_i) - z(y_i))(d(x_i) - d(y_i)) \\
 &- \frac{1}{|v_{x,y}|^2} \sum_{v_{x,y}} z(x_i) - z(y_i) \sum_v d(x_i) - d(y_i)
 \end{aligned} \tag{2.12}$$

ce que l'on notera de manière synthétique

$$\gamma_{Z+d}^*(x, y) = \gamma_Z^*(x, y) + \gamma_d^*(x, y) + 2\Gamma_{Z,d}(x, y) \tag{2.13}$$

Ce résultat est à rapprocher du fait que  $\gamma_{Z+d} = \gamma_Z$  (éq. 2.10). On retrouve ici un fait bien connu : le variogramme expérimental est perturbé par l'addition d'une dérive lorsqu'on l'estime à partir d'une seule réalisation. En particulier, si l'on observe un phénomène dont la moyenne présente une variation systématique dans l'espace (dérive), mais pas la variance, le variogramme expérimental calculé s'apparentera à un variogramme non stationnaire alors que le modèle approprié est un modèle stationnaire.

### 2.3.2 Estimation à l'aide de plusieurs réalisations

On peut faire un calcul analogue ou plus simplement constater que dans le cas où l'on dispose de plusieurs réalisations, l'estimateur du variogramme aux points  $(x_{j_1}, x_{j_2})$  apparaît comme une variance, celle d'une v.a dont la loi est la loi empirique des accroissements  $Z(x_{j_1}) - Z(x_{j_2})$ . Par conséquent

$$\hat{\gamma}_{Z+d}(x, y) = \hat{\gamma}_Z(x, y) \tag{2.14}$$

Dans ce cas rien ne s'oppose à l'estimation du variogramme de  $Z$ , même lorsque la dérive n'est pas connue.

## 2.4 Sur le variogramme climatologique

Pour l'interpolation des champs de pluie, et plus généralement dans l'étude des phénomènes reconnus par plusieurs réalisations, on peut se demander si, pour

l'interpolation d'un champ particulier, le modèle de variogramme doit être choisi au vu

- du variogramme mono-réalisation  $\gamma^*$  estimé à partir du champ à interpoler,
- du variogramme  $\hat{\gamma}$  multi-réalisation,
- d'une moyenne pondérée des variogrammes mono-réalisations  $\gamma_1^*, \dots, \gamma_n^*$ .

En hydro-météorologie, les champs présentant une forte dispersion spatiale sont souvent les champs ayant produit une lame d'eau moyenne importante. Mais ils sont peu nombreux.

On désigne alors par variogramme climatologique des champs de pluie, la moyenne des variogrammes mono-réalisation pondérée par l'inverse des variances spatiales respectives des champs [12, 71]. Ceci s'écrit :

$$\gamma_{clim}^*(x, y) = \sum_{i=1}^n \frac{1}{s_i^2} \gamma_i^*(x, y) \quad (2.15)$$

Où

$$s_i^2 = \frac{1}{p} \sum_{j=1}^p (z_i(x_j) - m_i)^2 \quad (2.16)$$

Par cette pondération, le variogramme obtenu est peu influencé par la structure de ces champs très dispersés.

L'intérêt de ce variogramme doit être évalué à partir de données en fonction de la précision des estimations obtenues, toutefois, du point de vue théorique et dans la perspective de réaliser des simulations, l'usage de ce variogramme présente un certain nombre de défauts :

#### Du point de vue de l'inférence :

- En normant chaque variogramme mono-réalisation dans le calcul de la moyenne (éq. 2.15), on obtient un variogramme dont le palier est toujours égal à 1. On n'a donc pas de repères pour comparer la dispersion des accroissements aux grandes distances avec la variance ponctuelle ;
- En tant que moyennes de variogrammes mono-réalisations, le variogramme climatologique est influencé par la présence d'une dérive (Cf éq. 2.13). Son utilisation peut conduire à interpréter à tort un variogramme expérimental semblant croître indéfiniment comme une non-stationnarité de la variance ;

#### Du point de vue de l'interprétation :

- C'est un variogramme qui dépend du domaine sur lequel on le calcule (à travers les dispersions  $s_i^2$ ), les conclusions qu'on peut tirer de son

observation sont propres à ce domaine, en particulier à sa taille par rapport aux échelles caractéristiques du phénomène ;

**Du point de vue de la simulation :**

- Si on simule une fonction aléatoire en prescrivant un modèle de variogramme qui aura été ajusté sur le variogramme climatologique (éventuellement renormé par une variance ponctuelle stationnaire), on n'obtient pas une fonction aléatoire ayant la même structure que les données (et dans le cas gaussien, pas les mêmes lois bi-variables, tri-variables, etc.)

## Chapitre 3

# Analyse des données à l'échelle de l'événement

### Présentation

Dans le chapitre précédent, on a donné quelques raisons qui motivent le choix du modèle gaussien anamorphosé stationnaire pour représenter les champs de pluie au pas de temps de l'événement. L'objectif de ce chapitre est d'en estimer les paramètres. Il ne s'agira pas à cette étape de déterminer si le modèle est approprié ou pas, mais seulement de faire le meilleur choix possible de paramètres à l'intérieur d'un cadre méthodologique donné. L'estimation des paramètres permettra d'améliorer nos connaissances de certaines propriétés des champs pluie (notamment de quantifier des phénomènes tels que l'anisotropie ou l'occurrence de valeurs extrêmes), d'interpréter ces propriétés dans un cadre théorique bien déterminé et de réaliser des simulations dans ce cadre. L'analyse est menée en deux temps selon la procédure décrite à la section 1.5.4 . La première étape consiste à ajuster une loi monovariante commune aux 30 stations du réseau de base (et par suite, à tous les points du domaine d'étude), il s'agira d'une loi Gamma mélangée avec des valeurs nulles. Le deuxième point consiste à choisir une fonction de covariance. On retiendra une somme de deux covariances exponentielles anisotropes d'échelles différentes.

## 3.1 Distribution monovariante

### 3.1.1 Introduction

On dispose de 258 événements mesurés en 30 stations du réseau de base pendant 6 ans. On va se placer dans un cadre paramétrique. Ceci peut constituer

une première approximation préalable à la mise en œuvre de méthodes plus sophistiquées, nécessaire tant que l'on ne connaît pas les défauts prédominants du modèles.

### 3.1.2 Choix d'une loi et estimation des paramètres

En chaque station  $x_j$  on dispose d'une série  $(r_i(x_j))_{i=1,n}$  à laquelle on veut ajuster une fonction de répartition, commune à toutes les stations.

On définit les fonctions de répartition empiriques en chaque station par

$$F_j^*(u) = \frac{1}{n} \sum_{i=1}^n \mathbb{I}_{r_i(x_j) \leq u} \quad (3.1)$$

$$\text{où } \mathbb{I}_{r_i(x_j) \leq u} = \begin{cases} 1 & \text{si } r_i(x_j) \leq u \\ 0 & \text{sinon.} \end{cases}$$

On prend en compte les valeurs nulles en choisissant une loi présentant un atome à l'origine. La probabilité d'occurrence d'une valeur nulle est estimée par

$$\sum_{j=1}^p F_j^*(0) = \frac{1}{np} \sum_{i=1}^n \sum_{j=1}^p \mathbb{I}_{r_i(x_j) \leq 0} \quad (3.2)$$

Les séries de valeurs non nulles sont fortement dissymétriques, nous allons les représenter par une loi gamma, dont la densité est de la forme

$$f_{\alpha,\lambda}(x) = \begin{cases} \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} & \text{si } x > 0 \\ 0 & \text{si } x \leq 0 \end{cases}$$

Avec cette paramétrisation on a ,

$$E[R|R > 0] = \frac{\alpha}{\lambda} \quad (3.3)$$

$$Var[R|R > 0] = \frac{\alpha}{\lambda^2} \quad (3.4)$$

Les paramètres de la partie continue de la loi seront estimés à partir des moments des lois empiriques conditionnelles notées  $F_{j|}$

$$m_{j|}^* = \int u dF_{j|}^*(u) \quad (3.5)$$

$$\sigma_{j|}^{2*} = \int (u - m_{j|}^*)^2 dF_{j|}^*(u) \quad (3.6)$$

On estimera  $E[R|R > 0]$  par  $\frac{1}{p} \sum_{j=1}^p m_{j|}^*$

et  $Var[R|R > 0]$  par  $\frac{1}{p} \sum_{j=1}^p \sigma_{j|}^{2*}$

Ce qui conduit à

$$\begin{aligned} P[R > 0] &= 0,26 \\ E[R|R > 0] &= 14.15 \text{ mm} \\ Var[R|R > 0] &= 214.4 \text{ mm}^2 \end{aligned}$$

### 3.1.3 Discussion

#### 3.1.3.1 Test

On peut se faire une idée de la qualité de la loi gamma pour représenter les séries de valeurs positives en calculant les valeurs du test de Kolmogorov-Smirnov :

$$D_n = \text{Sup}_i |F_n^*(r_i) - F(r_i)|$$

qui mesure l'écart entre la distribution empirique pour un échantillon  $(r_1, \dots, r_n)$  de taille  $n$  et la distribution théorique.

La statistique  $D_n$  a été tabulée par Kolmogorov et si l'échantillon est issu de la loi  $F$ , asymptotiquement on a [116], indépendamment de la loi  $F$  :

$$P(\sqrt{n}D_n < y) \xrightarrow{n \rightarrow +\infty} \sum_{k=-\infty}^{+\infty} (-1)^k \exp(-2k^2 y^2) \quad (3.7)$$

Ceci fournit donc un test de

$$\begin{cases} H_0 & : \text{ l'échantillon est issu de la loi } F \\ H_1 & : \text{ l'échantillon n'est pas issu de la loi } F \end{cases}$$

L'allure des fonctions de répartition des séries positives est présentée à la figure 3.1. On peut observer un bon agrément général, néanmoins le test calculé sur notre échantillon de 258 événements rejette  $H_0$  pour 8 stations sur 30 au seuil de 10%. Si les données relatives à chaque station étaient indépendantes les unes des autres, (i.e en cas d'absence de dépendance spatiale), ce taux de rejet serait considérable (car dans le cas où  $H_0$  est vraie, on doit s'attendre à rejeter  $H_0$  à tort une fois sur dix en moyenne). Mais ici, les différentes séries  $(r_i(x_j))_{1 \leq i \leq n}$  ne peuvent pas être considérées comme telles<sup>1</sup>.

Il convient de noter de plus que l'on n'améliore pas la qualité de l'ajustement :

- en choisissant des paramètres locaux (différents pour chaque station) ;

---

1. Ce taux de rejet peut être interprété dans le cadre de la théorie des tests pour les processus [63], mais ceci nécessite de toute façon des hypothèses probabilistes assez fortes sur le phénomène étudié (normalité).

- en ayant recours à d'autres types de loi paramétriques comme la loi log-normale utilisée de manière courante en hydro-météorologie [14, 62, par exemple].

### 3.1.3.2 Valeurs nulles

La carte des fréquences d'occurrence de valeurs nulles est peu compatible avec une hypothèse de stationnarité (figure 3.1.3.2). Les valeurs de la fréquence de pluie nulle sont en effet très dispersées. Le minimum est de 0,19 et le maximum de 0,38. Sous nos hypothèses, chaque fréquence, calculée sur un échantillon de 258 valeurs, suit une loi binomiale  $\mathcal{B}(n = 258, p = 0,26)$ .

Or,

$$P[\mathcal{B}(n = 258, p = 0,26) \geq 0,38] = 8,8 \cdot 10^{-6}$$

et

$$P[\mathcal{B}(n = 258, p = 0,26) \leq 0,19] = 5 \cdot 10^{-3}$$

On pouvait s'attendre à ce type de résultat parce qu'il est bien connu par ailleurs que les isohyètes du cumul annuel sont orientés d'Est en Ouest dans cette région, avec une augmentation du cumul annuel d'environ 1 mm par km [74, 73].

Un fait plus inattendu est que les lignes d'isovaleur de la fréquence de pluie nulle semblent centrées sur le réseau de mesure, avec un minimum global au centre. Ceci peut s'expliquer par le fait que les événements étudiés sont sélectionnés en fonction de la manière dont ils recouvrent la fenêtre d'observation. Les événements qui frôlent la fenêtre (en recouvrant seulement une petite bande au nord ou au sud) ne sont pas sélectionnés.

On peut penser toutefois que le fait de conserver un modèle stationnaire permettra de corriger cet artefact introduit par le mode de sélection des événements pluvieux.

### 3.1.3.3 Saisonnalité

L'analyse réalisée sur tous les événements ne tient pas compte de la position des événements dans la saison, or il a été montré [28] qu'au cœur de la saison des pluies, la lame d'eau ponctuelle par événement était systématiquement plus élevée qu'en début et en fin de saison. La présence de valeurs élevées et de valeurs faibles est bien prise en compte par le modèle mais leur ordre est rendu aléatoire.

La présence de ce type de tendance à l'intérieur de la saison peut avoir des conséquences importantes sur le cycle hydrologique et la végétation, par exemple en entretenant une humidité du sol plus élevée au cœur de la saison.

L'utilisation de ce modèle en mode de simulation (c'est à dire non-conditionné par une information de grande échelle) ne permettra pas de tenir compte de ce type de tendances. En revanche, on peut penser qu'en mode de désagrégation, l'information de grande échelle, elle-même soumise aux fluctuations saisonnières, le permettra de facto.

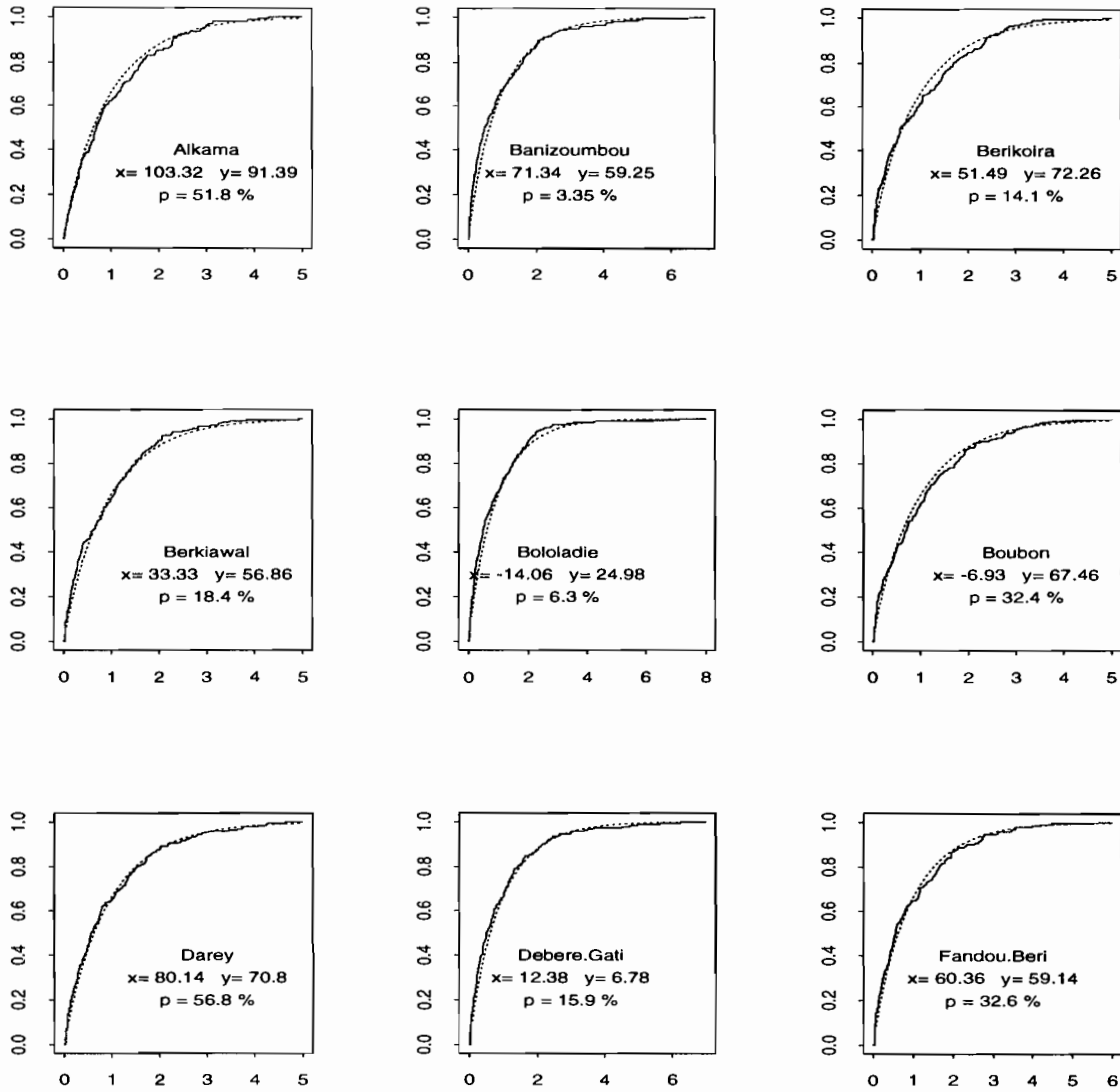


FIG. 3.1: Fonctions de répartition empiriques et théoriques pour les valeurs non-nulles normées des séries de 258 événements. La même loi Gamma est ajustée à toutes les stations. La probabilité associée au test de Kolmogorov-Smirnov (éq. 3.7) est donnée pour chaque station. (L'axe des abscisses est adimensionnel.

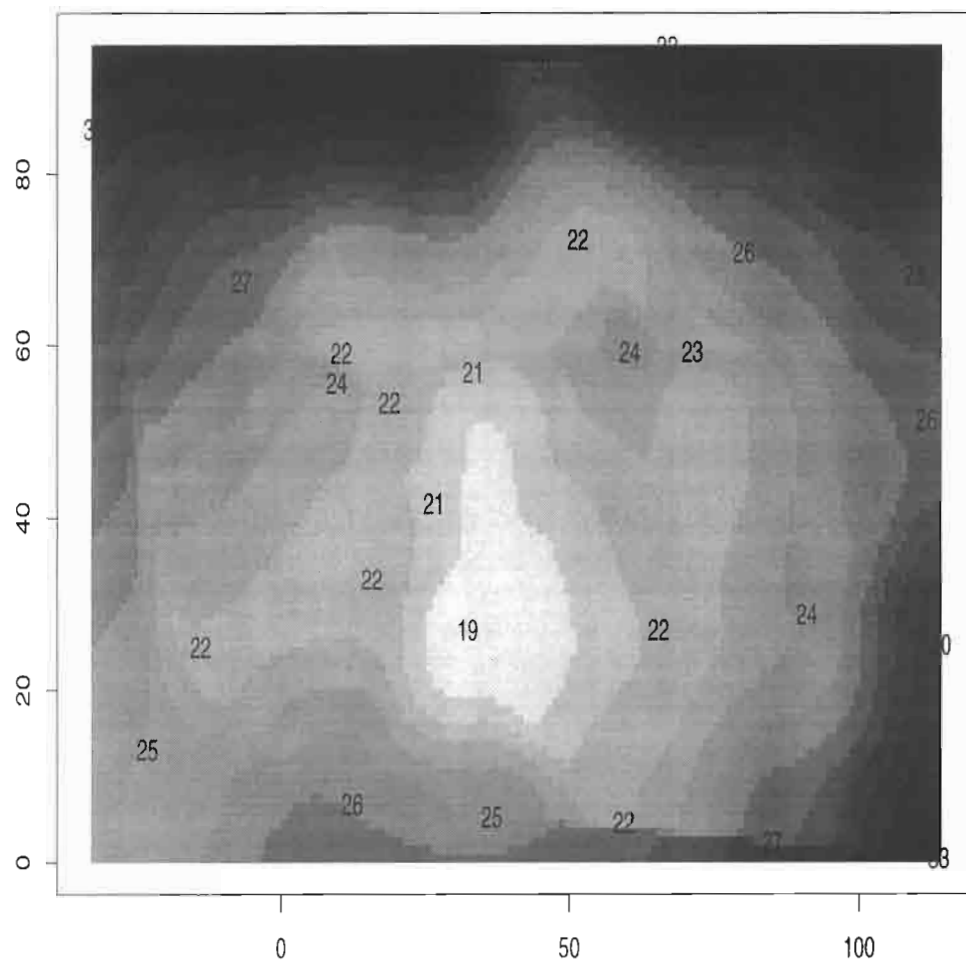


FIG. 3.2: Fréquence d'occurrence de valeurs nulles exprimée en % du nombre d'événements affectant la zone. Les lignes d'isovaleurs semblent centrées sur le réseau. Il s'agit d'un effet dû au mode de sélection des événements, les systèmes précipitants passé seulement au Nord ou au Sud du réseau n'étant pas inclus dans le jeu de données ayant servi à la constitution de cette carte.

### 3.1.4 Conclusion

Le principal intérêt de ce travail d'ajustement est de montrer que les problèmes de non stationnarité ne sont pas prédominants dans l'écart de la loi aux données. Concernant la confiance que l'on peut placer dans ce modèle, il convient d'être assez prudent, en particulier en ce qui concerne l'extrapolation à des valeurs non observées, comme le souligne Nicolas Bouleau [19]. En effet, rien ne nous permet a priori d'affirmer que les pluies suivent une loi gamma, et le fait de remplacer une loi empirique discrète par cette loi continue arbitraire peut conduire à des conclusions très éloignées de la réalité.

Ce problème très important a été abordé par Kedem et al. [62]. En partant de considérations physiques assez générales sur la formation des précipitations, ces auteurs parviennent à spécifier une famille de lois à laquelle devrait appartenir la loi de la pluie (en supposant qu'on soit fondé à parler de *la loi de la pluie*). Lorsque la loi est choisie à partir de considérations physiques, et seulement dans ce cas, il est légitime d'extrapoler les propriétés du modèle à des valeurs non observées. Ces auteurs ont mis en œuvre leur méthode sur des pluies tropicales de Mousson et de pré-Mousson et obtiennent une loi log-normale, ce qui est à la fois encourageant parce que c'est une loi couramment utilisée et décevant parce qu'on sait qu'elle est très approximative.

## 3.2 Fonction de covariance

### 3.2.1 Introduction

On a supposé qu'on pouvait représenter les champs de pluie par une fonction aléatoire de variance finie. On peut donc s'intéresser aux covariances expérimentales. L'intérêt de regarder la covariance et non pas le variogramme expérimental (qui n'est pas tout à fait équivalent) réside dans le fait que c'est la covariance qui sera spécifiée dans l'algorithme de simulation.

### 3.2.2 Calcul de la covariance empirique

Pour tous les couples  $(x_{j_1}, x_{j_2})$  de stations, on calcule sur les  $n$  événements :

$$\widehat{C}(x_{j_1}, x_{j_2}) = \frac{1}{n} \sum_{i=1}^n r_i(x_{j_1}) r_i(x_{j_2}) - \frac{1}{n^2} \sum_{i=1}^n r_i(x_{j_1}) \sum_{i=1}^n r_i(x_{j_2}) \quad (3.8)$$

Les valeurs  $\widehat{C}(x_{j_1}, x_{j_2})$  sont séparées en deux groupes selon l'orientation  $\theta_{j_{12}}$  du vecteur  $h_{j_{12}} = x_{j_1} - x_{j_2}$

$$H_\alpha = \{h_{j_{12}}/\theta_{j_{12}} \in [\alpha + \epsilon, \alpha - \epsilon]\}$$

et

$$H_{\alpha + \frac{\pi}{2}} = \{h_{j_{12}}/\theta_{j_{12}} \in [\alpha + \frac{\pi}{2} + \epsilon, \alpha + \frac{\pi}{2} - \epsilon]\}$$

Le demi-angle  $\epsilon$  au sommet du cône est pris égal à  $10^\circ$ .

On peut alors chercher quel est l'angle  $\alpha$  qui maximise visuellement la différence entre les deux nuages

$$\{(h_j, \widehat{C}(h_j)), h_j \in H_\alpha\}$$

et

$$\{(h_j, \widehat{C}(h_j)), h_j \in H_{\alpha + \frac{\pi}{2}}\}$$

Il apparait qu'un angle  $\alpha = 70^\circ$  (avec la convention géographique), maximise la différence entre ces deux nuages, ils sont représentés sur la figure 3.4.

Sans anticiper sur la suite de ce travail, on peut remarquer que cet angle  $\alpha = 70^\circ$  définit un axe d'anisotropie pointant dans une direction E-NE qui correspond à la direction moyenne préférentielle de déplacement des S.C.M sahéliens.

### 3.2.3 Analyse

#### 3.2.3.1 Stationnarité

Si on se donne une classe de distances, on voit sur la figure 3.3 que les valeurs des covariances expérimentales peuvent varier dans un rapport de un à quatre. On

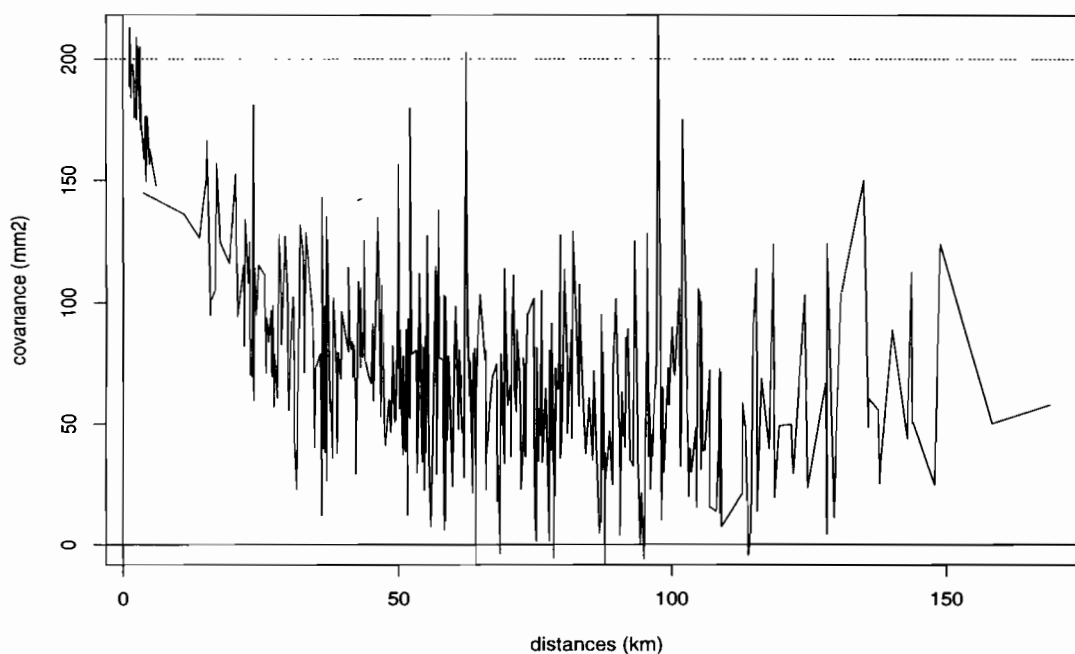


FIG. 3.3: Covariances expérimentales non directionnelles. Les covariances relatives aux faibles distances ( $\leq 2$  km) sont calculées à partir du réseau complet ayant fonctionné de 1990 à 1993. Les autres points sont calculés à partir du réseau de base sur la période 1990-1995. On observe un taux de décroissance rapide aux faibles distances ( $\leq 10$  km) qui tend à devenir nul aux grandes distances. Toutefois, les corrélations ne s'annulent pas.

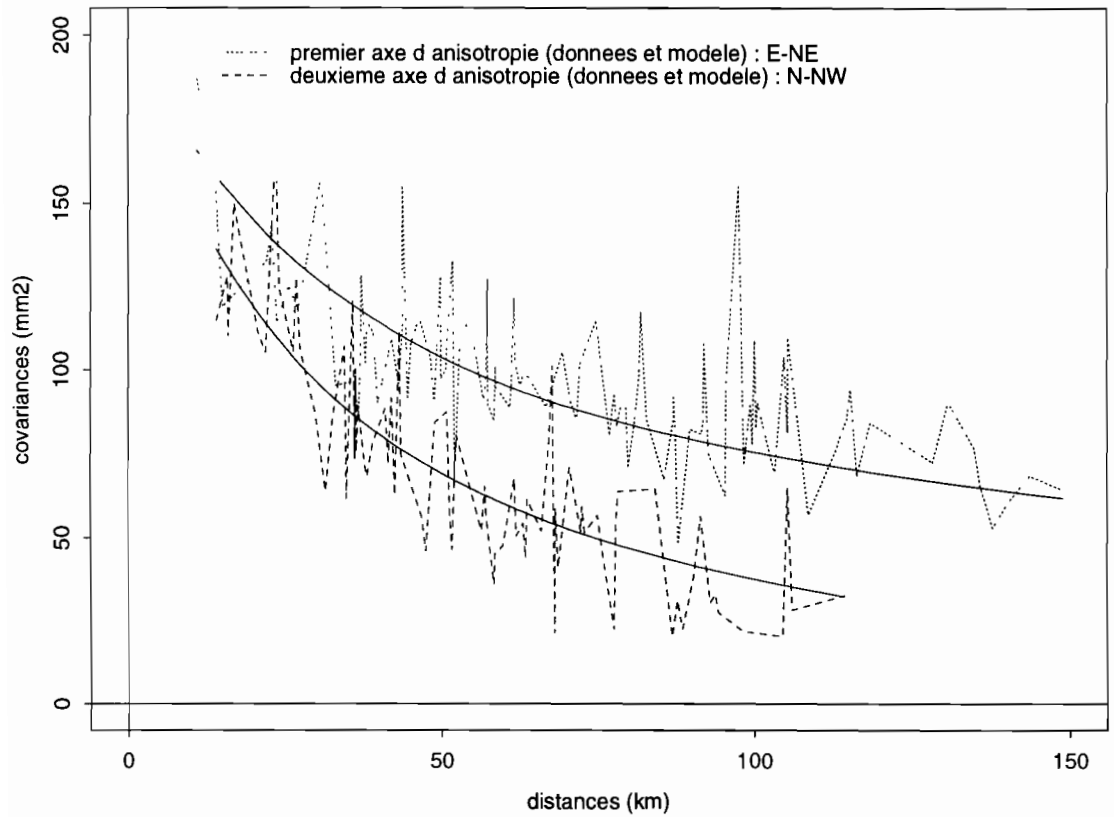


FIG. 3.4: Covariances expérimentales directionnelles,  $(h_j, \hat{C}(h_j))$ , selon un axe orienté à  $70^\circ$  (pointillés) et selon un axe orienté à  $160^\circ$  (tirets). Les corrélations décroissent beaucoup moins vite dans un axe orienté E-NE qui correspond à la direction moyenne de propagation des systèmes convectifs.

voit sur le graphique 3.4 que ce fait est partiellement imputable à l'anisotropie. Mais une certaine variabilité subsiste à l'intérieur de quelques classes de distances. On peut penser qu'une part de cette variabilité résiduelle est imputable au fait qu'on dispose d'un échantillon fini. Nous admettrons donc le caractère stationnaire de la covariance des champs étudiés même s'il est théoriquement possible (mais numériquement assez lourd) de tester et modéliser la non stationnarité des structures de covariance [95, 115].

### 3.2.3.2 Comportement à l'origine

Il apparaît sur la figure 3.3 que la covariance expérimentale aux faibles distances ( $\leq 2\text{km}$ ) est très proche de la variance stationnaire estimée.

On sait par ailleurs que les mesures effectuées par les pluviomètres sont entachées d'une certaine erreur non corrélée spatialement. Celle-ci contribue à la destructuration apparente des champs observés, c'est à dire à l'apparition d'une discontinuité de la covariance à l'origine. Vu la grande proximité entre la valeur de la covariance pour les faibles valeur de  $|h_{j12}|$  et la variance stationnaire estimée, il ne reste pas beaucoup de place pour une variance non imputable à la variabilité propre des champs de pluie.

Nous allons donc représenter les champs de pluie par une covariance continue à l'origine. Comme la décroissance de la covariance est assez rapide, on retiendra un modèle linéaire (non dérivable) à l'origine. Ceci signifie que les champs seront représentés par une fonction aléatoire continue en moyenne quadratique, mais non différentiable.

### 3.2.3.3 Structures emboîtées et comportement aux grandes distances

Si on essaye d'ajuster au mieux une covariance aux nuages empiriques, on ne parvient pas à rendre compte simultanément de la décroissance assez rapide aux faibles distances, et de la décroissance lente aux grandes distances (en particulier du fait que les corrélations ne sont pas nulles même au-delà de 100 km). Pour rendre compte de ce changement de régime dans la décroissance des corrélations, il est nécessaire de modéliser la covariance par une somme de covariances élémentaires. Ceci revient à supposer que les champs observés résultent de la superposition (additive) de deux phénomènes indépendants manifestant des échelles caractéristiques différentes [135].

En terme spectral, ceci revient à choisir une densité spectrale bi-modale: des hautes fréquences pour la structure de petite échelle, et des très basses fréquences qui correspondraient à une porteuse associée aux fluctuations de grande échelle.

### 3.2.3.4 Anisotropie

À partir du moment où l'on suppose l'existence de deux structures, l'anisotropie de chaque structure doit être analysée séparément [27]. Il apparaît que l'on rend

	variance ponctuelle	paramètre d'échelle (portée)	rapport d'anisotropie
structure n° 1	$\sigma_1^2 = 100 \text{ mm}^2$	$s_1 = 30 \text{ km}$	$a_1 = 2/3$
structure n° 2	$\sigma_2^2 = 100 \text{ mm}^2$	$s_2 = 300 \text{ km}$	$a_2 = 1/3$

TAB. 3.1: Paramètres de la fonction de covariance théorique retenue

mieux compte du nuage observé en choisissant une covariance de grande échelle fortement anisotrope, et une structure de petite échelle à l'anisotropie moins marquée. Cette observation admet une interprétation, mais qui doit être évidemment considérée avec prudence : la covariance de petite échelle traduirait la présence de cellules convectives aux mouvements assez désordonnés à l'intérieur d'une bande frontale, la covariance de grande échelle correspondrait à la dynamique de méso-échelle des systèmes précipitants, plus régulière.

### 3.2.4 Synthèse

#### 3.2.4.1 Modèle retenu

À partir des considérations précédentes, il y a évidemment de nombreux modèles possibles.

En voici un :

$$C(h) = \sigma_1^2 e^{-|h|_1/s_1} + \sigma_2^2 e^{-|h|_2/s_2} \quad (3.9)$$

où

$$|h|_i = \sqrt{h_E^2 + \left(\frac{h_N}{a_i}\right)^2} \quad (3.10)$$

$(h_E, h_N)$  étant les coordonnées de  $h$  sur les axes d'anisotropie ( $h_E$  le long de l'axe principal E-N-E et  $h_N$  le long de l'axe secondaire N-W-N) et  $a_i$  le rapport d'anisotropie de la structure  $i$ . Les valeurs des paramètres sont données dans le tableau 3.1.

Le modèle est ergodique au sens où les fluctuations d'une moyenne d'espace de la forme  $\int_D \frac{1}{|D|} R(x) dx$  tendent vers 0 lorsque  $D$  tend vers  $+\infty$ .

Autrement dit,

$$\text{Var} [\bar{R}(D)] = \text{Var} \left[ \int_D \frac{1}{|D|} R(x) dx \right] \rightarrow 0, \text{ lorsque } D \rightarrow \infty \quad (3.11)$$

Ceci peut se voir en remarquant que

$$\frac{1}{|D|^2} \int_D \int_D Cov[R(x), R(y)] dx dy \rightarrow 0, \text{ lorsque } D \rightarrow \infty \quad (3.12)$$

Une propriété faisant intervenir des points à l'infini n'est pas à prendre à la lettre puisque notre modèle est une représentation des systèmes précipitants valable seulement localement. Mais elle permet d'introduire la portée intégrale [65]. Celle-ci est définie par

$$A = \lim_{D \rightarrow \infty} |D| \frac{Var[\bar{R}(D)]}{\sigma_R^2} \quad (3.13)$$

A s'interprète comme l'échelle du phénomène représenté par  $R$ . En particulier, si  $D$  est un domaine du plan que l'on peut recouvrir par  $n$  domaines disjoints d'aires  $A$ , alors

$$Var[\bar{R}(D)] \approx \frac{\sigma_R^2}{n} \quad (3.14)$$

qui est la variance de la somme de  $n$  v.a indépendantes de variance  $\sigma_R^2$ .

Dans le cas d'une covariance exponentielle isotrope sur  $\mathbb{R}^2$  (i.e de la forme  $h \rightarrow e^{-h/s}$ ), la portée intégrale est  $A = 2\pi s^2$ , soit ici une valeur de l'ordre de 500 000 km<sup>2</sup>.

Dans le cas présent, le sens d'une telle valeur est physiquement discutable, mais ceci montre qu'il est incorrect d'utiliser des résultats supposant l'absence de corrélation spatiale, même aux grandes distances.

### 3.2.4.2 Effet de l'anamorphose

Cette analyse a porté sur les données brutes, mais on a vu à la section 1.5.4 que pour obtenir une covariance compatible avec une loi marginale donnée, il était nécessaire de transformer la covariance de départ ajustée aux données.

Il se trouve que pour la loi marginale considérée ici, la transformation  $\zeta$  définie par l'équation 1.4 à la section 1.5.4 est très proche de l'identité. Plus exactement, l'ordre de grandeur de la transformation induite par  $\zeta$  est bien inférieur à la précision avec laquelle on peut ajuster la covariance (ce point est étudié dans [46]).

Par conséquent, on peut considérer que les champs étudiés peuvent être obtenus par anamorphose de fonctions aléatoires gaussiennes admettant la covariance composée de deux modèles exponentiels anisotropes, définies à la section 3.2.4.

La covariance résultante n'est plus une somme d'exponentielles, mais elle en est très proche. En particulier, comme l'anamorphose est continue, il s'agit toujours d'une covariance continue mais non dérivable à l'origine.



## Chapitre 4

# Propriétés du modèle et comparaison aux données

### Présentation

On étudie maintenant le modèle en le confrontant aux données EPSAT-Niger du point de vue de certaines propriétés n'ayant pas été incluses explicitement dans les équations mais que l'on souhaiterait retrouver. Ces propriétés sont : (i) l'aptitude à reproduire l'hétérogénéité du jeu de données telle qu'elle apparaît à partir de la classification, (ii) les distributions à différentes échelles d'agrégation spatiale, (iii) la morphologie des surfaces où la lame d'eau dépasse un certain seuil (surfaces sup-seuil).

### 4.1 Procédure de mise en œuvre numérique et exemples graphiques

La plupart des propriétés du modèle que l'on voudrait étudier, comme les distributions à différentes échelles par exemple, sont inaccessibles analytiquement. Nous allons donc les étudier expérimentalement. À cette fin, un premier jeu de 400 champs a été simulé sur une grille régulière de  $100 \text{ km} \times 100 \text{ km}$  contenant 10000 points. Afin d'étudier les fluctuations d'échantillonnage, on a simulé un deuxième jeu de 1200 champs sur une grille de  $100 \text{ km} \times 100 \text{ km}$  contenant 496 points : une grille régulière assez lâche avec une grille plus dense emboîtée au centre (figure 4.1).

Deux exemples de simulations (réalisés sur une troisième grille régulière de 10000 points) sont présentés (figures 4.2 et 4.3). Le premier présente une forte intermittence. Les cumuls ne dépassent pas 30 mm. Le deuxième est représentatif des lignes de grain, il ne présente pas de valeurs nulles et les cumuls sont très élevés (jusqu'à 100 mm).

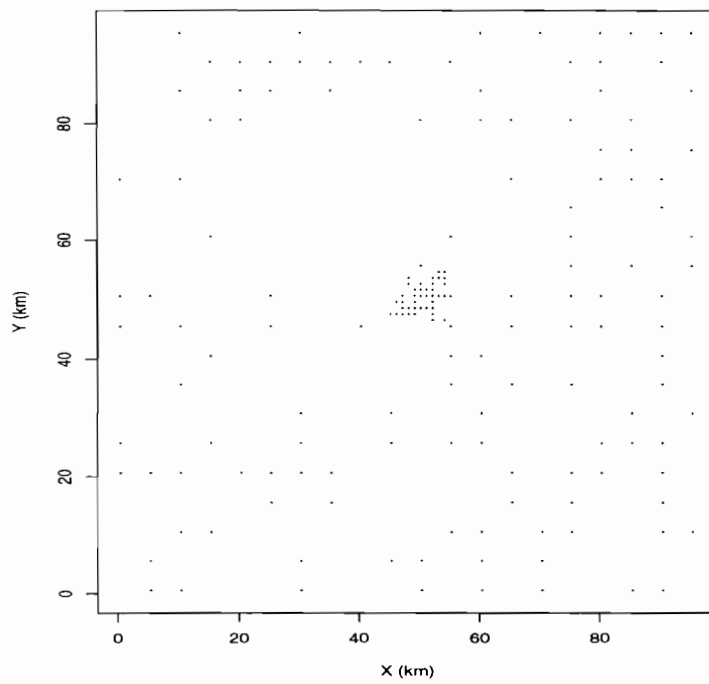


FIG. 4.1: Grille irrégulière de 496 points utilisée pour l'étude expérimentale du modèle. La maille de la grande grille est de  $5 \text{ km} \times 5 \text{ km}$ , celle de la petite grille au centre est de  $1 \text{ km} \times 1 \text{ km}$

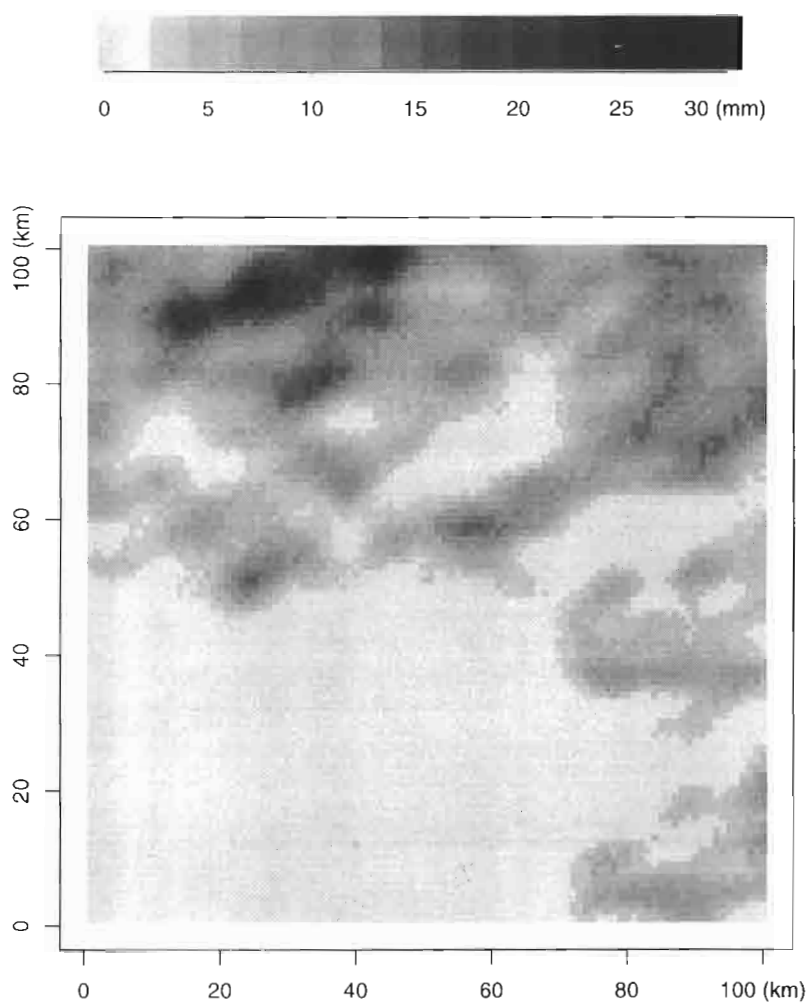


FIG. 4.2: *SCM simulé présentant une forte intermittence. Les zones les plus claires (principalement en bas à gauche), correspondent à des zones non affectées par la pluie.*

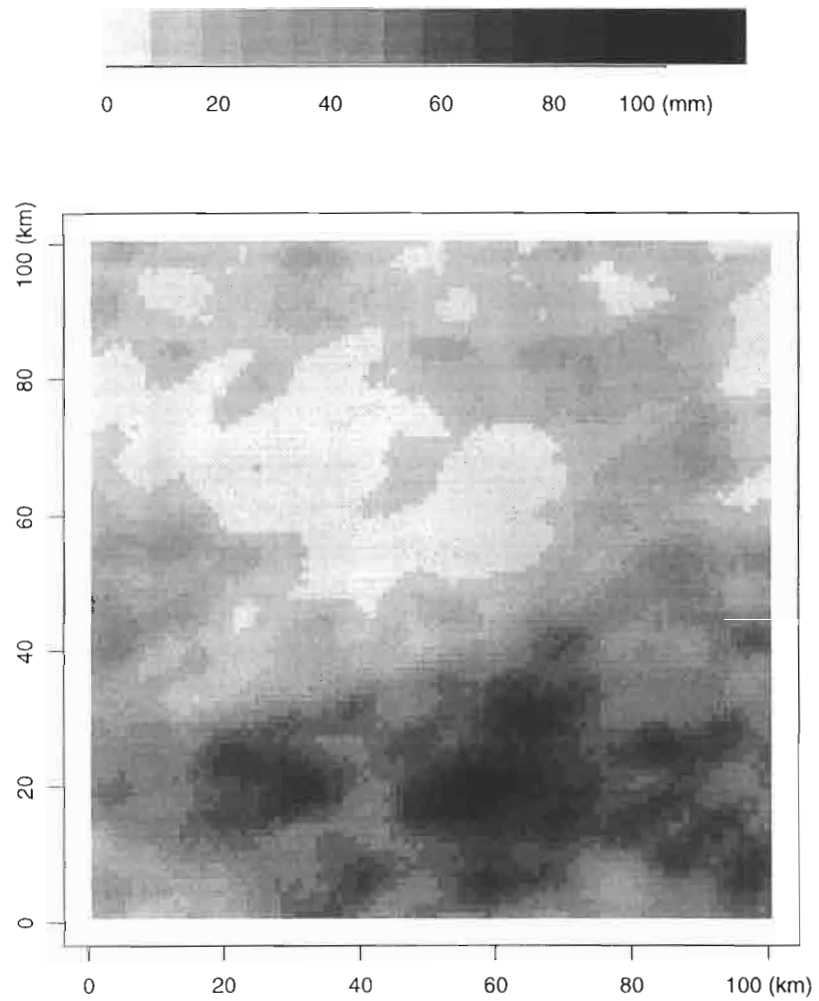


FIG. 4.3: *SCM* simulé présentant une faible intermittence, (la palette de gris est différente de la figure précédente). Tout le domaine a été affecté par la pluie.

## 4.2 Analyse par groupe

### 4.2.1 Position du problème

Les événements pluvieux survenant au cours de la saison des pluies présentent des caractéristiques très diverses en terme d'extension spatiale, d'organisation, de cinématique, d'intensité ou encore d'intermittence. Cette diversité peut résulter des différences de maturité entre les différents systèmes précipitants affectant la zone d'étude, ou des différences des conditions leur donnant naissance et des phénomènes atmosphériques mis en jeu.

Dans une perspective d'amélioration des méthodes géostatistiques d'interpolation des champs de pluie, Amani a testé certains critères de classification des événements pluvieux basés sur des données pluviographiques [7, 9]. Son travail montre que la distinction des événements selon la fraction du réseau non affectée par la pluie (intermittence) au cours de l'événement constituait un bon critère.

Suite aux travaux de Laurent [67, 68] montrant une bonne adéquation générale entre une classification basée sur des observations atmosphériques et une classification basée sur les données sol d'EPSAT-Niger, D'Amato dans une étude de la climatologie sahélienne [28] a été amenée à distinguer trois groupes d'événements pluvieux :

**groupe 1** : intermittence  $\leq 5\%$

**groupe 2** :  $5\% < \text{intermittence} \leq 25\%$

**groupe 3** :  $25\% < \text{intermittence} \leq 70\%$

Par rapport à la terminologie utilisée par les atmosphériciens [30, 78], on a la correspondance suivante :

**Lignes de grain**  $\longleftrightarrow$  groupe 1

**Masses Convectives Complexes (MCC)**  $\longleftrightarrow$  groupe 1  $\cup$  groupe 2

**Systèmes Convectifs de Mésos-échelle**  $\longleftrightarrow$  groupe 1  $\cup$  groupe 2  $\cup$  groupe 3

Il est alors naturel de penser qu'à l'hétérogénéité physique correspond une hétérogénéité statistique, et cette correspondance justifierait l'étude statistique des événements pluvieux si ces paramètres permettaient de caractériser (partiellement) les différents groupes.

Étant donné cette classification en trois groupes, deux mises en œuvre du modèle sont possibles.

- La première consiste à réaliser une analyse statistique sur chaque groupe et à modéliser pour chaque groupe une loi et une covariance  $(F^{(i)}, C^{(i)})$  pour  $i = 1, 2, 3$ . Ceci permet de simuler des événements à partir du modèle gaussien anamorphosé avec des paramètres différents selon chaque groupe.

- Le deuxième mode consiste à simuler un ensemble d'événements conformément aux paramètres estimés sur l'ensemble des événements observés.

Avec la première méthode, on peut espérer qu'en délimitant plus précisément l'objet d'étude, on peut le décrire plus correctement. En contre-partie, on perd en généralité et on alourdit la description puisqu'il faut trois fois plus de paramètres. Un autre inconvénient est que le critère de classification est mal adapté à la structure du modèle. En effet, les événements du groupe 1 par définition affectent plus de 95% de la fenêtre. Or en simulant une fonction aléatoire gaussienne anamorphosée on n'est jamais certain de satisfaire ce critère<sup>1</sup>.

Par conséquent, pour avoir un algorithme cohérent avec la définition des groupes, nous avons choisi le deuxième mode. On peut par ailleurs remarquer que le modèle ne peut pas être parfait selon les deux modes à la fois car le modèle gaussien anamorphosé n'est pas stable par mélange. Nous verrons par la suite qu'il est évidemment imparfait selon les deux modes. Ayant fait le choix d'analyser et de simuler les événements sans distinction entre les groupes, il convient en premier lieu de vérifier que le modèle reproduit l'hétérogénéité statistique de la population modélisée.

#### 4.2.2 Effectifs et statistiques simples

Si on scinde l'ensemble des événements simulés en trois groupes selon la fraction du domaine simulé non couvert par la pluie on voit qu'on obtient sensiblement les mêmes proportions que sur les données (figure 4.1). À l'intérieur de chaque groupe, la fraction moyenne du domaine non couvert par la pluie est identique sur les données et sur les simulations. La moyenne ponctuelle est également sensiblement identique entre les données et les simulations. La variance ponctuelle est bien reproduite par le modèle pour les simulations des groupes 1 et 2 et la réunion de ces deux groupes. En revanche les valeurs ponctuelles des événements du groupe 3 simulés sont deux fois moins dispersées que leurs analogues réelles.

	proportion (%)	$m$ (mm)	$\sigma^2$ (mm <sup>2</sup> )	$F_0$ (%)
groupe 1	30 / 25	20,8 / 22	268 / 275	1,4 / 0,8
groupe 2	23 / 25	10,9 / 10	130 / 147	11,3 / 12
groupe 1 $\cup$ groupe 2	53 / 51	16,4 / 16,3	231 / 250	5,8 / 6,5
groupe 3	47 / 49	3,3 / 4,5	35,8 / 78,5	48,3 / 46
ensemble	100 / 100	10,8 / 10,8	190 / 200	26 / 26

TAB. 4.1: *Statistiques de base (simulation/données).*

1. Dès que la loi monovariante admet un atome à l'origine, on a même une probabilité 1 qu'une suite infinie de réalisations indépendantes sur un domaine fini donné contienne au moins une réalisation uniformément nulle sur ce domaine.

### 4.2.3 Distribution monovariante

Compte tenu du bon résultat observé pour la distribution ponctuelle des valeurs du groupe 1U2 à l'ordre deux, il est intéressant d'examiner si la fonction de répartition dans son ensemble est bien reproduite pour ce groupe.

Les fonctions de répartition empiriques ont été calculées en neuf stations régulièrement espacées sur le domaine d'étude. Il apparaît qu'elles sont bien reproduites par le modèle, une seule station étant rejetée au seuil de 10%.

En particulier si on ajuste une loi  $\Gamma \vee \delta_0$  au groupe 1U2 réel (figures 4.5 et 4.6), le résultat en terme de proportion de stations acceptées au seuil de 10% par un test de Kolmogorov-Smirnov est moins bon que celui obtenu ici (figure 4.4).

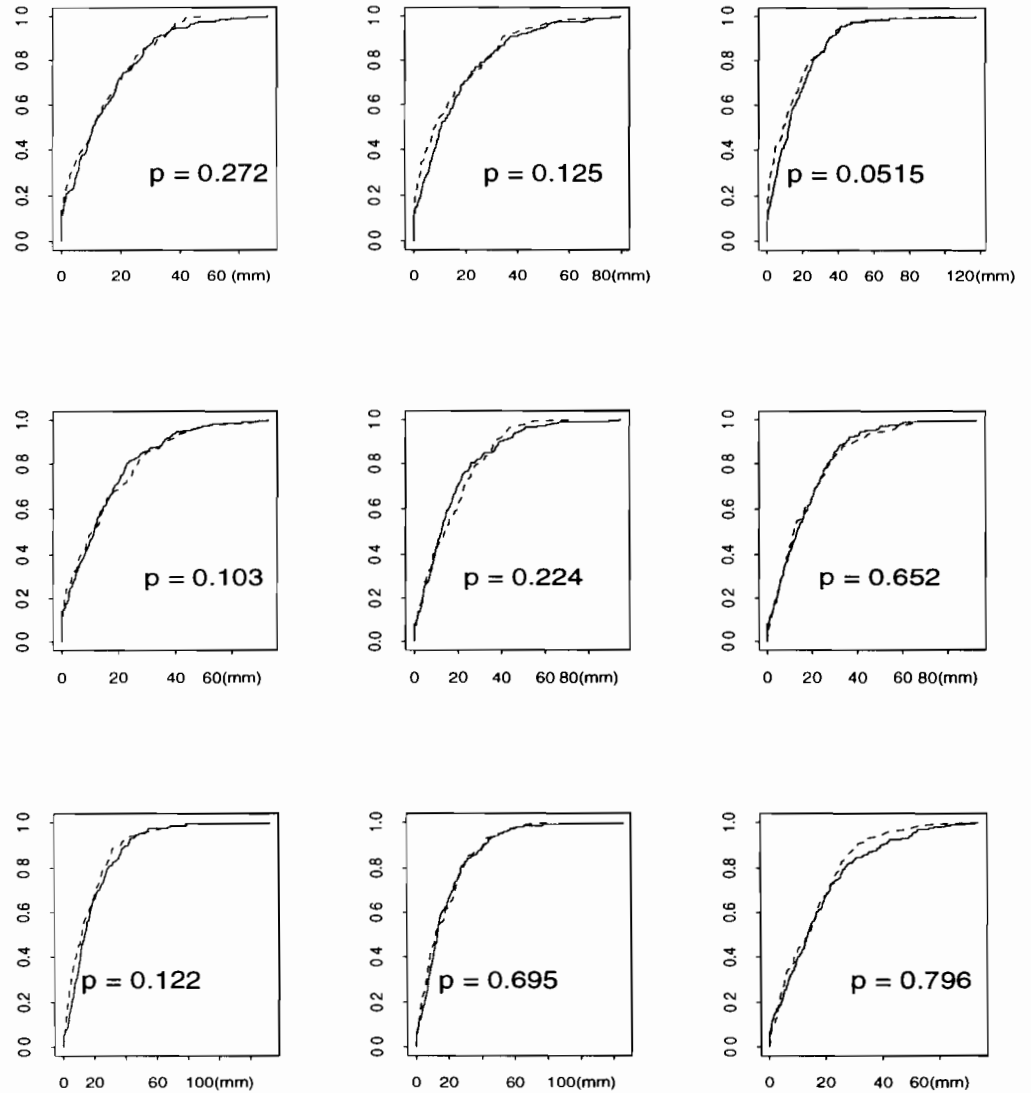


FIG. 4.4: Fonction de répartition en 9 stations pour les valeurs appartenant aux événements du groupe  $1 \cup 2$  (trait continu = données / tiret = simulation), et la  $p$ -valeur associée au test de Kolmogorov-Smirnov

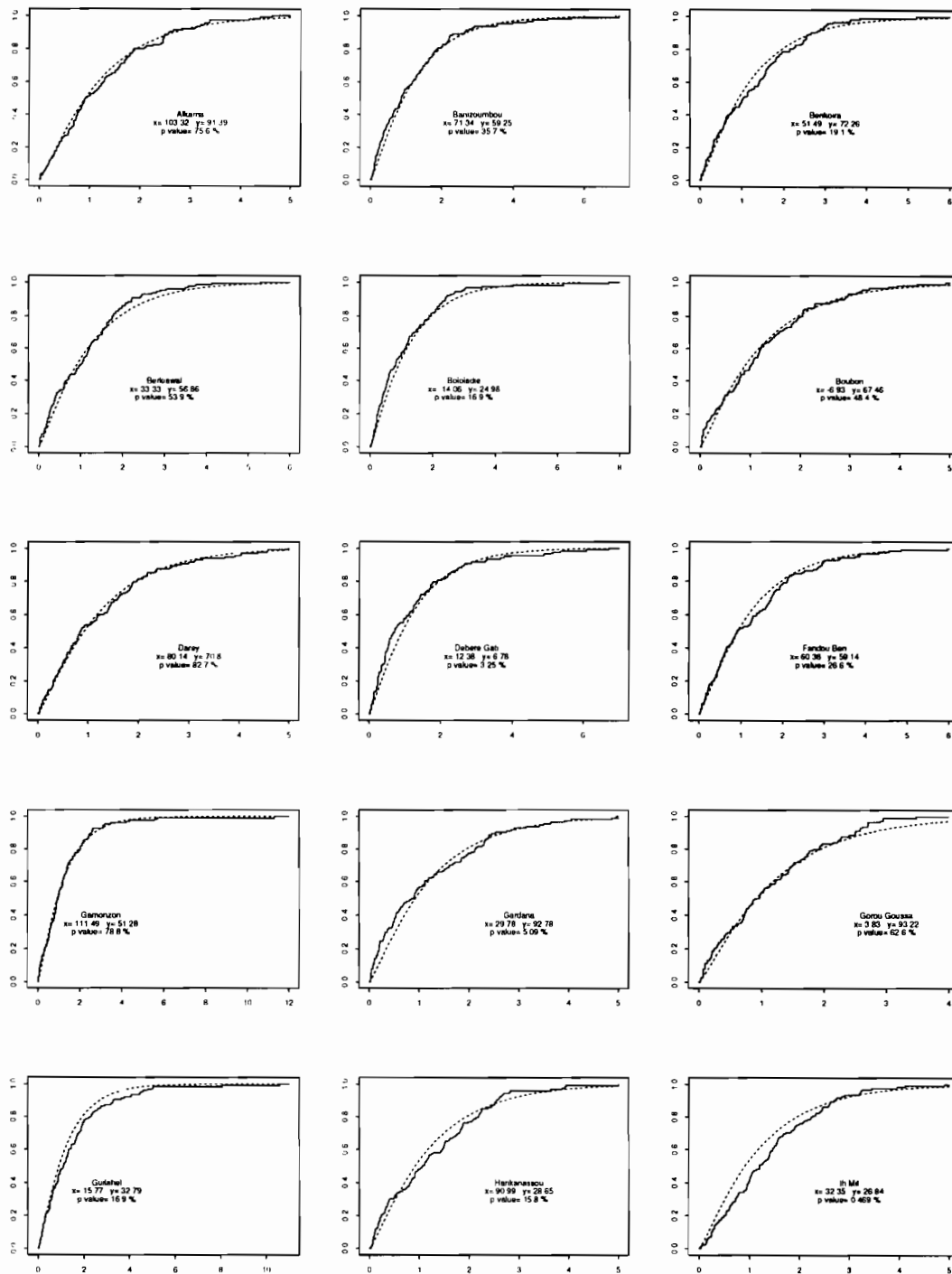


FIG. 4.5: Fonction de répartition pour les valeurs non nulles appartenant aux événements du groupe  $1 \cup 2$  (trait continu = données / tiret = loi gamma), et la p-valeur associée au test de Kolmogorov-Smirnov

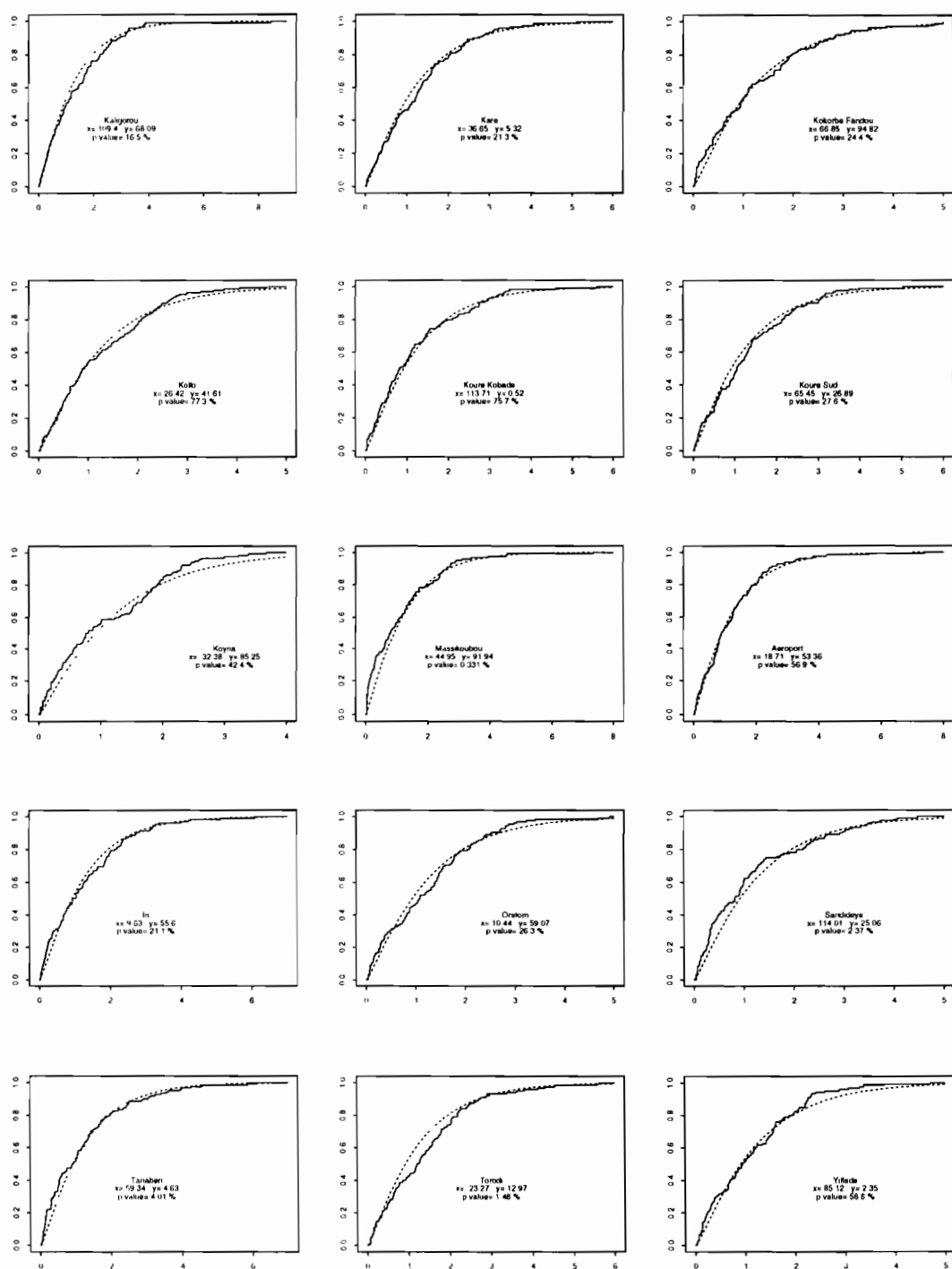


FIG. 4.6: Fonction de répartition pour les valeurs non nulles appartenant aux événements du groupe  $1 \cup 2$  (trait continu = données / tiret = loi gamma), et la p-valeur associée au test de Kolmogorov-Smirnov

#### 4.2.4 Fonction de structure

On examine maintenant la fonction de structure spatiale des champs par groupe. Partant d'un ensemble d'événements simulés, possédant le même variogramme que l'ensemble des événements observés (voir figure 4.7), on calcule les variogrammes pour chaque sous-groupe (simulés et observés).

**Groupe 3** Le variogramme n'est pas bien reproduit. Non seulement la variance n'est pas correcte mais le variogramme des champs simulés dépasse cette variance à 70 km (effet de trou).

**Groupe 1U2** L'agrément est bien meilleur. Néanmoins, alors que la variance est légèrement sous-évaluée par le modèle, le variogramme simulé est partout au-dessus du variogramme réel, si bien que la décorrélation est atteinte sur les simulations à 100 km alors qu'une corrélation subsiste pour les données.

**Groupe 1** On observe un bon accord global. On retrouve pour le groupe 1 ce phénomène de corrélation à très longue distance pour les champs réels qui est mal reproduit sur les champs simulés. (On pouvait s'attendre à observer une décorrélation rapide sur les champs simulés, étant donné que la sélection de champs sur la base de l'intermittence revient à sélectionner des champs gaussiens à valeur élevée qui sont réputés destructurés [90]).

**Groupe 2** Ici aussi l'adéquation est bonne, on peut noter que la décroissance de la covariance est plus rapide (la variance des incréments croît plus vite) pour le groupe 2 que pour l'ensemble de l'échantillon, aussi bien pour les champs réels que pour les simulations.

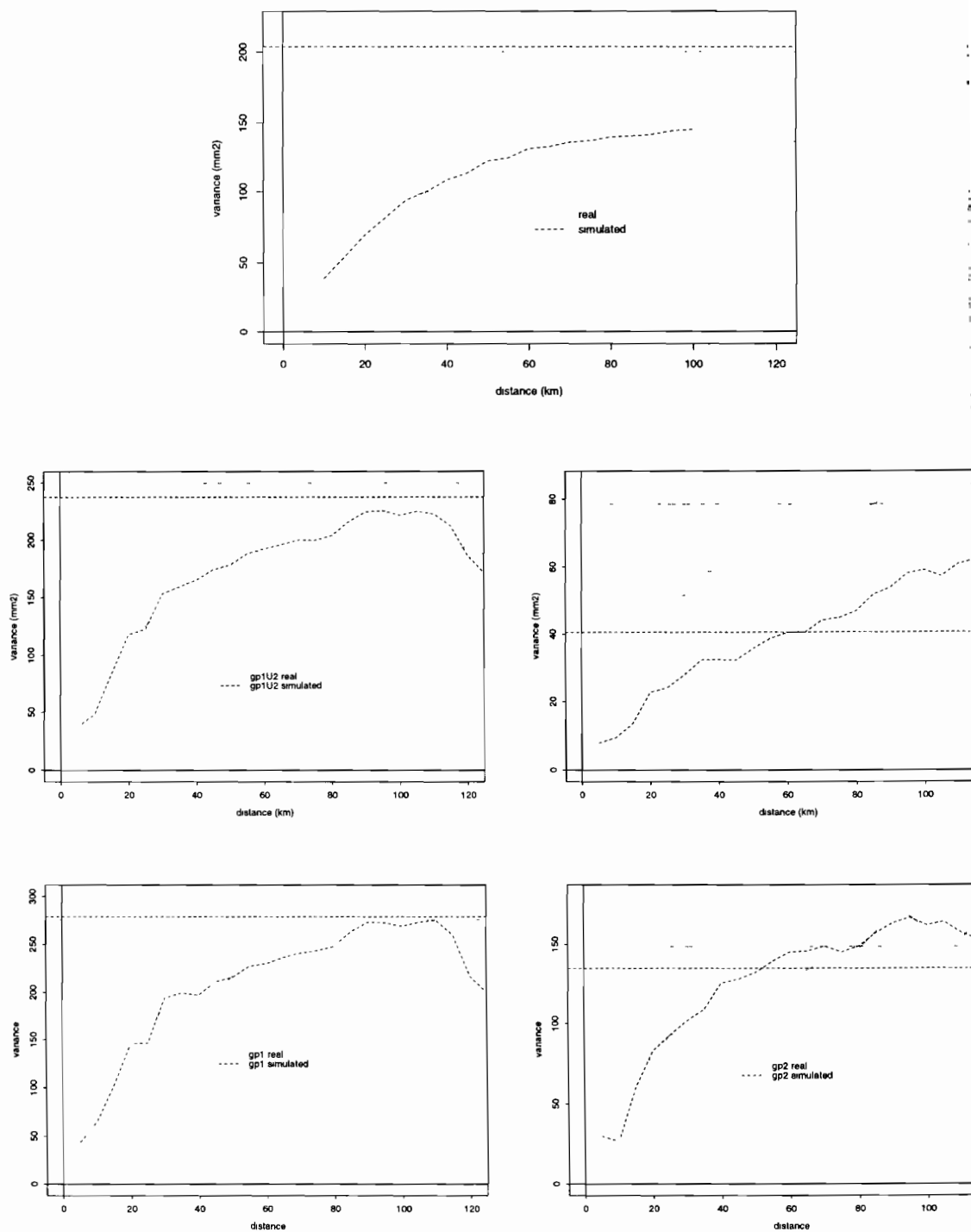


FIG. 4.7: Variogrammes par groupe. De haut en bas et de gauche à droite : ensemble, groupe 1U2, groupe 3, groupe 1, groupe 2 (tirets = simulations, pointillés = données). On notera la bonne adéquation des courbes expérimentales aux courbes théoriques, excepté pour le groupe 3, y compris pour la variance (droite horizontale).

## 4.3 Changement d'échelle

### 4.3.1 Position du problème

Le présent modèle a été spécifié à partir de valeurs ponctuelles, or des données à cette échelle sont de peu d'intérêt pour l'hydrologie. En effet les processus (et les modèles) de surface font intervenir des mécanismes à des échelles allant d'un hectare à quelques  $\text{km}^2$  (en particulier en ce qui concerne l'intermittence [61]).

Par ailleurs les modèles d'infiltration ou d'échange sol-plante-atmosphère qui sont principalement des modèles décrivant les phénomènes en un point de l'espace (modèles ponctuels) sont précisément limités par le manque de données pluviométriques en plusieurs points. L'examen des propriétés multivariées qui va suivre peut constituer une validation (ou une invalidation) pour cet usage.

Enfin, pour l'utilisation du modèle en mode de désagrégation, on conditionne les simulations par une moyenne d'espace de grande échelle. Il est donc intéressant de vérifier si les distributions des sorties du modèle sont identiques à celles des données à grande échelle.

Nous allons donc étudier les propriétés de quantités de la forme

$$\bar{R}(\lambda) = \int_{D(\lambda)} \frac{1}{|D|} R(x) dx$$

où  $D(\lambda)$  est un carré de côté  $\lambda$

### 4.3.2 Propriétés à l'ordre deux

Les propriétés statistiques à l'ordre deux des moyennes spatiales sont entièrement déterminées par les propriétés à l'ordre deux des valeurs ponctuelles. Ceci provient de ce que

$$\begin{aligned} C_h(\lambda) &= \text{Cov} \left[ \int_{D(\lambda)} R(x) dx, \int_{D_h(\lambda)} R(x) dx \right] \\ &= \int_{D(\lambda)} \int_{D_h(\lambda)} C_R(x, y) dx dy \\ \text{avec } D_h(\lambda) &= \{x + h, \quad x \in D(\lambda)\} \end{aligned} \tag{4.1}$$

Regarder les propriétés d'ordre deux des moyennes spatiales revient à tester les hypothèses d'ordre deux du modèle (stationnarité et modèle particulier de covariance).

À partir du réseau de pluviographes, on peut difficilement évaluer des grandeurs telles que  $C_h(\lambda)$  parce qu'on n'a pas suffisamment de points pour calculer la moyenne de  $R$  sur des domaines  $D_h(\lambda)$  suffisamment éloignés de  $D(\lambda)$  et pas trop petits.

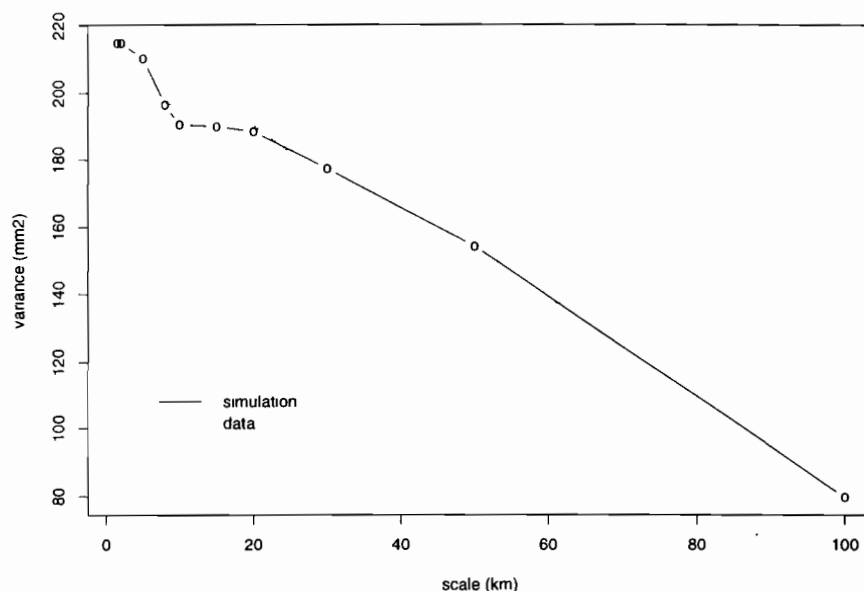


FIG. 4.8: Variance à différentes échelles spatiales. On notera la même décroissance linéaire de la variance en fonction de la largeur du carré.

On peut seulement calculer des variances à différentes échelles. On observe alors (figure 4.8), la même décroissance approximativement linéaire pour des domaines allant de  $2\text{km} \times 2\text{km}$  à  $100\text{km} \times 100\text{km}$ , avec une sous-estimation du modèle à grande échelle (de l'ordre de 30%) qui est peut être une manifestation de la non stationnarité.

Il convient de noter qu'un tel résultat ne peut être obtenu sans prendre en compte l'anisotropie des systèmes précipitants à travers la fonction de covariance.

Ceci étant, il est difficile de conclure sûrement sur ce point car, comme le montre la figure 4.9, il apparaît que cette courbe de variances à différentes échelles est sujet à des fluctuations d'échantillonnage assez importantes.

### 4.3.3 Intermittence et fonctions de répartitions

#### 4.3.3.1 Intermittence

Nous avons vu au paragraphe 4.2.2 que l'intermittence à l'échelle ponctuelle était correctement reproduite par le modèle (on retrouve bien la proportion de valeurs nulles prescrites via la fonction de répartition). Mais la connaissance de  $Pr[R(x) = 0]$  peut donner lieu à des comportements très divers à d'autres échelles selon la loi spatiale de la fonction aléatoire (qui, rappelons le, a été choisie pour des raisons de commodité numérique et conceptuelle, mais n'a pas fait l'objet de

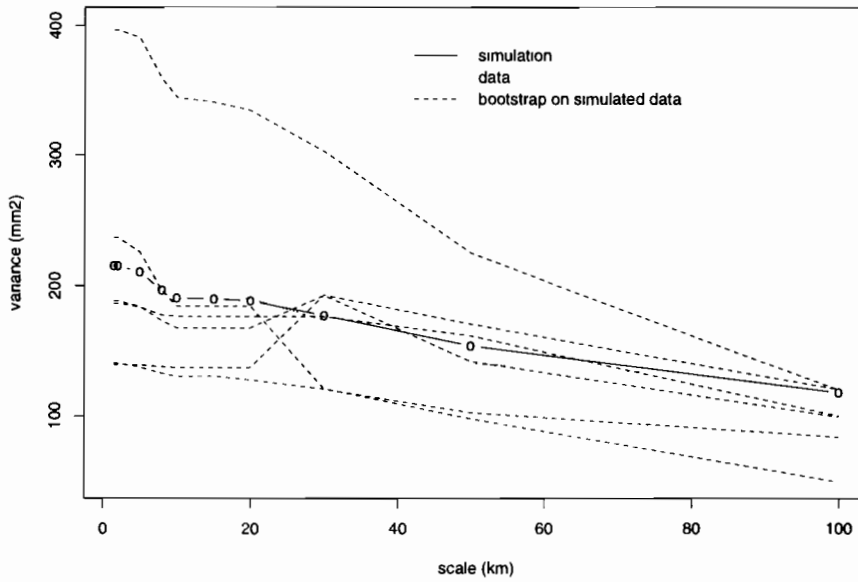


FIG. 4.9: Variance à différentes échelles spatiales calculées sur des sous-échantillons de 200 événements.

tests par rapport aux données).

Le calcul par simulation montre que le modèle reproduit bien la fréquence d'occurrence de valeurs nulles de l'échelle ponctuelle à environ 400 km<sup>2</sup>. Au delà, le modèle surestime cette fréquence (figure 4.10).

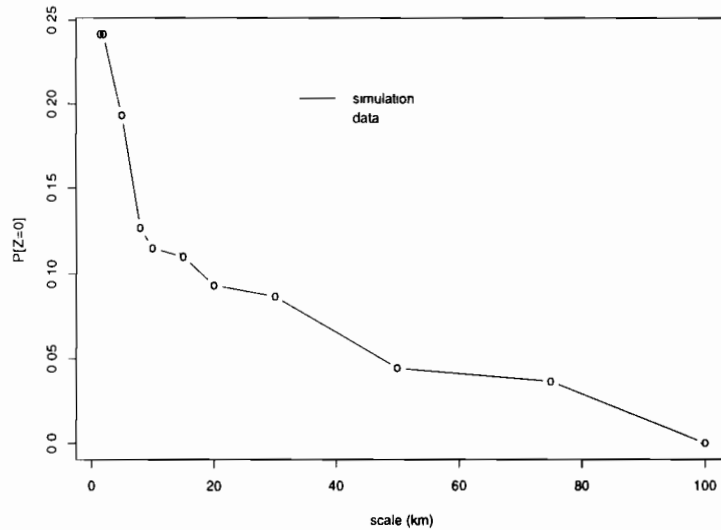


FIG. 4.10: Intermittence à différentes échelles. En abscisse : échelle. En ordonnées : Fréquence de valeurs nulles (trait continu = simulation, pointillés donnée).

#### 4.3.3.2 Fonction de répartition

La forme générale de la distribution est reproduite correctement au sens où les ordres de grandeurs des quantiles associés aux différentes fréquences empiriques observées sont respectés. Toutefois, l'hypothèse d'égalité des distributions est rejetée pour les quatre échelles présentées figure 4.11, même au seuil de 5%. On observe également une dégradation progressive de la distribution à travers les échelles.

Par ailleurs, on observe la présence d'un événement simulé produisant une valeur remarquablement élevée à toutes les échelles (en particulier plus de 100 mm sur 2500 km<sup>2</sup>). Cette valeur ne possède pas d'équivalent dans le jeu de données sur la période 90-95 et pourrait constituer une faiblesse du modèle. Néanmoins cette affirmation est à nuancer dans la mesure où un événement très fort (provoquant des dégâts de très grande ampleur) a été réellement observé autour de Niamey en 1989 [131] et plus récemment en 1998.

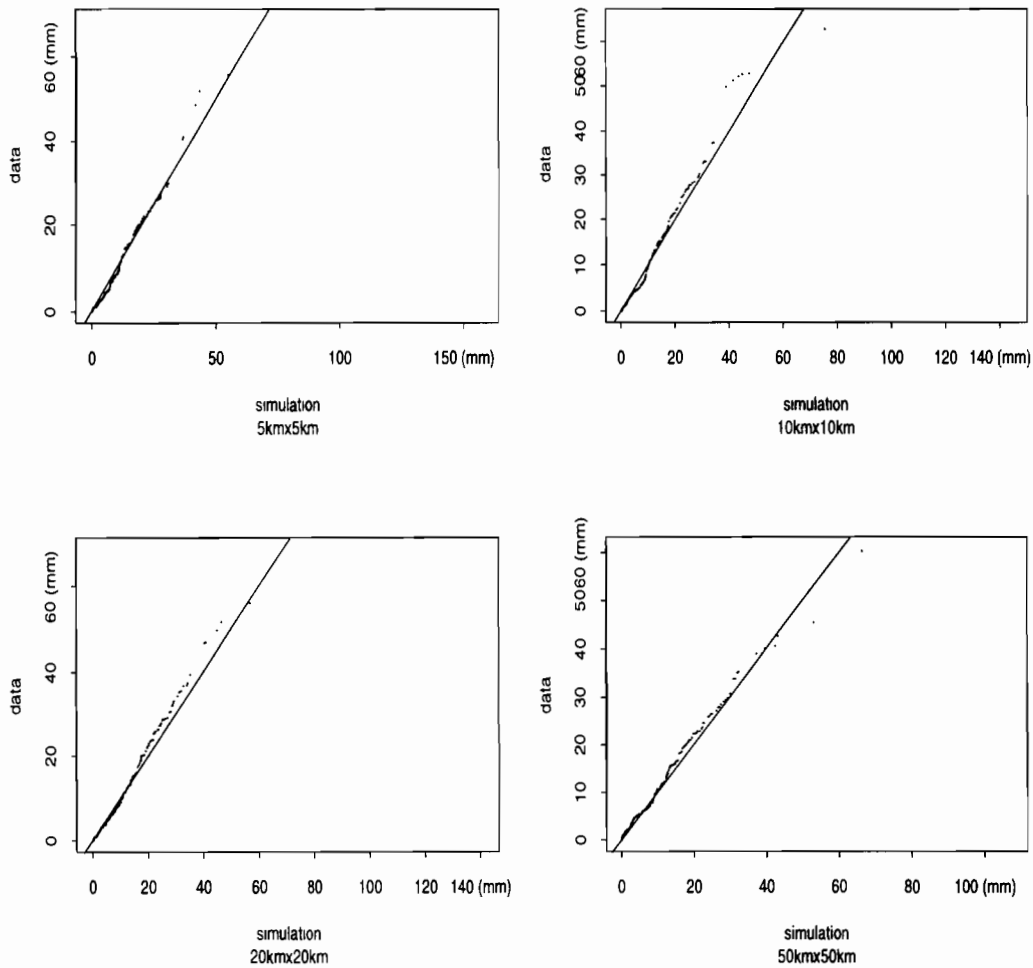


FIG. 4.11: *Graphiques quantile-quantile de  $R_D(\lambda)$  pour  $\lambda = 5, 10, 20, 50km$ . (Simulations en abscisse, données en ordonnées. L'hypothèse d'égalité des deux distributions est rejetée par un test de Kolmogorov-Smirnov aux quatre supports d'intégration à 10%. Néanmoins, on observe une bonne adéquation dans les quantiles faibles et médians, et une relative concordance des ordres de grandeur dans les valeurs extrêmes.*

## 4.4 Relation entre aire des surfaces sup-seuil et moyenne spatiale

### 4.4.1 Définition

Dans la littérature consacrée à l'estimation des précipitations par satellite, de nombreux auteurs mentionnent une relation croissante entre l'aire des surfaces où la pluie dépasse un certain seuil (en intensité ou en volume) et la moyenne spatiale sur le domaine d'étude. L'observation et la modélisation de cette relation permet d'estimer un cumul pluviométrique sur un domaine à partir de la seule donnée de l'aire de la surface pluvieuse, celle-ci pouvant être estimée à partir d'images satellitales.

L'objet de cette section est de voir si une telle relation est observée sur les données EPSAT et si c'est le cas, d'examiner comment elle est reproduite par le modèle.

On notera  $\bar{I}_\tau$  l'aire du domaine défini par  $\{x \in D / R(x) > \tau\}$  rapportée à l'aire de  $D$  (où  $D$  est le domaine d'étude), et  $\bar{R}$  la lame d'eau moyenne sur ce domaine.

### 4.4.2 Résultats et analyse

Les nuages  $(\bar{I}_\tau, \bar{R})$  sont présentés figure 4.12 pour différents seuils.

On observe, d'une part, sur les données une relation croissante très marquée entre les deux variables, et d'autre part que le modèle reproduit cette liaison de manière très fidèle. On peut quantifier cette liaison en calculant le coefficient de corrélation linéaire correspondant à chaque nuage :

	données	simulation
$\tau = 0$ mm	0.51	0.53
$\tau = 5$ mm	0.79	0.78
$\tau = 10$ mm	0.86	0.85
$\tau = 20$ mm	0.91	0.95
$\tau = 50$ mm	0.51	0.68

Les valeurs du modèle sont très proches de celles calculées sur les données, en particulier aux seuils  $\tau$  pour lesquels  $\rho(\tau)$  est proche de 1. On peut voir que cette propriété n'a rien d'automatique.

En effet, notons  $I_\tau(x)$  l'indicatrice de  $R$  au seuil  $\tau$ , alors

$$\begin{aligned} \text{Cov} [\bar{I}_\tau, \bar{R}] &= \text{Cov} \left[ \int_D I_\tau(x) dx, \int_D R(x) dx \right] \\ &= \int_D \int_D \text{Cov} [I_\tau(x), R(y)] dx dy, \end{aligned} \quad (4.2)$$

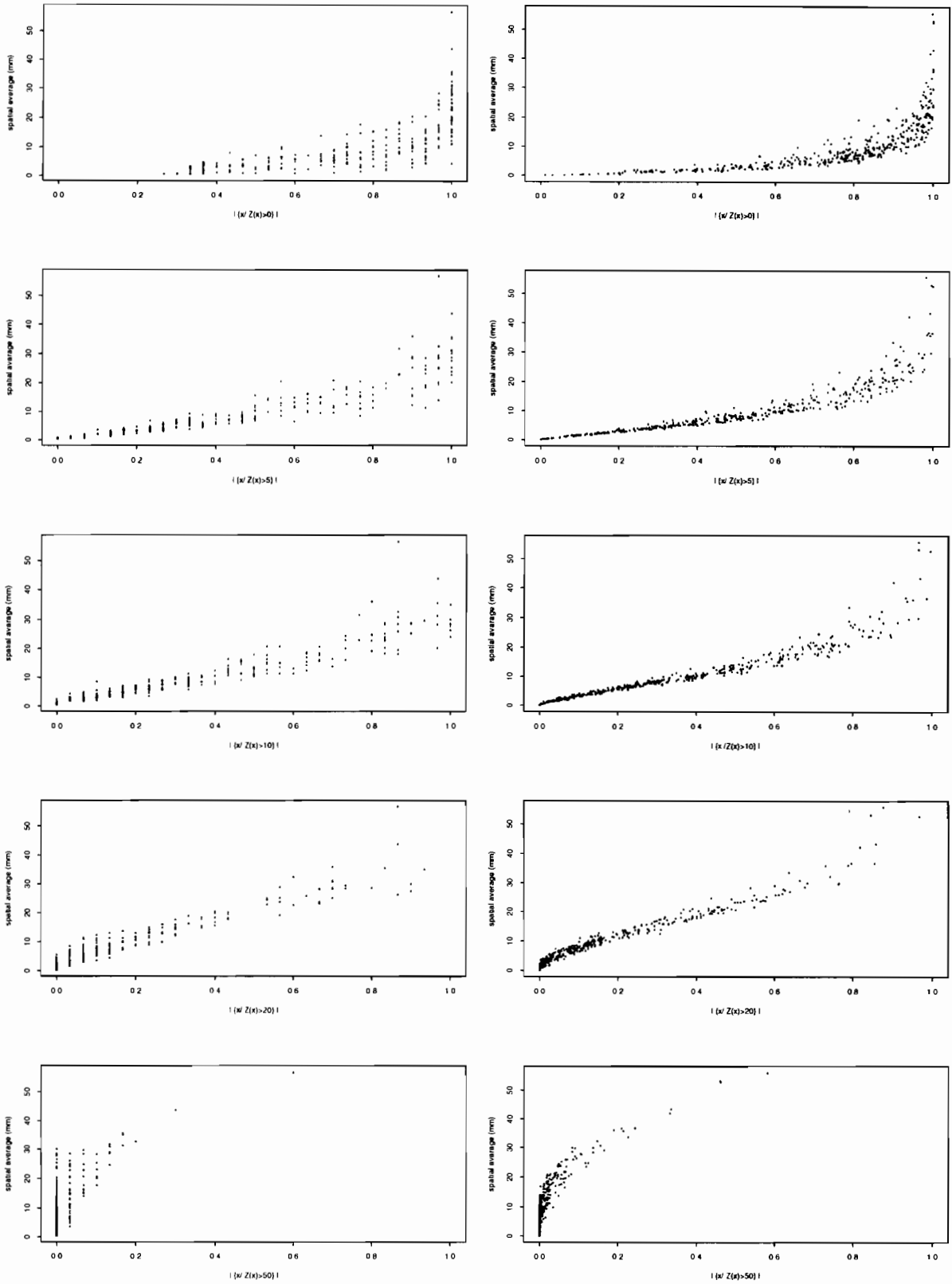


FIG. 4.12: Relation entre la fraction du domaine au-dessus d'un certain seuil (en abscisse) et la moyenne spatiale (en ordonnée). (De haut en bas :  $\tau = 0, 5, 10, 20, 50$  (mm) à gauche : données , à droite : simulations.

Le coefficient de corrélation linéaire entre  $\overline{I_\tau}$  et  $\overline{R}$  dépend donc de la covariance entre  $I_\tau(x)$  et  $R(x)$  qui n'a pas été spécifiée explicitement dans le modèle.

Le fait qu'une telle relation assez universellement observée sur les champs de pluie soit bien reproduite par le modèle constitue un point encourageant. Toutefois il convient de noter que ce type de relation ne donne pas d'information sur la géométrie des champs étudiés du point de vue de la connexité des surfaces sup-seuil, propriétés susceptible de jouer un rôle important dans les phénomènes hydrologiques de surface et de sub-surface. On sait que du point de vue théorique et dans le cadre gaussien, les propriétés de connexité sont gouvernées par le comportement à l'origine de la covariance de la f.a [6]. Celui-ci a été examiné en détail, toutefois une validation à partir de données constituerait un complément intéressant qui est malheureusement inaccessible à partir des données EPSAT. L'étude directe de ces propriétés de connexité, requiert donc un travail supplémentaire à l'aide d'un jeu de données adapté, constitué par exemple d'images radar.

## 4.5 Conclusion

La capacité du modèle à reproduire les propriétés de chaque classe en terme d'effectifs et de statistiques simples est une propriété très forte puisqu'elle n'a pas été incluse explicitement dans le modèle. Ceci permet de simuler de manière cohérente les principaux groupes d'événements S.C.M (lignes de grain et non lignes de grain) avec un seul modèle. La mauvaise représentation des événements de type 3 est à retenir mais son importance est de deuxième ordre puisqu'ils totalisent moins de 20% du cumul annuel en moyenne.

Concernant les propriétés de changement d'échelle, reposant sur la loi multivariée de la f.a, on observe une certaine capacité du modèle à reproduire les propriétés des champs de pluie pour les valeurs faibles et médianes, en particulier l'intermittence aux échelles inférieures à 500 km<sup>2</sup>. En revanche, le modèle semble avoir tendance à produire des valeurs excessivement fortes sur de grands domaines. Ce type de limitations n'est pas propre à cette famille de modèles mais survient en général dès que l'on compare un grand nombre de réalisations à un échantillon de taille limitée. Le réalisme de telles valeurs devrait donc faire l'objet de tests supplémentaires à partir des données de la période 1990-1999.

En regard des utilisations qui seront faites de ce modèle, on peut distinguer trois situations.

**La simulation en mode non conditionnel :** Dans ce cas, la tendance du modèle à sur-représenter les valeurs extrêmes peut constituer un problème important. Il conviendrait de le quantifier, par exemple en examinant à l'aide d'un modèle hydrologique quels types de conséquences résulteraient des valeurs très fortes remarquées à la section précédente.

**La désagrégation d'estimations satellitales :** Ici, le problème serait moins grave puisque les valeurs aux différentes échelles sont forcées par une climatologie réelle (celle qui découle de l'estimation satellitale) à travers la valeur de grande échelle à désagréger. C'est en fait plutôt la question inverse qui est posée : à partir d'une climatologie de grande échelle correcte et d'un modèle transférant imparfaitement les propriétés à travers les échelles, quelle est la distribution aux échelles inférieures qui en résulte ? En particulier, la distribution ponctuelle de départ, qui est nécessairement altérée, est-elle beaucoup modifiée ?

**La désagrégation de sorties de M.C.G.A :** Dans ce dernier cas, il convient de prendre en compte le fait que le M.C.G.A lui-même représente la climatologie des champs de pluie de manière très approximative, par exemple en terme de durée de vie, de vitesse et de pluviosité des SCM [104, 29]. Il est probable que dans cette configuration, l'écart commis par le modèle de désagrégation en terme de distribution est très inférieur à l'écart commis par un M.C.G.A, en terme de lame d'eau moyenne à grande échelle. Un travail de

comparaison entre les deux types de modèle est donc nécessaire avant toute application du présent modèle.

# Chapitre 5

## Cinématique

### Présentation

On étudie maintenant les S.C.M du point de vue de la cinématique de l'onde de pluie. Ce mouvement est modélisé à partir des champs de temps d'arrivée. On utilise ici un modèle gaussien non-stationnaire pour représenter ces champs. Les paramètres du modèle de champ de temps d'arrivée (une dérive climatologique, représentant une vitesse moyenne et la covariance des temps d'arrivée) sont ajustés pour chacun des trois sous-groupes de la classification. Ce modèle reproduit les moyennes climatologiques des mouvement observés et certaines propriétés plus fines. Il permet de simuler les temps d'arrivée qui seront utilisés en entrée du schéma de désagrégation temporel. À partir de maintenant, on travaillera sur les 170 événements de la période 1990-1993, qui sont fournis sur un réseau plus dense, utile pour l'analyse aux petits pas de temps.

## 5.1 Introduction

### 5.1.1 Difficultés associées à la notion de mouvement

On peut donner des sens très divers à la notion de mouvement des systèmes précipitants et plus généralement d'un système déformable [106, 107, 108, 109]) selon l'objectif que l'on poursuit et la nature des systèmes étudiés. Pour notre modèle de simulation, la structure interne des systèmes (c'est à dire les propriétés à l'intérieur des surfaces pluvieuses) étant modélisée via les hyétogrammes (voir chapitre 6), il est seulement nécessaire d'avoir une représentation des caractéristiques extérieures des systèmes.

Concernant les lignes de grain, on peut facilement définir un front (convectif) comme la frontière pluie/non pluie située à l'Ouest du système. L'étude des caractéristiques extérieures revient à étudier la progression de ce front en général à

peu près rectiligne et mu d'Est en Ouest. À l'opposé, les systèmes les moins organisés sont morcelés et la frontière pluie/non pluie est en général beaucoup plus tourmentée dans sa forme et dans son évolution (phénomènes de regroupement ou de scission des surfaces pluvieuses, par exemple).

Un autre problème est de faire un lien entre ces définitions et les données disponibles. Si on cherche à représenter le mouvement d'une ligne de grains, on est amené à estimer le champ de vitesses sur le front. En approximant localement le front par un segment de droite se déplaçant à vitesse constante, on peut estimer une vitesse pour tous les triplets de stations du réseau. Ceci conduit à estimer des vitesses de nature assez différente en terme d'échelle temporelle. En effet, certains triplets concernent des stations très proches et d'autres des stations très éloignées, les premiers permettent d'estimer des vitesses locales moyennées sur quelques minutes, les seconds des vitesses qui ne sont pas vraiment locales et qui sont relatives à des intervalles de temps qui peuvent excéder une heure. Pour avoir des estimations de quantités homogènes, on peut exclure les vitesses calculées sur des triplets de stations trop éloignées (ce qui nécessite un critère pour déterminer quand un triangle est trop grand).

Le résultat d'un tel exercice est reporté à la figure 5.1. On voit alors qu'on dispose de mesures en des points très irrégulièrement espacés. De plus, l'événement considéré était animé d'un mouvement global d'Est en Ouest ce qui n'apparaît pas sur les vitesses locales. Ceci peut-être imputé à des variations de petite échelle de la vitesse par rapport au mouvement de meso-échelle, ce qui implique qu'un modèle du mouvement des systèmes via le champ de vitesse locale devrait prendre en compte ces propriétés de petite échelle.

### 5.1.2 Mouvement et temps d'arrivée

On voit donc que la notion de mouvement des systèmes précipitants fait émerger d'une part des problèmes de définition, d'autre part des difficultés de mise en oeuvre pratique à partir des mesures du réseau au sol.

Ces difficultés peuvent être évitées si au lieu de chercher à estimer et modéliser directement le mouvement, c'est à dire relier des mesures de temps à des coordonnées d'espace, on se limite à la description de mesures de temps. Pour cela on va considérer le champ des temps d'arrivée étudié initialement pour le krigeage lagrangien [8].

Ce champ est défini de la manière suivante: pour un événement donné, la première station touchée par la pluie est prise comme origine des temps. Pour chaque station, le temps d'arrivée du système est la différence entre la date d'arrivée à la station considérée et la date d'arrivée à la station de référence.

Il se trouve qu'un modèle statistique de temps d'arrivée est suffisant pour notre algorithme de désagrégation. L'étude des temps d'arrivée a l'avantage de s'appliquer à une variable directement mesurée, et de faire l'économie d'hypothèses sur

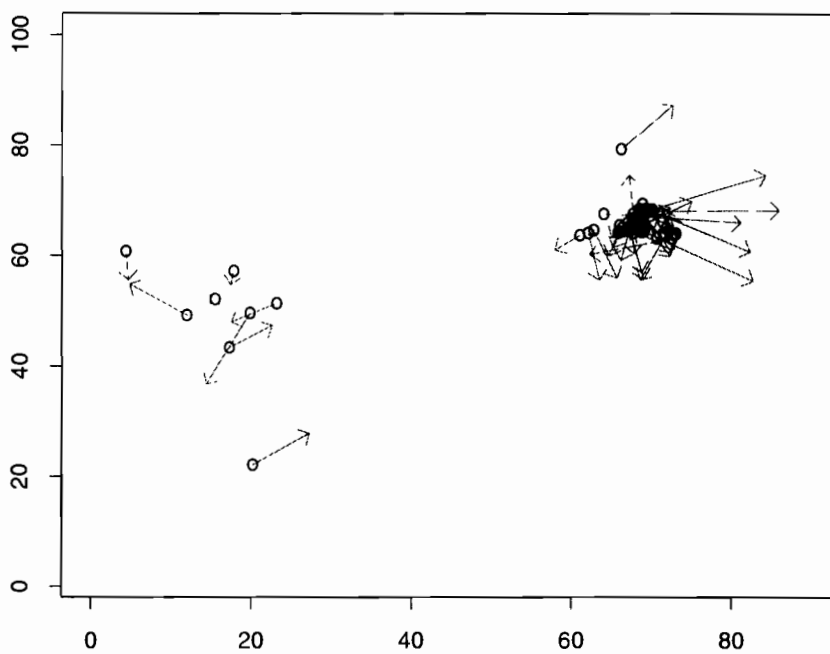


FIG. 5.1: *Champ de vitesse calculé au passage d'un front convectif à partir des stations du réseau EPSAT. Le champ de vitesse peut sembler isotrope alors que le mouvement général du système correspondant était plutôt orienté vers l'Ouest. Cette impression est partiellement imputable au fait que cette carte ne présente pas des vitesses estimées aux mêmes échelles de temps et d'espace selon les points considérés.*

	unités	groupe 1	groupe 2	groupe 3
$a$	$min$	155.8	157	133
$b$	$min.km^{-1}$	-0.9	-0.74	-0.62
$c$	$min.km^{-1}$	-0.29	-0.22	-0.11
pente	$km.h^{-1}$	63.7	77.9	95.76

TAB. 5.1: Coefficients des plans moyens de temps d'arrivée obtenus par minimisation de l'erreur quadratique.

la régularité du mouvement. Une fois le modèle spécifié, on pourra toujours déduire des propriétés relatives à des vitesses locales à partir du modèle de temps d'arrivée.

## 5.2 Analyse des données

### 5.2.1 Moyenne et variance

Étant donnée la prédominance des flux d'Est en Ouest dans cette région, on doit s'attendre à observer des champs de temps d'arrivée fortement non stationnaires en moyenne. Ceci est mis en évidence en calculant la moyenne et la variance du temps d'arrivée en chaque station, par groupe d'événements. Ces valeurs aux stations sont interpolées linéairement et présentées aux figures 5.2, 5.3 et 5.4.

La prédominance du mouvement d'Est en Ouest apparaît nettement pour chaque groupe. De plus, on retrouve un résultat attendu : le gradient de ces plans pointe dans la direction de l'axe principal d'anisotropie des champs de cumuls événementiels décrits dans la section 3.2.

Le champ de la moyenne relatif aux événements du groupe 1 semble plus régulier (proche d'un plan) que ceux relatifs aux autres groupes. Cet écart au caractère linéaire du temps moyen s'interprète comme un écart au caractère uniforme du mouvement.

La différence entre les groupes est plus nette en terme de variance. L'écart type est de l'ordre d'une heure pour chaque groupe (environ 50 minutes pour le groupe 1, 80 minutes pour le groupe 2) mais les valeurs de la variance sont beaucoup plus dispersées dans l'espace pour les groupe 2 et 3 que pour le groupe 1. Pour tous les groupes, la variance ne semble pas présenter de variations systématiques dans l'espace.

Si on ajuste un plan (de la forme  $a + b\xi_1 + c\xi_2$  où  $x = (\xi_1, \xi_2)$ ) à ces champs moyens par moindre carré, la pente du plan est homogène à une vitesse. On obtient les résultats du tableau 5.2.1.

Pour les événements du groupe 1 cette «vitesse» est conforme aux observations radar obtenues sur la zone [72]. Pour les deux autres groupes, cette pente n'a pas grand sens puisqu'elle correspond à une vitesse globale et on trouve des valeurs qui n'ont évidemment jamais été observées pour ces systèmes. Il faut d'ailleurs noter que cette «vitesse» globale est d'autant plus grande que l'on a affaire à un

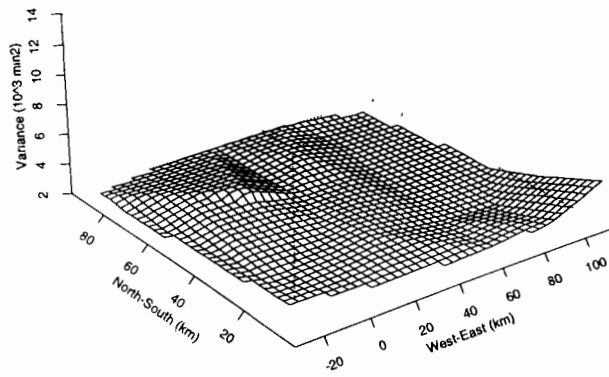
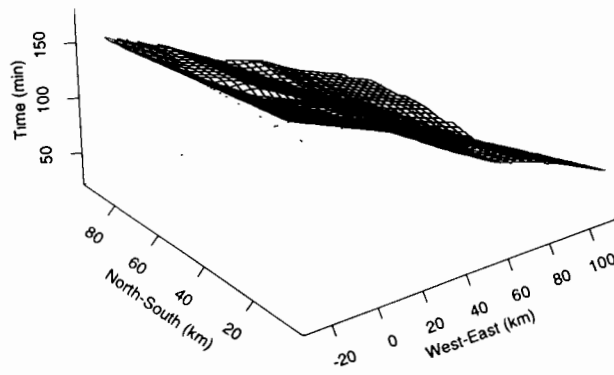


FIG. 5.2: Moyenne (en haut) et variance (en bas) du temps d'arrivée des événements du groupe 1

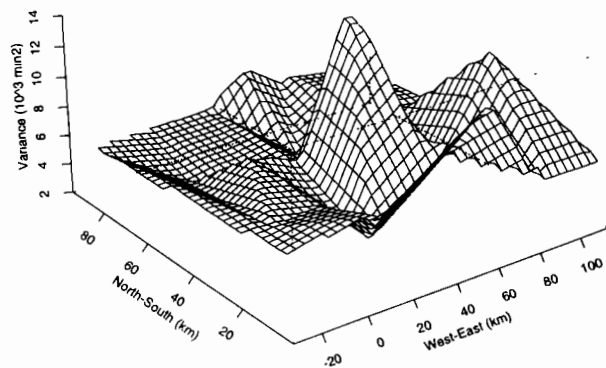
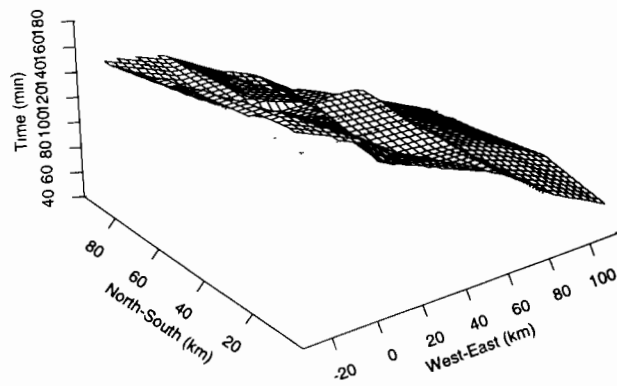


FIG. 5.3: Moyenne (en haut) et variance (en bas) du temps d'arrivée des événements du groupe 2

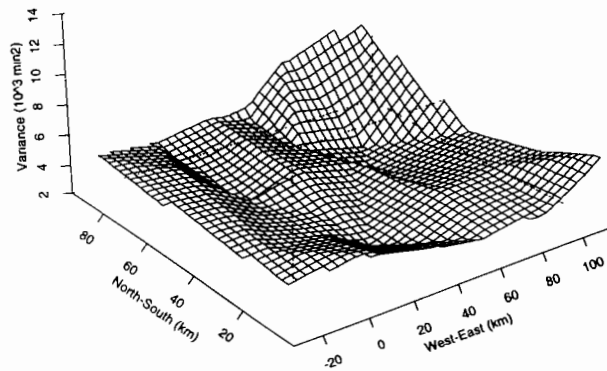
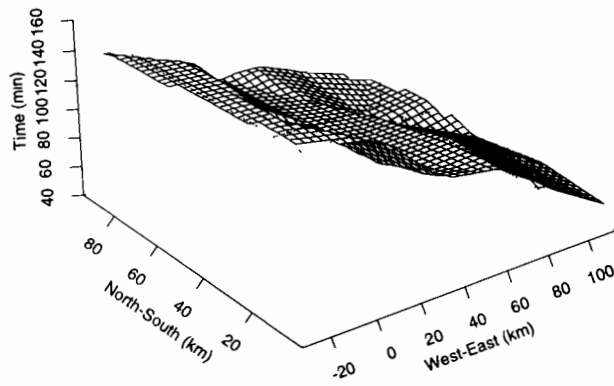


FIG. 5.4: Moyenne (en haut) et variance (en bas) du temps d'arrivée des événements du groupe 3

événement peu organisé. Or les observations satellitales et sol montrent que ce sont les SCM les mieux structurés qui présentent les vitesses moyennes les plus stables et les plus élevées.

### 5.2.2 Covariance

On retrouve également des différences marquées en terme de structure. La variance est plus élevée pour les groupes 2 et 3 et les covariances sont très dispersées. L'anisotropie est assez marquée pour le groupe 1, peu marquée pour le groupe 2 et inexistante pour le groupe 3.

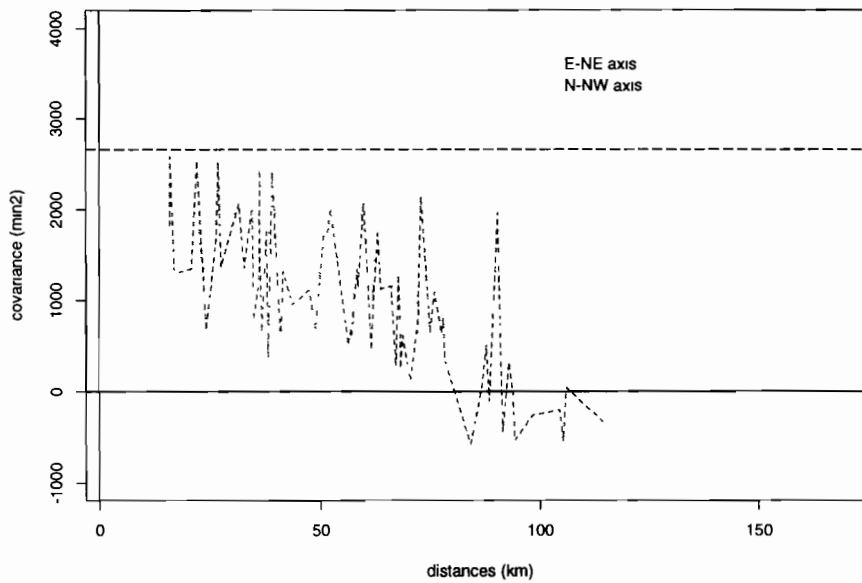


FIG. 5.5: Covariance des temps d'arrivée des événements du groupe 1

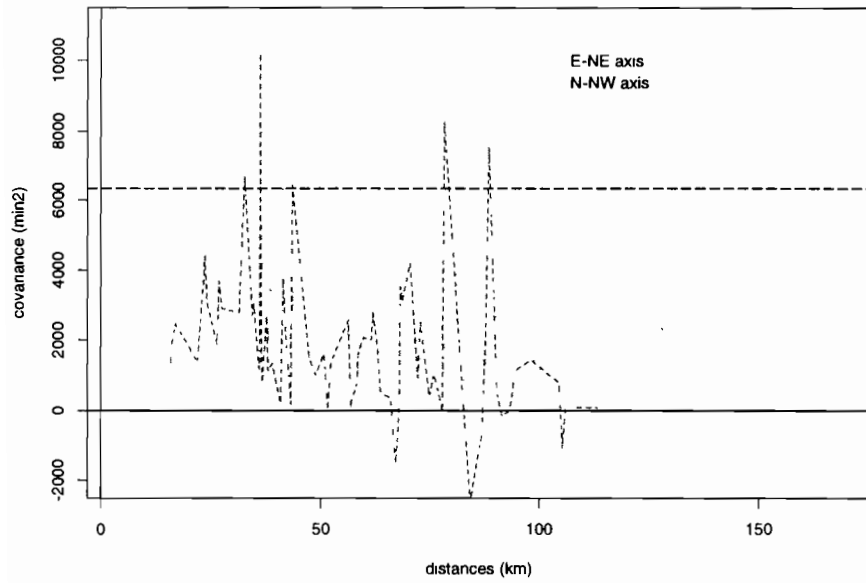


FIG. 5.6: Covariance des temps d'arrivée des événements du groupe 2

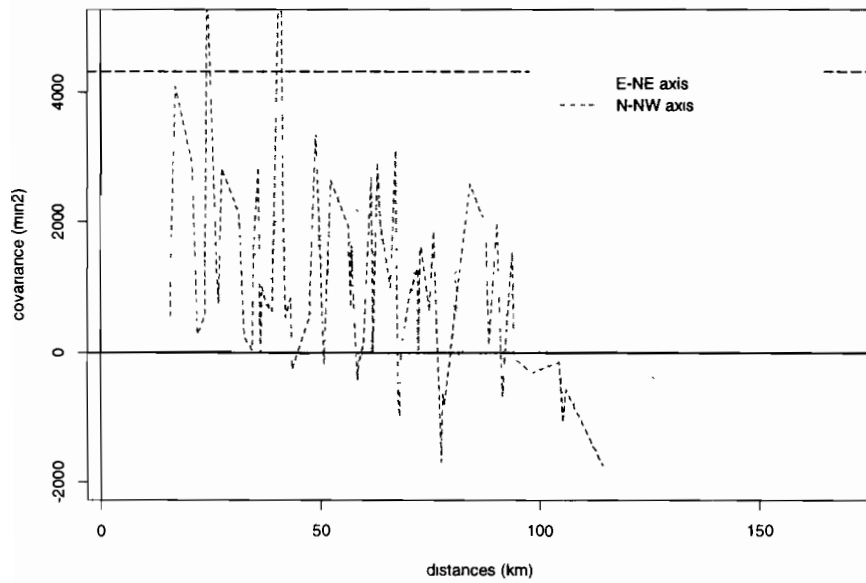


FIG. 5.7: Covariance des temps d'arrivée des événements du groupe 3

### 5.3 Modèle

On va représenter les temps d'arrivée de chaque groupe par une fonction aléatoire  $T$  sur  $\mathbb{R}^2$ . On décompose  $T$  en une partie déterministe et une partie aléatoire :

$$T(x) = E[T(x)] + \eta(x) \quad (5.1)$$

On considèrera que  $E[T(x)]$  est une fonction affine des coordonnées d'espace :

$$E[T(x)] = a + b\xi_1 + c\xi_2 \text{ où } x = (\xi_1, \xi_2) \quad (5.2)$$

Le résidu  $\eta(x)$  sera supposé gaussien stationnaire. Il s'agit d'une hypothèse très forte qui peut sembler abusivement simplificatrice. En effet, les temps d'arrivée sont positifs par définition ce qui n'est plus assuré dès qu'on les suppose gaussiens ; de plus les histogrammes des lois monovariabiles sont assez dissymétriques (figure 5.8).

Considérant que les champs de temps d'arrivée doivent présenter une certaine régularité (compte tenu de leur définition), nous supposerons en outre qu'ils admettent une covariance gaussienne (ce qui est probablement trop régulier). C'est à dire de la forme

$$C(h) = \exp[-|h|^2/s^2] \quad (5.3)$$

où  $|\cdot|$  est une norme anisotrope qui dépend du groupe d'événement considéré.

On peut maintenant calculer le champ de vitesse  $V$  associé à ce modèle par

$$V(x) = \frac{1}{|\text{grad } T(x)|^2} \text{grad } T(x) \quad (5.4)$$

On peut remarquer que cette écriture, qui traduit une intuition physique, soulève un certain nombre de questions. On peut se demander par exemple en quel sens et à quelle condition on peut parler du gradient d'une fonction aléatoire. Ces questions sont difficiles et dépassent très nettement le niveau de formalisation mathématique de ce travail. Faisant abstraction de ces difficultés, et d'un point de vue plus pragmatique, on peut remarquer que  $V$  n'est pas gaussien si  $T$  l'est, ce qui contredit une hypothèse qui est couramment faite dans la littérature et qui a montré son utilité [21], [100]. Nous allons voir néanmoins dans la section suivante que ce modèle est relativement bien adapté aux systèmes précipitants sahéliens, dans une certaine plage d'approximation.

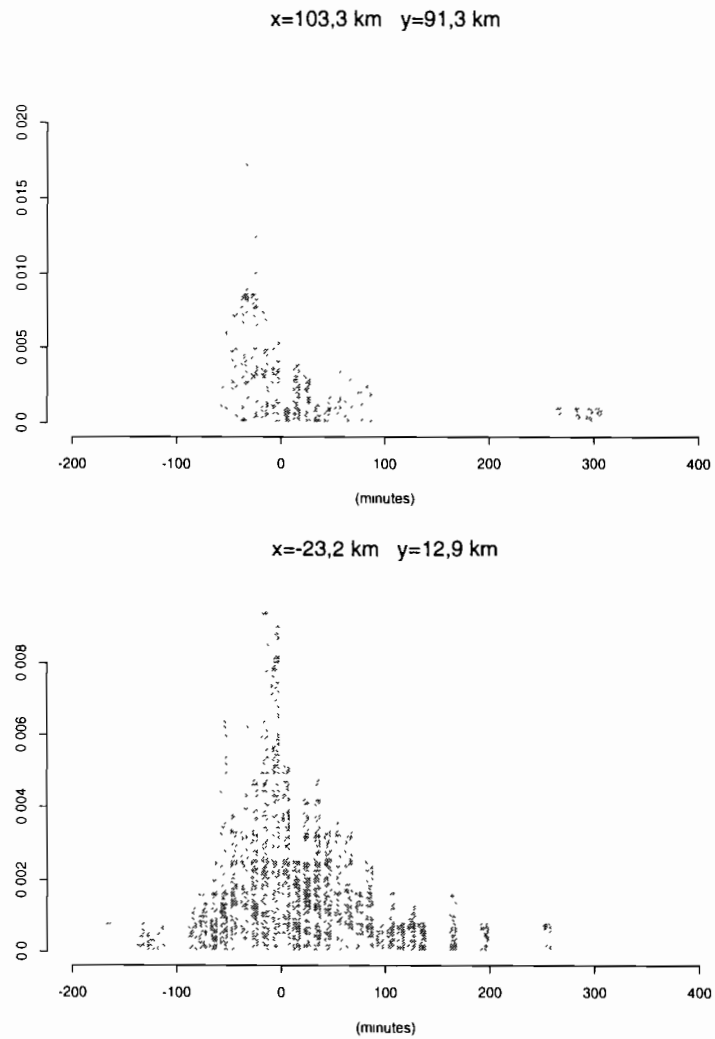


FIG. 5.8: *Histogramme des résidus de temps d'arrivée, calculés à partir des 170 événements de la période 1990-1993, en deux stations de part et d'autre du réseau (Alkama en haut et Yillade en bas)*

## 5.4 Comparaison aux données

La capacité de ce modèle à rendre compte des mouvements observés est illustrée par un exercice de simulations. Un ensemble de champs de temps d'arrivée a été simulé pour chaque groupe avec des paramètres propres. Pour chaque événement simulé on ajuste un plan de moindres carrés qui donne une idée du mouvement associé à ces temps d'arrivée. Le même calcul est réalisé pour les événements réels.

Ce calcul a une interprétation en terme de vitesse seulement pour les événements du groupe 1, mais la pente et l'orientation de ce plan donnent une description synthétique de chaque champ simulé. On observe alors (figures 5.9, 5.10 et 5.11) un bon agrément en terme de module et de direction entre les données et les simulations.

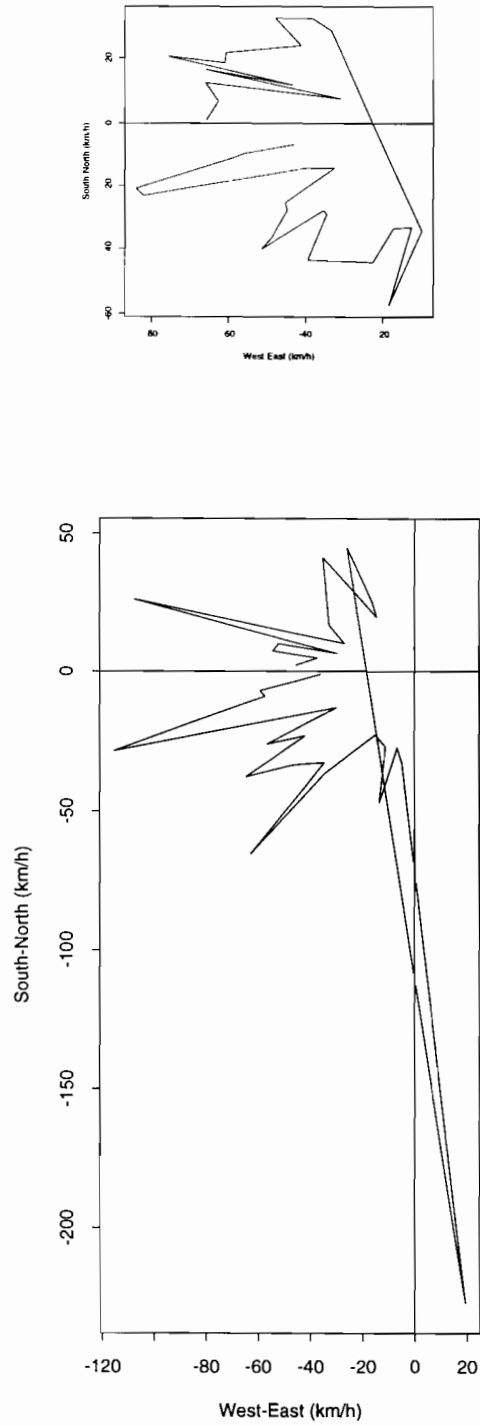


FIG. 5.9: Rose des vents des mouvements des événements du groupe 1. (Données en haut, simulations en bas)

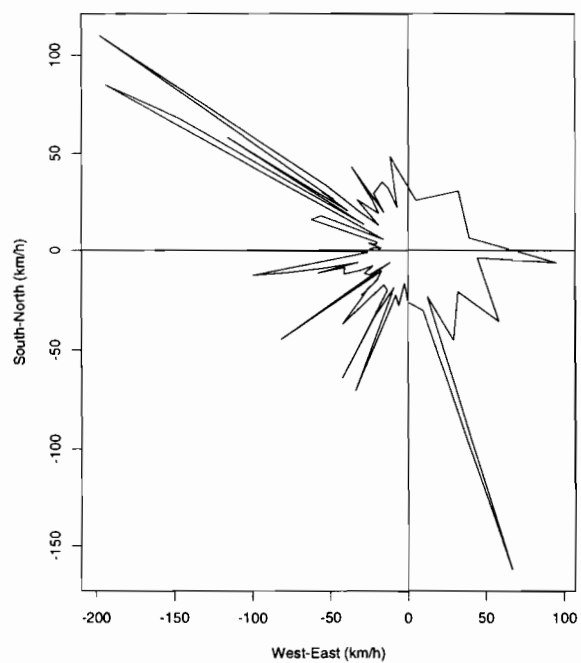
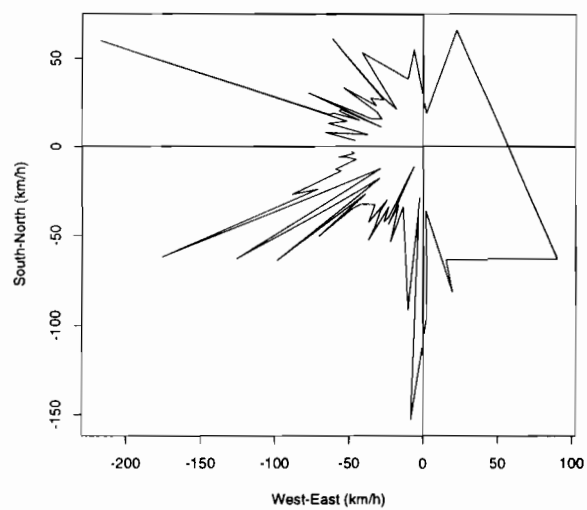


FIG. 5.10: *Rose des vents des mouvements des événements du groupe 2. (Données en haut, simulations en bas)*

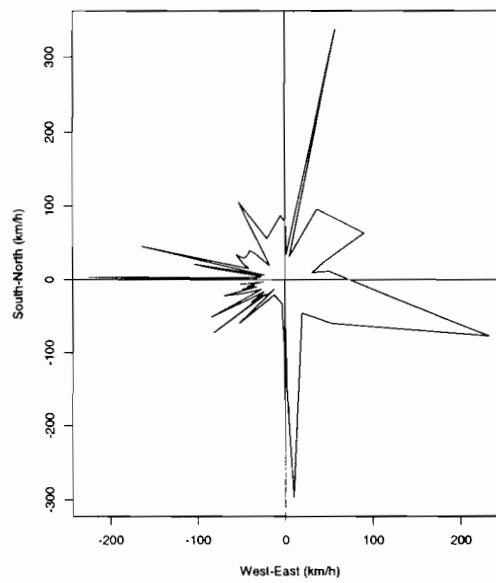
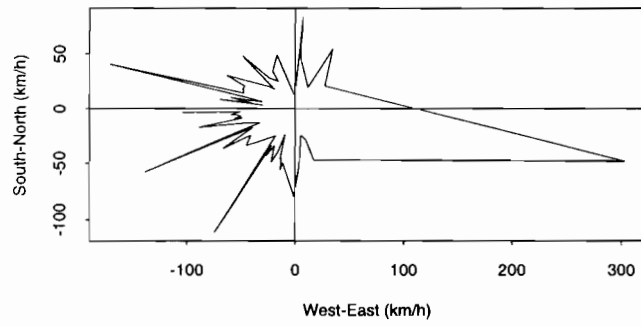


FIG. 5.11: *Rose des vents des mouvements des événements du groupe 3. (Données en haut, simulations en bas)*

## Chapitre 6

# Hyétogrammes

### Présentation

Nous avons spécifié un modèle pour les propriétés spatiales au pas de temps de l'événement et un modèle de la cinématique des systèmes. Pour être en mesure de simuler un événement aux petits pas de temps il faut disposer d'une information sur la manière dont se répartissent les cumuls en chaque point du domaine. L'objet de ce chapitre est donc d'étudier les propriétés statistiques des champs de pluie aux petits pas de temps. Ceci sera fait via un modèle statistique de hyétogramme représentant la manière dont se répartissent les intensités pluvieuses au cours du temps en un point générique du domaine. On verra que l'on perd peu d'information en représentant un hyétogramme par un schéma géométrique simplifié, puis que l'on peut relier statistiquement les paramètres de ce schéma à la lame d'eau totale au point considéré.

### 6.1 Introduction, position du problème

La démarche la plus générale pour étudier et simuler les champs de pluie à un pas de temps inférieur à la durée d'un événement consiste à

- caractériser les champs de pluie aux petits pas de temps
- simuler ces champs conditionnellement au champ de cumuls événementiels (c'est à dire en respectant les propriétés statistiques des hyétogrammes conditionnés par la lame d'eau de l'événement simulé)

Ceci pose plusieurs problèmes difficiles :

- caractériser des champs non stationnaires
- modéliser la dépendance entre les champs appartenant à un même événement qui ne peuvent pas être considérés comme indépendants

Nous allons donc aborder le problème d'une manière plus simple en étudiant la dépendance temporelle localement, c'est à dire en examinant le lien entre le champ de cumul-événement et la forme du hyétoگرامme en un point du domaine.

Le problème de la désagrégation temporelle en un point a été abordé très souvent dans la littérature depuis les travaux de Mandelbrot [81, 82, 83] et a donné lieu à des modèles très élaborés (basé sur des processus ponctuels ou sur la covariance temporelle) permettant de rendre compte de certains aspects de la loi temporelle des hyétoگرامmes. Nous allons montrer que concernant les systèmes convectifs sahéliens, un schéma extrêmement simple permet de rendre compte de certains traits essentiels des hyétoگرامmes.

En observant quelques hyétoگرامmes (figure 6.1), en particulier ceux relatifs aux lignes de grains, on s'aperçoit d'une part que l'essentiel du cumul-événement est dû à la contribution convective, d'autre part que la forme du hyétoگرامme semble relativement stable d'un événement à l'autre. Autrement dit, il semble que l'essentiel de l'information relative au hyétoگرامme soit contenue dans un paramètre de localisation en temps (la maximum d'intensité convective) et une amplitude (l'intensité du pic convectif). Ceci signifie que la donnée du cumul-événement et d'une relation générique (valable en tout point du domaine) entre ce cumul-événement et les deux paramètres pourrait suffire à retrouver l'évolution temporelle des intensités à de petites fluctuations près.

C'est cette idée générale que nous allons mettre en oeuvre. Ceci nécessitera de définir des paramètres relatifs à la forme du hyétoگرامme susceptibles de rendre compte le plus précisément possible de l'évolution temporelle, d'examiner les éventuels liens entre ces paramètres et le cumul-événement (autrement dit de regarder quelle est l'information apportée par le cumul-événement sur la forme du hyétoگرامme), enfin d'utiliser cette information pour réaliser la désagrégation en temps, en tout point du domaine.

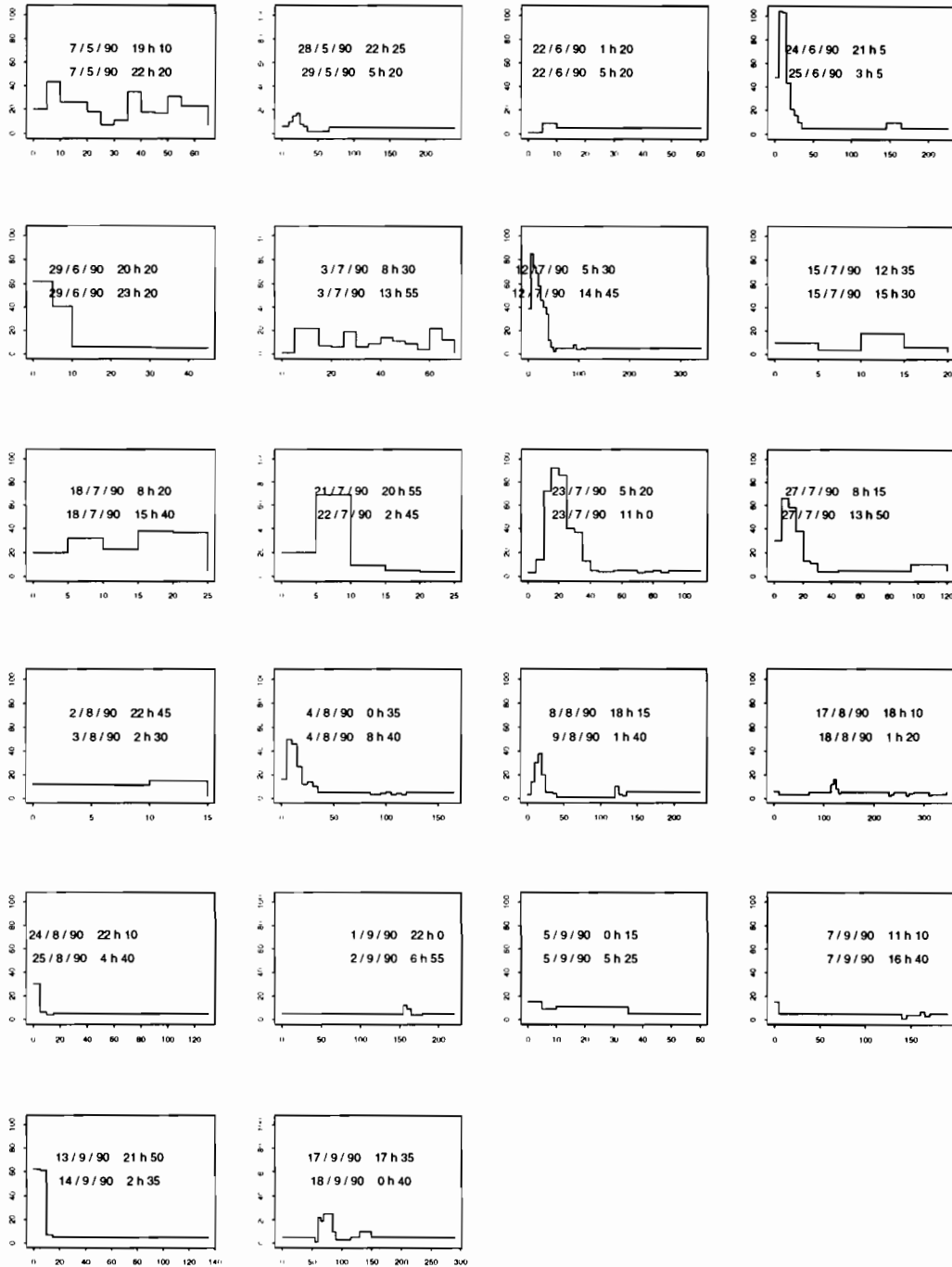


FIG. 6.1: Quelques hyétogrammes mesurés à la station de Berkiawal. (Temps en minutes, intensités en  $\text{mm.h}^{-1}$ ).

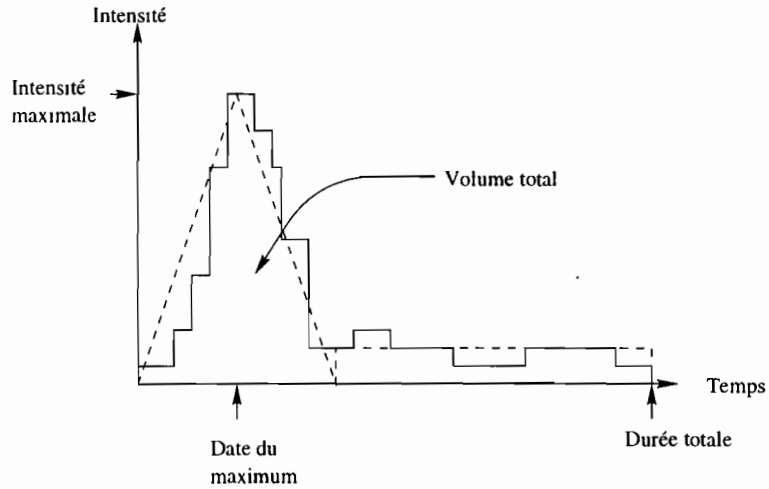


FIG. 6.2: Schéma d'un hyétogramme vrai et de sa représentation par quatre paramètres

## 6.2 Étude de la forme des hyétogrammes

### 6.2.1 Représentation simplifiée

À partir de maintenant nous allons travailler sur les données à des pas de temps inférieurs à la durée d'un événement. Comme il a été dit en introduction, les données brutes sont les dates de basculement des augets, ce qui signifie qu'il n'y a pas d'échelle de temps caractéristique attachée à cette mesure, ou plutôt, l'échelle de temps dépend de l'intensité pluvieuse et varie au cours du temps. On peut donc en principe se fixer une échelle de travail arbitrairement petite. Nous allons ici travailler sur des données au pas de temps de cinq minutes, ce qui permettra d'obtenir des propriétés à des échelles plus grossières par intégration.

Un hyétogramme au point  $x$  se présente alors comme la donnée des volumes  $R(x, t_1), \dots, R(x, t_n)$  sur un certain nombre d'intervalles consécutifs.

Pour l'étudier nous allons nous intéresser à quatre paramètres :

- deux paramètres de position à savoir la durée totale  $D$  et la date du maximum  $T_m$ ,
- deux paramètres de taille à savoir l'intensité maximale  $I_m$  et le volume total  $R$ .

Ces quatre paramètres peuvent être directement calculés sur chaque hyéto-gramme mesuré.

De plus, si l'on suppose que l'intensité croît à un rythme constant jusqu'au maximum, qu'elle décroît symétriquement après ce maximum et que l'intensité est constante après, alors le hyéto-gramme est entièrement déterminé par ces quatre paramètres (figure 6.2).

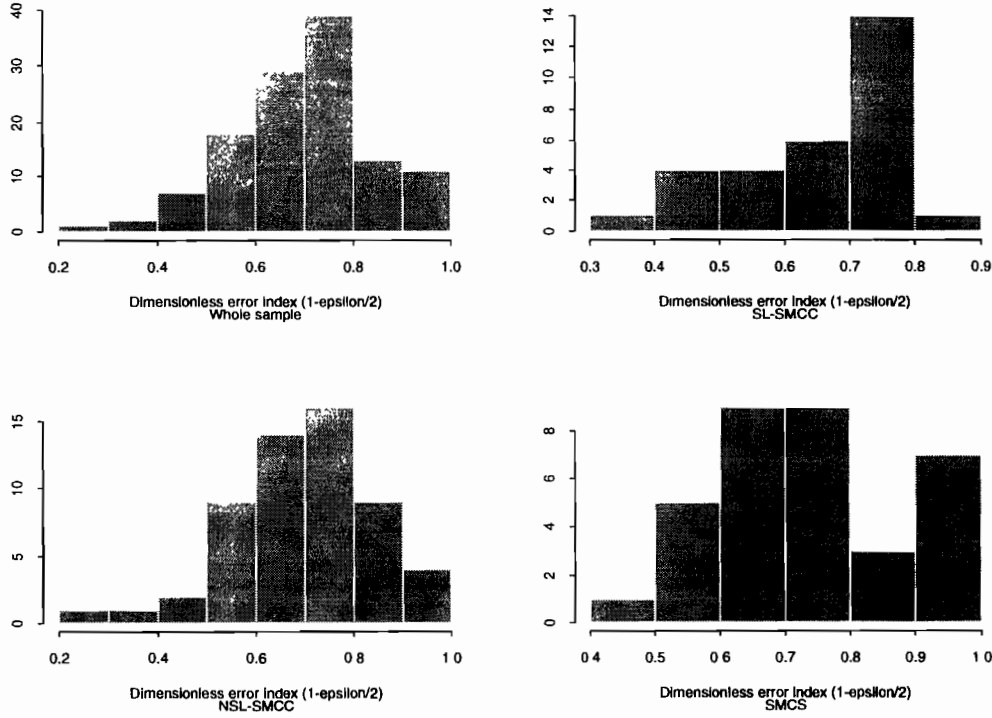


FIG. 6.3: *Histogramme (en nombre d'événements) de l'index  $1 - \epsilon/2$  représentant la fraction du volume total bien localisé par le modèle de hyétogramme*

### 6.2.2 Erreur commise par le modèle de hyétogramme

On peut quantifier l'écart entre ce schéma très simple et les intensités vraies en calculant un index d'erreur  $\epsilon$ .

$$\epsilon(x) = \frac{1}{\int r_x(t)dt} \int |r_x(t) - r_x^*(t)|dt \tag{6.1}$$

où  $r_x^*(t)$  est l'intensité modélisée

$1 - \epsilon/2$  représente la fraction du volume bien localisée par le hyétogramme modélisé  $r_x^*$ .

Il apparaît d'après l'historgramme des erreurs (figure 6.3) illustré par quelques exemples (figure 6.4) que les hyétogrammes sont relativement bien représentés par ce schéma. En moyenne sur 170 événements, 70% du cumul événement sont bien localisés en temps par le hyétogramme modélisé. Il semble toutefois que la qualité de cette représentation ne soit pas liée au cumul-événement (figure 6.5).

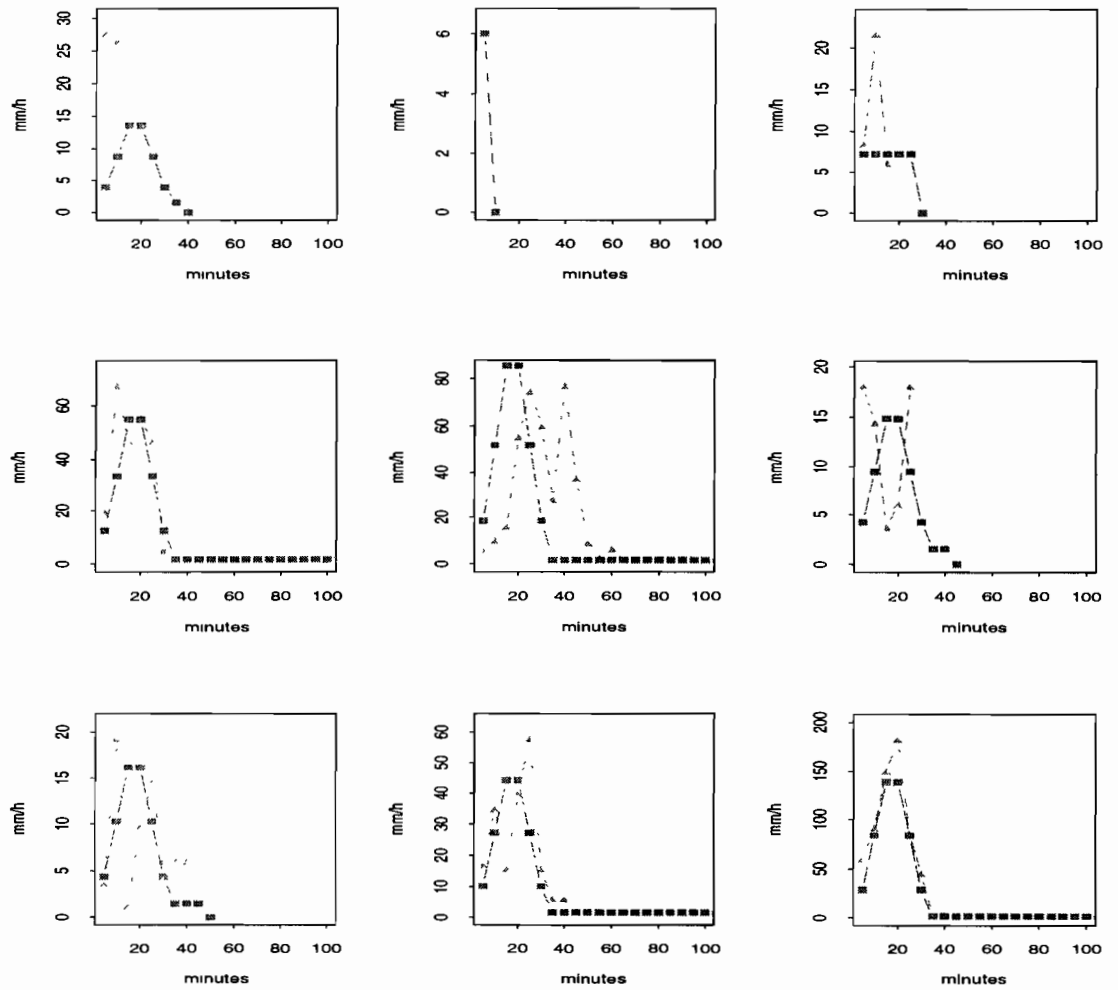


FIG. 6.4: Exemple de hyétogrammes observés à Berkiawal en 9 événements survenus en 1990 et de leur représentation par un schéma à quatre paramètres.

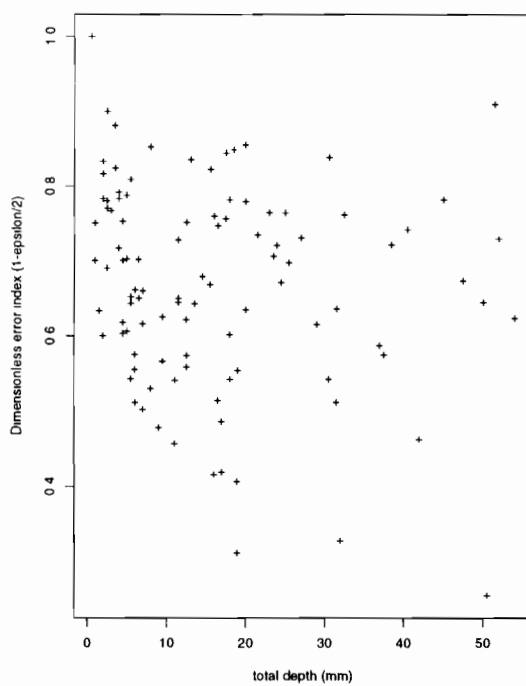


FIG. 6.5: *Index d'erreur en fonction du cumul-événement (le coefficient de corrélation linéaire vaut -0.32)*

### 6.2.3 Relation entre la forme du hyétogramme et le cumul-événement

Si l'on s'en tient au schéma précédent, on peut considérer que la donnée des quatre paramètres apporte une information substantielle quant à la localisation en temps du cumul événement. Le problème est évidemment que ces quatre paramètres ne sont pas connus lorsqu'on veut réaliser une désagrégation temporelle.

Nous allons donc examiner les liens entre ces paramètres et la seule information connue à cette étape, c'est à dire les cumuls événements.

Nous allons d'abord regarder quelques statistiques relatives aux distributions monovariante de ces paramètres afin de connaître leur variabilité.

	Min.	Médiane	Moyenne	Max.	Variance	c.v
Intensité max. (mm.h <sup>-1</sup> )	0	28,8	42,5	347	1438	0,89
Volume (mm)	0	10	14,7	162	263	1,1
Date du max. (min.)	0	5	18,4	575	1940	2,39
Durée totale (min.)	0	90	116,9	740	12771	0,97

TAB. 6.1: *Statistiques monovariante simples des paramètres relatives aux 170 événements de la période 1990-1993*

Pour examiner la relation locale entre tous les paramètres du hyétogramme, nous avons calculé les coefficients de corrélation linéaire (tableau 6.2) des couples de paramètres. Les nuages correspondant à chaque couple sont présentés aux figures 6.6, 6.7, et 6.8.

Du tableau 6.2, il apparaît qu'en terme de liaison linéaire, l'information apportée localement par le volume total sur la forme du hyétogramme est assez pauvre. Les coefficients de corrélation sont faibles sauf pour le couple intensité maximale / volume.

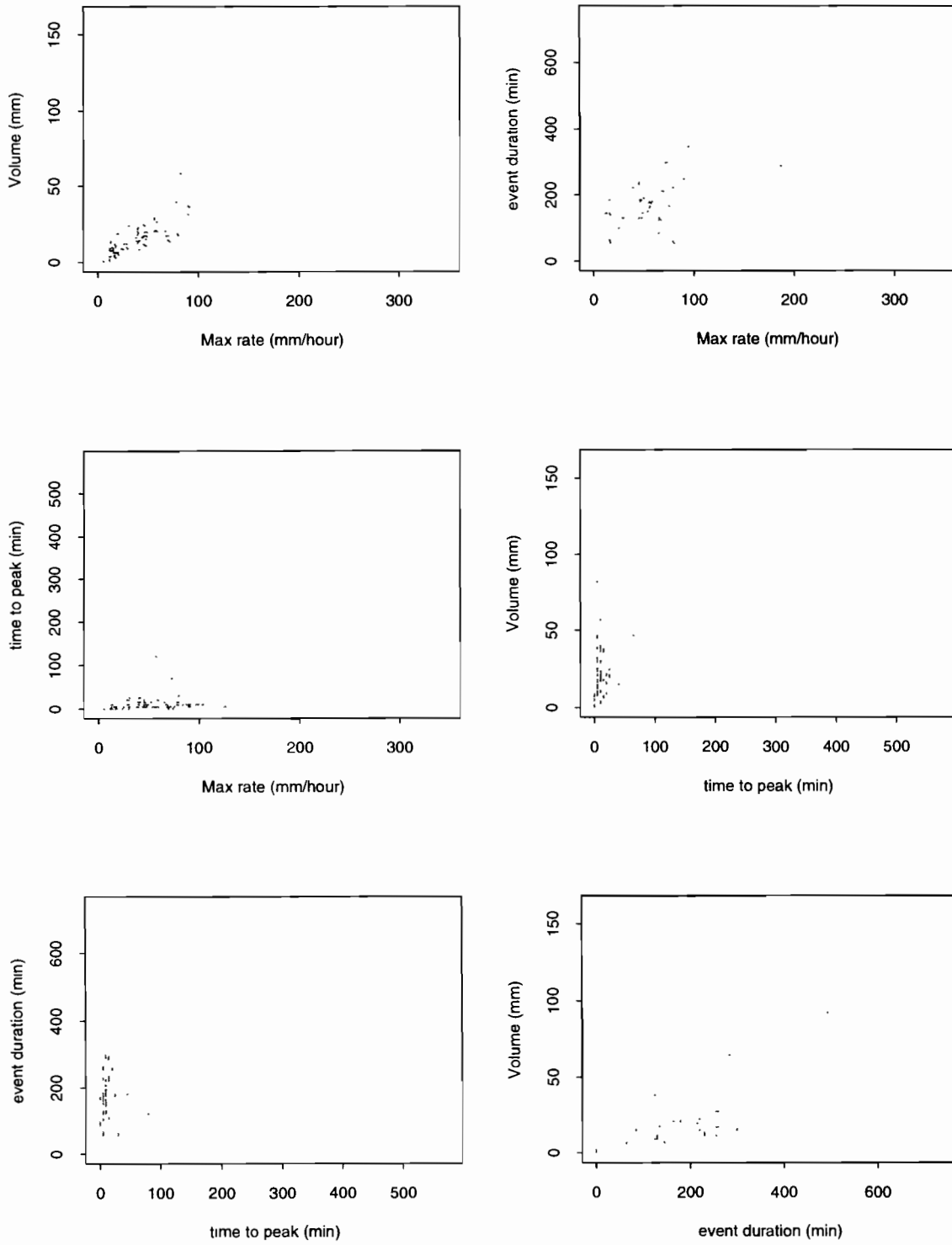


FIG. 6.6: Relation entre les paramètres de hyétogrammes. Groupe 1

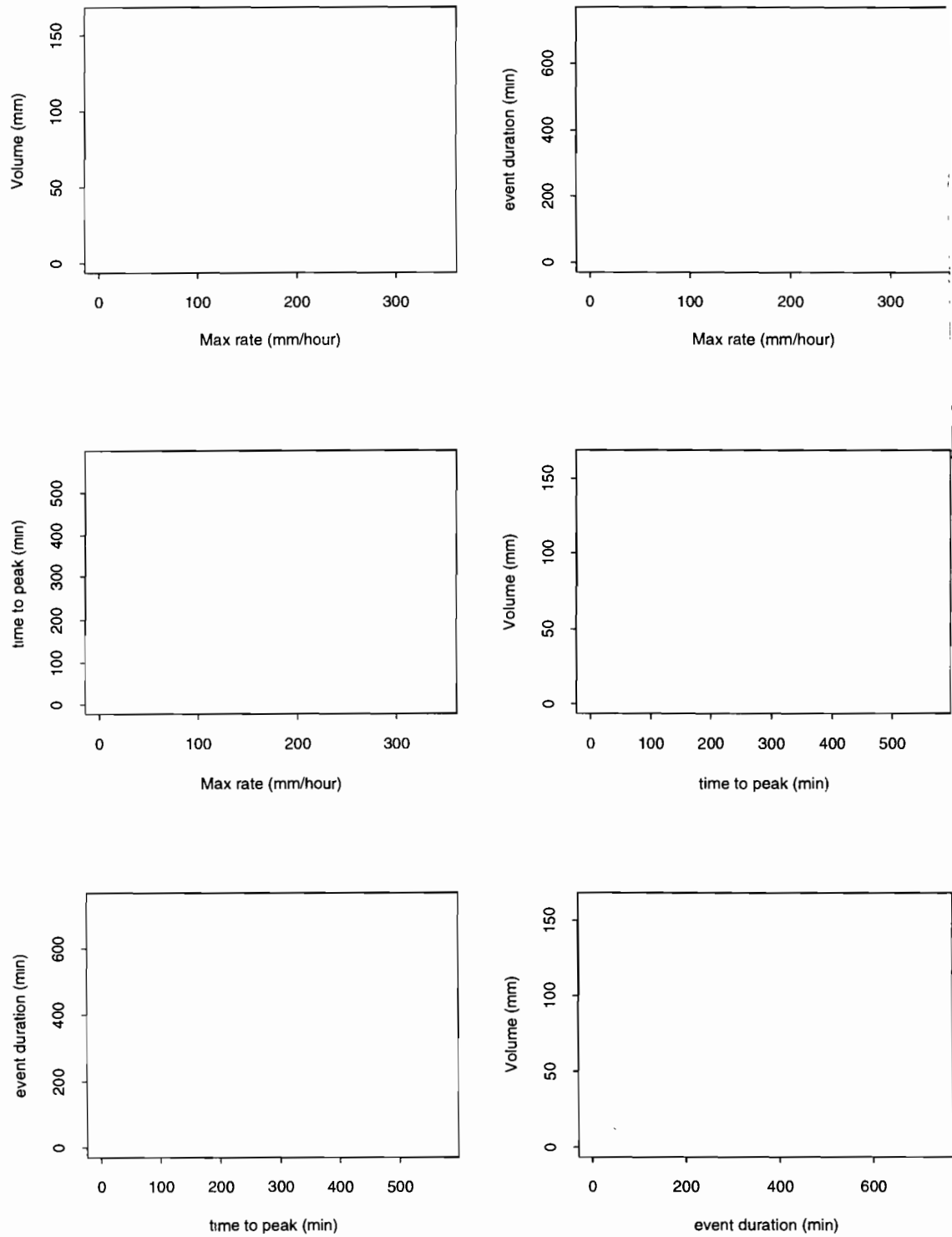


FIG. 6.7: Relation entre les paramètres de hyétogrammes. Groupe 2

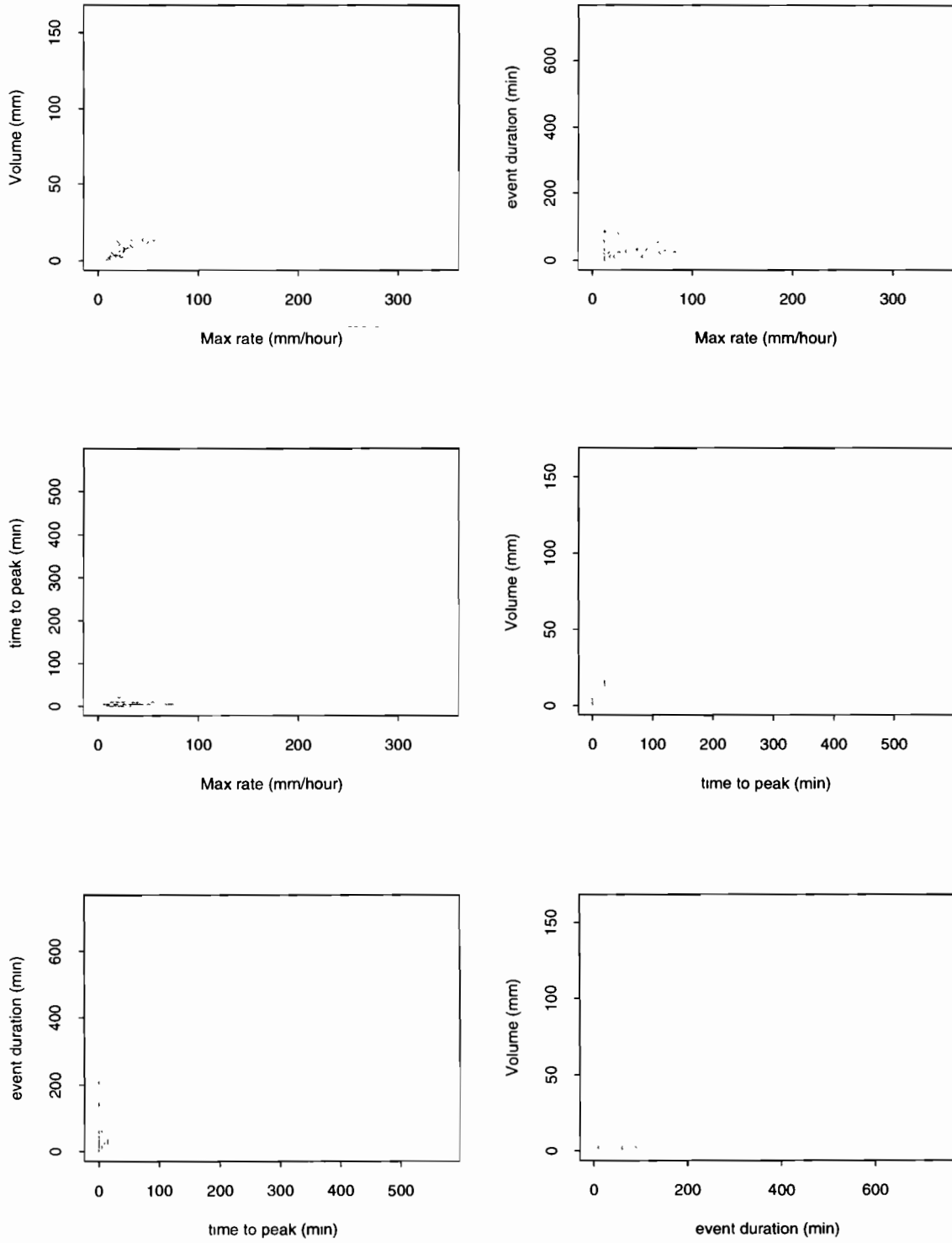


FIG. 6.8: Relation entre les paramètres de hyétogrammes. Groupe 3

	Ensemble	groupe 1	groupe 2	groupe 3
Intensité max. / Volume	0,78	0,71	0,76	0,85
Intensité max. / Durée totale	0,3	0,22	0,27	0,23
Intensité max. / Date du max.	0,025	-0,15	0,06	0,05
Date du max. / Volume	0,227	$10^{-4}$	0,33	0,16
Date du max. / Durée totale	0,42	0,1	0,54	0,45
Durée totale / Volume	0,57	0,52	0,56	0,43

TAB. 6.2: Coefficient de corrélation linéaire pour tous les couples de paramètres

### 6.3 Schéma de désagrégation

On peut rendre compte de la dépendance entre le volume total  $R$  et l'intensité maximale  $I_m$  par une relation linéaire, le cas des autres couples est moins simple. Le couple  $(D, R)$  peut également faire l'objet d'une modélisation par une relation linéaire, mais la représentation sera artificiellement rigide. Concernant la position du maximum, on peut observer que c'est apparemment la variable la plus dispersée (au sens du coefficient de variation). Cette dispersion provient probablement de la présence de hyétogrammes ne présentant pas de pic convectif et pour lesquels le maximum n'est pas obtenu au passage du front. Cette affirmation est corroborée par le fait que les hyétogrammes pour lesquels on observe un maximum tardif correspondent à des volumes faibles (figures 6.6, 6.7 et 6.8). On voit par ailleurs que pour les fortes valeurs du volume, le maximum est précoce, en général il survient avant une heure. Autrement dit, lorsque la contribution du pic convectif est forte, c'est à dire lorsqu'il est important de bien localiser ce pic, il se trouve en général au début.

En choisissant le schéma suivant :

- la durée totale  $D$  est proportionnelle au volume total  $R$ :  $D = \tau V$
- une fraction  $\alpha$  du volume total est répartie uniformément sur cette durée totale
- le volume résiduel est réparti selon un schéma triangulaire de durée  $Inf(D, 2\delta)$  où  $\delta$  est un paramètre constant.

Alors,

- $D(x) = \tau R(x)$  en tout point  $x$ , où  $\tau$  est uniforme et non aléatoire
- $R_{strat}(x) = \alpha R(x)$  en tout point  $x$ , où  $\alpha$  est uniforme et non aléatoire
- $I_m = \begin{cases} \alpha/\tau + 2V(x)(1 - \alpha)/\delta & \text{si } D \geq 2\delta \\ 2 - \alpha/\tau & \text{si } D \leq 2\delta \end{cases}$

Ce schéma dépend de trois paramètres :  $\tau$ ,  $\alpha$  et  $\delta$ . On peut les choisir de manière à restituer la durée moyenne observée sur les données et la moyenne et la variance des intensités maximales.

Ceci conduit à  $\tau = 1/8h.mm^{-1}$ ,  $\alpha = 0,2$  et  $\delta = 30min$ .

## 6.4 Éléments de validation

### 6.4.1 Distribution des intensités maximales

Il se trouve qu'en reproduisant les propriétés à l'ordre deux des intensités maximales avec ce schéma, on reproduit également très bien leurs propriétés aux ordres supérieurs (figure 6.9).

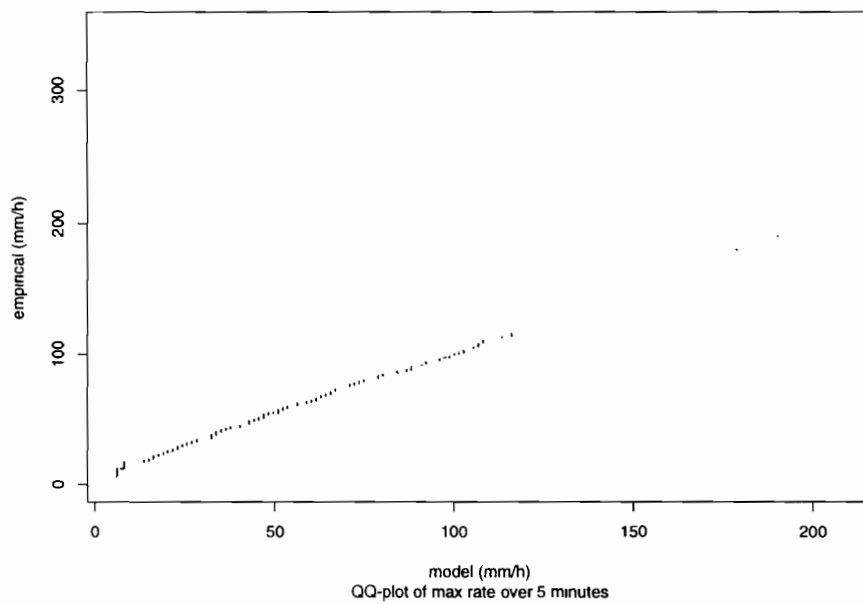


FIG. 6.9: Graphique quantile-quantile des intensités maximales au pas de temps de cinq minutes

### 6.4.2 Structure spatiale des champs d'intensité maximale

Avec ce schéma, l'intensité maximale se déduit du volume total par une transformation linéaire (à l'exception des volumes très faibles), la covariance spatiale des champs d'intensité maximale et la covariance spatiale croisée du couple  $(I_m, R)$  sont donc identiques (à la variance près) à celle des champs de cumul-événement.

Nous allons ici examiner si cette propriété du modèle est conforme aux observations.

On calcule pour cela la covariance des champs d'intensité maximale définie par

$$C_{I_m}(h) = \frac{1}{n} \sum_{k=1}^n (I_{mk}(x) - m_I)(I_{mk}(x+h) - m_I) \quad (6.2)$$

et la covariance spatiale croisée définie par

$$C_{I_m, R}(h) = \frac{1}{n} \sum_{k=1}^n (I_{mk}(x) - m_I)(R_k(x+h) - m_V) \quad (6.3)$$

On observe une assez grande similitude entre ces trois covariances (figure 6.10). En particulier, on retrouve le fait que la covariance ne s'annule pas aux grandes distances. En revanche, il semble que les champs d'intensité maximale se destructurent plus vite que les champs de cumul (la covariance entre paires de stations distantes d'environ 50 km est de l'ordre du cinquième de la variance pour les champ d'intensités maximales alors qu'elle est de l'ordre du tiers de la variance pour le champ de cumuls).

Du point de vue des champs d'intensités maximales, le modèle aura donc tendance à produire des champs trop structurés.

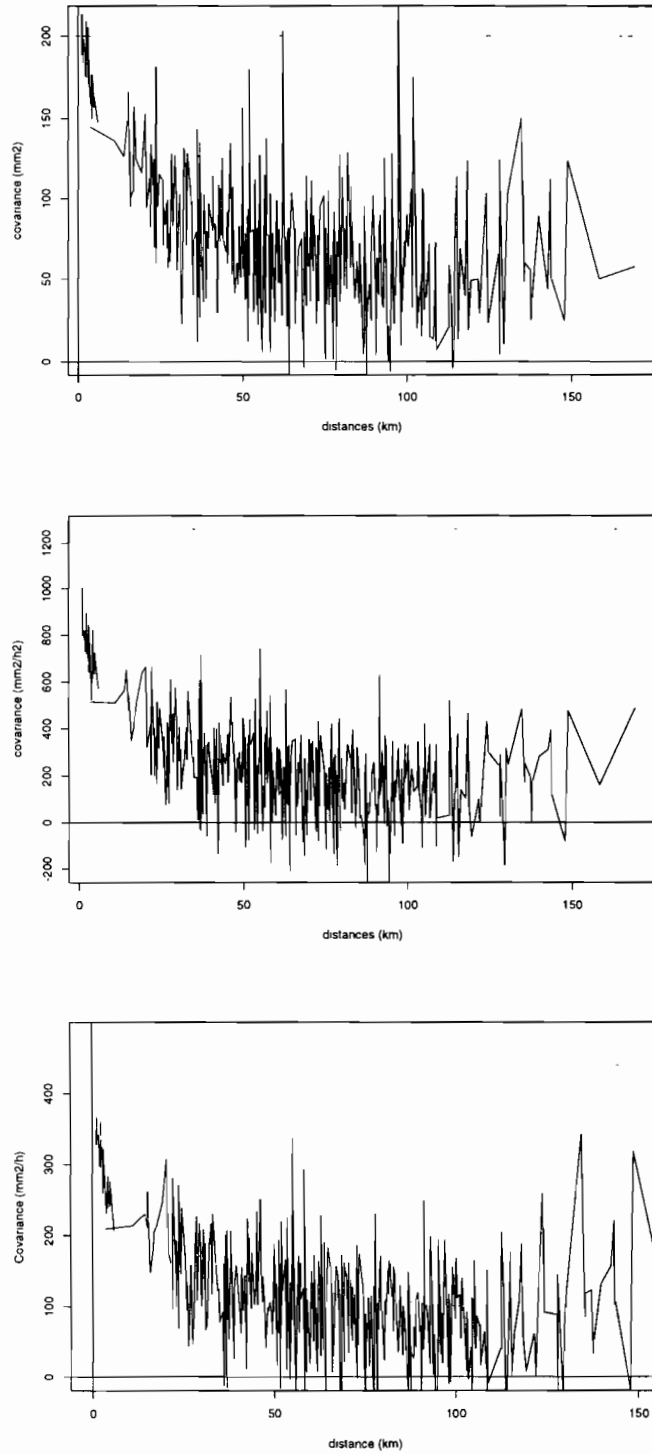


FIG. 6.10: De haut en bas : Covariance des champs de cumuls événementiels, covariance des champs d'intensités maximales (au pas de temps de 5 minutes), et covariance croisée (celle-ci est symétrique par rapport à l'axe des ordonnées)

## Chapitre 7

# Propriétés spatiales des champs à cinq minutes

### Présentation

Pour compléter l'analyse précédente, nous examinons les propriétés spatiales des champs de pluie faisant intervenir des moments d'ordre supérieur à deux. Pour cela, nous allons étudier les aires des surfaces où l'intensité pluvieuse dépasse un certain seuil. Ce travail est réalisé en comparant un jeu de 173 événements simulés aux 173 événements observés sur la période 1990-1993.

## 7.1 Problèmes de méthode

Le modèle développé dans les sections précédentes permet de simuler des champs de pluie au pas de temps de cinq minutes. Pour évaluer l'intérêt du modèle complet, il faut disposer d'un critère de comparaison entre les champs réels et ceux produits par le modèle.

À cet égard, on rencontre plusieurs problèmes :

- Les champs relatifs à un même événement sont fortement interdépendants (à travers le cumul événement et la vitesse du système notamment). La séquence des champs observés ne peut pas être considérée comme une suite de champs indépendants ;
- de plus, les champs observés ne sont pas stationnaires, au sens où ils présentent systématiquement des valeurs fortes le long du front.

Il n'est donc pas envisageable de caractériser la structure des champs au pas de temps de cinq minutes avec les outils statistiques habituels tels que la fonction de

covariance ou le variogramme. De plus, comme nous l'avons déjà souligné, les propriétés relatives à des échelles inférieures au  $\text{km}^2$  ne sont pas les seules à présenter un intérêt pour l'hydrologie de la région.

Comme il se trouve par ailleurs que les processus hydrologiques de surface sont sensibles aux phénomènes de dépassement de seuil, nous allons étudier les champs produits par le modèle à travers les tailles des surfaces sup-seuil.

Nous allons définir pour chaque champ  $R_k(x, t)$  ( $k$  étant le numéro de l'événement et  $t$  le numéro du champ dans cet événement) la surface sup-seuil de niveau  $s$

$$\{x/R(x, t) > s\}$$

L'aire de cette surface, exprimée en fraction du domaine total, sera notée  $f_{k,t}(s)$ , i.e

$$f_{k,t}(s) = \frac{1}{|D|} |\{x/R(x, t) > s\}|$$

Pour chaque champ simulé  $R_k^*(x, t)$ , on définit de manière analogue

$$f_{k,t}^*(s) = \frac{1}{|D|} |\{x/R^*(x, t) > s\}|$$

Il est important de noter qu'une suite du type  $\{f_{k,t}(s)\}_{1 \leq t \leq t_k}$  ne peut pas être considérée comme une suite de v.a indépendantes. Nous allons étudier ces variables seulement d'un point de vue descriptif.

Les distributions des tailles de ces surfaces n'étant pas accessibles analytiquement, nous avons simulé trois jeux d'événements (un pour chaque groupe). Les calculs relatifs à chaque groupe d'événements simulés sont réalisés sur une grille comparable à celle du jeu de données.

## 7.2 Aire moyenne des surfaces sup-seuil

En première approche, nous allons comparer le modèle et les données du point de vue de la valeur moyenne des fractions des aires des surfaces sup-seuil (f.a.s.s)

$$\bar{f}(s) = \sum_{k=1}^K \sum_{t=1}^{t_k} f_{k,t}(s) \quad (7.1)$$

$$\bar{f}^*(s) = \sum_{k=1}^K \sum_{t=1}^{t_k} f_{k,t}^*(s) \quad (7.2)$$

Les courbes  $(s, \bar{f}(s))$  et  $(s, \bar{f}^*(s))$  sont présentées pour  $s$  variant de  $0 \text{ mm.h}^{-1}$  (intermittence) à  $200 \text{ mm.h}^{-1}$  (figure 7.1).

On peut observer une bonne adéquation pour le groupe 1 pour les seuils supérieurs à  $20 \text{ mm.h}^{-1}$ . Pour les faibles intensités le modèle sous-estime les aires tandis qu'il les sur-estime pour le seuil  $s = 0 \text{ mm.h}^{-1}$ .

Pour les groupes 2 et 3, le modèle reproduit correctement l'intermittence, mais il sous-estime les aires moyennes correspondant aux autres seuils. Cette sous-estimation est sans doute une conséquence de la relation linéaire imposée entre le volume ponctuel et la durée de l'événement. Cette hypothèse très forte est évidemment assez arbitraire et crée une grande région derrière le front où règne une intensité uniforme et constante.

### 7.3 Distribution des aires des surfaces sup-seuil

On peut examiner de manière plus détaillée les f.a.s.s en comparant les  $(f_{k,t}(s))_{k,t}$  aux  $(f_{k,t}^*(s))_{k,t}$ .

Mais dans la mesure où un grand nombre de champs correspondent à la traîne stratiforme (pour les données), ou à cette grande région d'intensité uniforme (pour le modèle), nous n'allons considérer que les champs possédant au moins une valeur supérieure à  $2 \text{ mm.h}^{-1}$ .

Les fractions sont comparées pour chaque groupe et pour différents seuils (figures 7.2, 7.3 et 7.4).

De manière générale, il est intéressant de constater que les propriétés sont très différentes d'un groupe à l'autre du point de vue des données, du point de vue du modèle et du point de vue de l'adéquation du modèle aux données.

Pour le groupe 1, on observe une bonne concordance générale entre le modèle et les données. L'écart est maximal autour des faibles intensités avec une tendance du modèle à produire des surfaces trop grandes. Les comparaisons pour des seuils supérieurs à  $30 \text{ mm.h}^{-1}$  sont à considérer avec beaucoup de prudence dans la mesure où elles sont faites à partir d'un très petit nombre de champs.

Les résultats pour le groupe 2 sont moins homogènes. On observe un très bon accord entre les distributions à l'exception de quelques champs pour lesquels les surfaces sont très grandes (et ceci à tous les seuils).

Concernant le groupe 3, en revanche, les fractions sont mal reproduites par le modèle à tous les seuils. Le modèle sous-estime la taille des surfaces à tous les seuils. Le décalage linéaire du nuage quantile-quantile par rapport à la première bissectrice traduit la présence d'une relation du type  $F(x) = F^*(\alpha x)$  entre les fonctions de répartitions des deux variables. Cette circonstance traduit l'inadaptation du schéma de désagrégation temporel et peut-être aussi partiellement la qualité assez médiocre du modèle au pas de temps de l'événement.

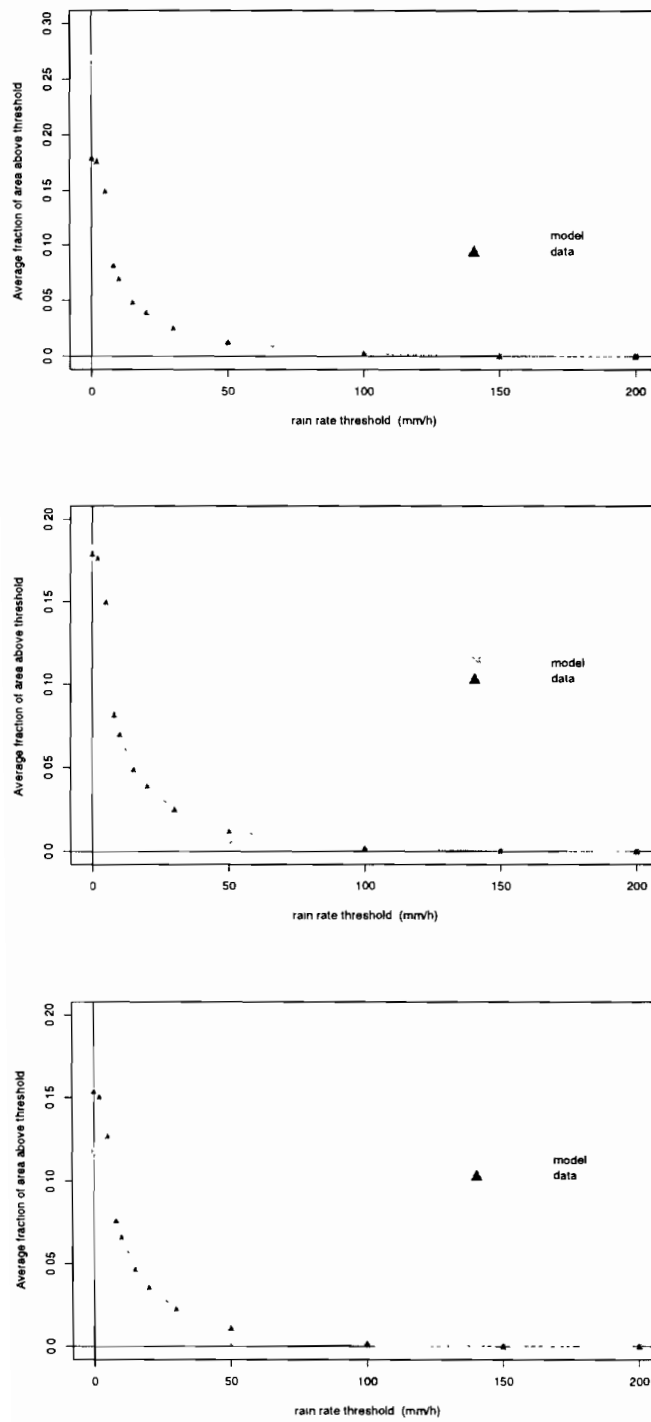


FIG. 7.1: Moyenne des aires des surfaces sup-seuil. De haut en bas : groupe 1, groupe 2, groupe 3

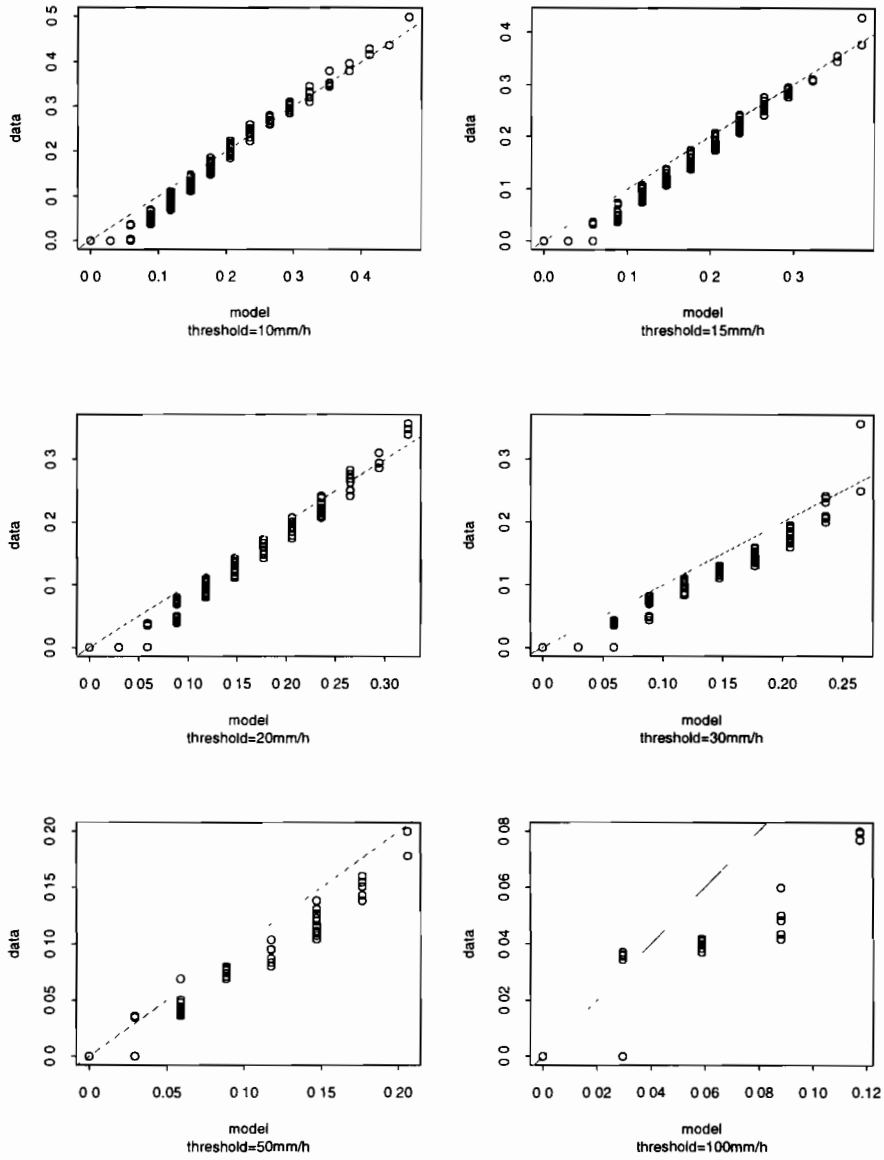


FIG. 7.2: Groupe 1: Graphiques quantile-quantile des fractions des aires des surfaces sup-seuil. (Le caractère discret des valeurs en abscisse provient de ce que le calcul des fractions est réalisé sur un nombre fixe de stations, ce qui n'est pas le cas pour les données à cause des valeurs manquantes)

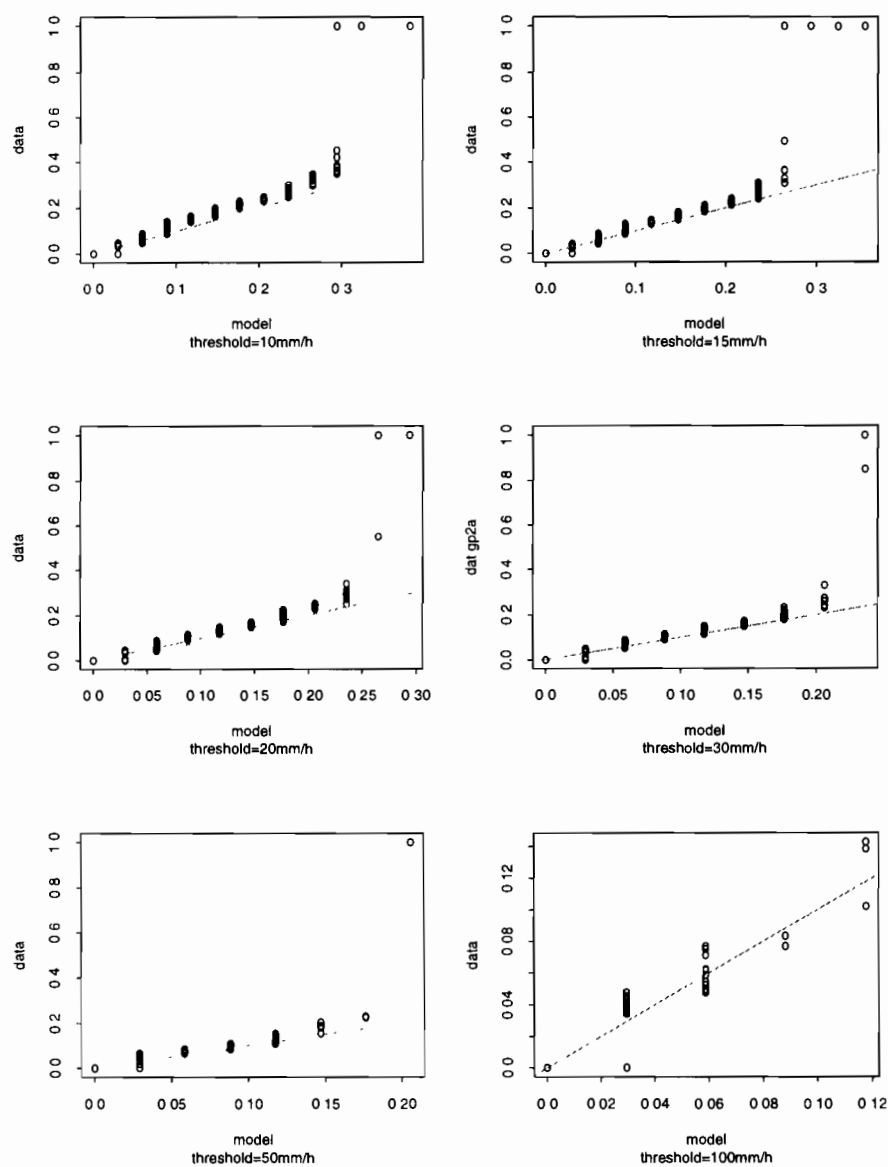


FIG. 7.3: Groupe 2: Graphiques quantile-quantile des fractions des aires des surfaces sup-seuil.

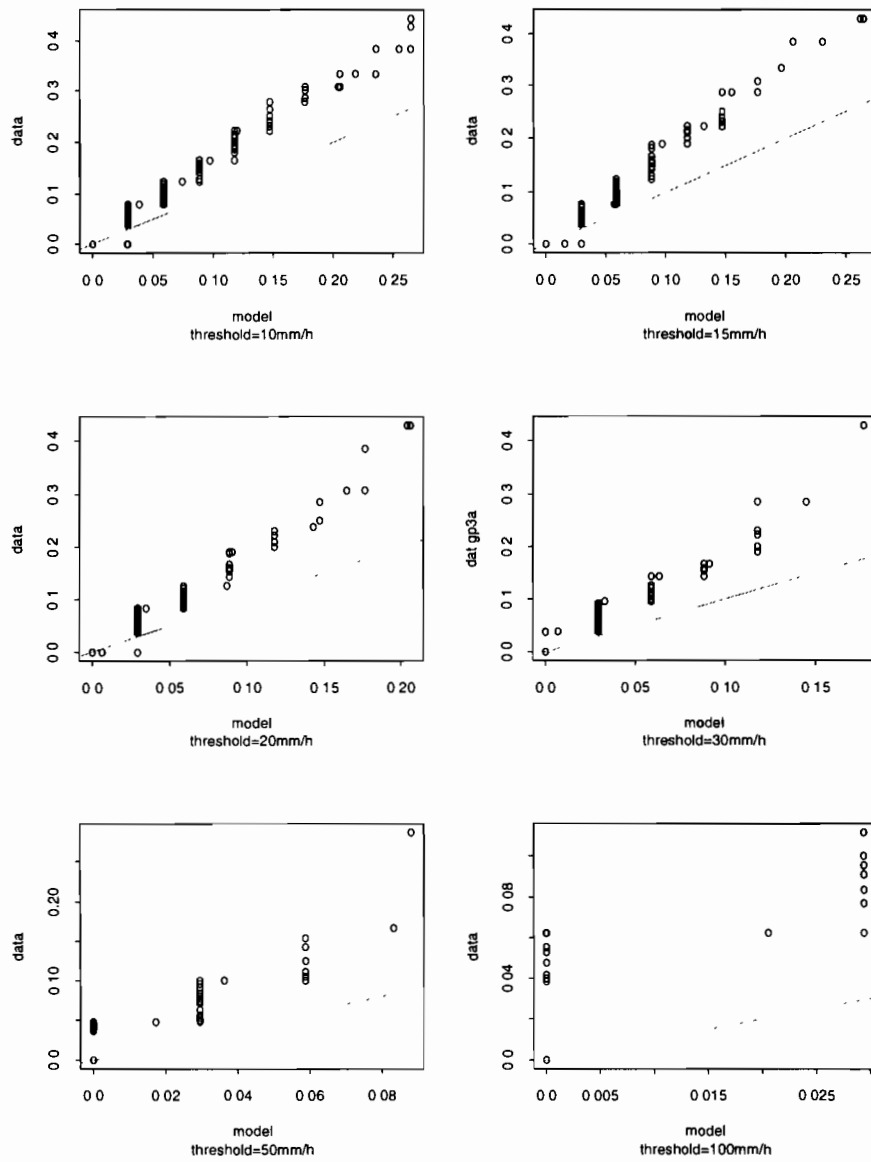


FIG. 7.4: Groupe 3: Graphiques quantile-quantile des fractions des aires des surfaces sup-seuil.



# Conclusion

## Résumé des principaux résultats

Le présent modèle, initialement développé d'après l'observation de lignes de grains sahéliennes, a montré une certaine aptitude à représenter convenablement l'ensemble des SCM de la région.

### Au pas de temps de l'événement

- Ce modèle permet à partir d'un unique jeu de paramètres de reproduire les proportions de chaque groupe d'événements observés, et leurs statistiques ponctuelles d'ordre deux. Le groupe des MCC est en outre bien représenté en terme de distribution monovariante ponctuelle et de structure spatiale. On peut considérer que ces propriétés sont très fortes, compte tenu de l'hétérogénéité de l'échantillon d'événements que l'on a observée au cours de ce travail.
- À partir d'observations ponctuelles, on parvient à rendre compte des fluctuations à l'ordre deux à toutes les échelles (de quelques  $\text{km}^2$  à plus de  $10000 \text{ km}^2$ ).
- Pour une gamme d'échelles plus restreinte, on parvient également à rendre compte de la distribution monovariante complète, en particulier des distributions statistiques spatiales de l'intermittence.

### Aux faibles pas de temps

- Le schéma temporel associé au modèle de champs de pluie au pas de temps de l'événement permet de rendre compte très fidèlement des intensités maximales ponctuelles.
- Concernant les MCC, il permet en outre de rendre compte de la distribution des aires des surfaces sup-seuil.

## Limites du modèle

### Structure du modèle

Un certain nombre de propriétés contenues dans le modèle sont criticables d'un point de vue conceptuel, même si elle n'ont pas fait l'objet d'une invalidation. Par exemple, le fait que les aspects *spatiaux*, *cinématiques* et *temporels* soient traités séparément, par des modèles indépendants au sens probabiliste n'est pas réaliste ; même si on peut penser que la représentation de la cinématique n'est pas un point crucial pour les études hydro-climatiques dans cette région.

### Paramétrisation

Les modèles hydro-météorologiques développés dans les années 80 étaient no-toirement surparamétrés au sens où les grandeurs physiques intervenant dans la paramétrisation n'étaient jamais directement observables (le diamètre et la vitesse d'une cellule convective par exemple), et où les méthodes indirectes d'estimation des paramètres menaient en général à des indéterminations ou à des résultats irréalistes [101, 39]. En conséquence, le présent modèle a été développé avec le souci de conserver un nombre de paramètres restreint et directement observables *à partir des données disponibles*. En dépit de cet effort, l'analyse détaillée des données a montré la nécessité de prendre en compte des propriétés assez fines, comme l'anisotropie des champs au pas de temps de l'événement. Ceci a contribué à une prolifération des paramètres alourdissant le modèle.

### Taille de la fenêtre

La conception du modèle et l'estimation de ses paramètres ayant été réalisées à partir d'observations sur une fenêtre de 140 km × 110 km (réseau complet EPSAT-Niger), le modèle peut-être considéré comme valide (aux limites près mentionnées plus haut) pour la représentation des champs de pluie sur un domaine de taille comparable. Son utilisation pour la désagrégation de champs à des échelles supérieures est sujette à caution :

- Dans une perspective d'étude climatique à partir d'un M.C.G.A, les champs sont donnés sur des mailles de 2° × 2° environ. Dans ces conditions, le gradient pluviométrique du cumul annuel ne serait pris en compte ni par le modèle atmosphérique (travaillant à une échelle trop grossière), ni par le modèle de désagrégation (ajusté sur une fenêtre sur laquelle le gradient n'est pas prédominant).
- Par ailleurs, pour utiliser le modèle sur un plus grand domaine, il conviendrait de prendre explicitement en compte l'extension spatiale limitée des systèmes précipitants. Ceci n'est le cas qu'indirectement dans le présent modèle à travers l'intermittence.

Ces deux points sont liés. Le premier peut être traité en relâchant l'hypothèse de stationnarité par exemple, en spécifiant une anamorphose non uniforme sur le domaine. Le deuxième point est plus délicat, dans la mesure où une prise en compte de la finitude des systèmes nuageux ne peut se faire sans remettre en cause la cohérence du modèle.

### Conditionnement

Enfin, une utilisation intensive du modèle numérique dans sa version actuelle est rendue très lourde par le fait qu'on ne peut pas conditionner directement par une moyenne d'espace.

### Perspectives

Le présent travail n'ouvre pas de perspectives à l'intérieur du formalisme utilisé. En effet ce travail apparaît en définitive comme le développement jusqu'à leurs extrémités, d'idées anciennes [20, 14], ayant été partiellement mises en œuvre dans des contextes divers de l'hydro-météorologie. En quelque sorte, tout a été fait ici, pour reproduire les propriétés des champs de pluie observée, à l'intérieur de ce formalisme. En conséquence, la qualité du modèle ne peut probablement guère être améliorée par une modification des paramètres.

On peut en revanche dégager trois axes de recherche ayant été suggérés par ce travail et s'inscrivant dans sa continuité :

### Comparaison à d'autres modèles, vers une généralisation

Les modèles stochastiques de champs de pluie sur-paramétrisés des années 80 ont laissé place au début des années 90 à des modèles très épurés [53, 52, 130, 103]. Ces derniers, partant de considérations physiques très simples, par analogie avec la physique de la turbulence [11], ont pour objectif de reproduire la loi spatiale des champs, à l'aide d'un très petit nombre de paramètres (2 ou 3 paramètres scalaires) censés caractériser les propriétés d'invariance d'échelle des statistiques observées.

Bien que le présent modèle ne contienne explicitement aucune hypothèse sur l'éventuelle invariance d'échelle des statistiques des champs, on a pu voir à la section 4.3, que la variance semblait présenter de telles propriétés, aussi bien dans le modèle que dans la réalité. Une comparaison détaillée du présent modèle avec les modèles de cascade constituerait une contribution intéressante au débat opposant les modèles gaussiens aux modèles non gaussiens.

### Extension de l'approche à différentes échelles

Ce travail a porté essentiellement sur la modélisation des champs de précipitations pour des échelles allant du  $\text{km}^2$  à un peu plus de  $10000 \text{ km}^2$ , et pour des pas de temps allant de 5 minutes à la durée de l'événement, ceci en mettant

l'accent sur la manière dont les propriétés statistiques se transfèrent d'une échelle à l'autre. Il est souhaitable d'explorer d'autres échelles.

Les modèles physiques de méso-échelle étant actuellement incapables de reproduire correctement la variabilité observée sur des domaines de cette taille, il serait intéressant d'étendre cette approche aux échelles supérieures, à partir des données des réseaux nationaux. Ces données sont disponibles sur la période 1950-1990 et ont l'intérêt de rendre compte de la période sèche et de la période humide. Elles couvrent un domaine de plusieurs centaines de milliers de km<sup>2</sup>, en fait pratiquement toute l'Afrique de l'Ouest.

Ces données pluviométriques sont fournies au pas de temps journalier. Le lien avec les échelles inférieures peut être fait par une approche stochastique, par exemple, à partir d'un modèle de déconvolution, fondé sur la loi des fuites [111, 69], en s'attachant spécifiquement à maintenir une cohérence entre les propriétés modélisées à chaque échelle.

### **Étude de l'impact des fluctuations pluviométriques sur le cycle hydrologique**

La variabilité des systèmes hydrologiques présente des aspects contradictoires selon les échelles auxquelles on s'intéresse. On voit donc l'importance d'être capable de caractériser la variabilité pluviométrique à toutes les échelles, en prenant en compte les effets des fluctuations des caractéristiques externes (cycle de vie, rythme d'occurrence) et internes (intensité de la pluie à différentes échelles de temps, dynamique) des systèmes convectifs.

L'étude de l'impact hydrologique pourrait consister à coupler des scénarios pluviométriques, observés ou possibles, à un modèle hydrologique. Ce modèle ne peut être unique. Sa formulation dépend de l'échelle traitée et de la région concernée. Il est par exemple possible d'utiliser un modèle hydro-dynamique pour étudier la sensibilité du système hydrologique très particulier formé par les mares et l'aquifère du continental terminal 3 de la région de Niamey [31] aux fluctuations pluviométriques observées au cours des 8 années de suivi à long terme d'EPSAT-Niger, d'une part, aux modifications des régimes pluviométriques sur de plus grandes échelles de temps, d'autre part. Dans ce dernier cas, c'est un M.C.G.A qui peut fournir les scénarios climatiques et c'est le présent modèle utilisé en mode conditionnel, qui permettrait de désagréger les sorties pluviométriques du M.C.G.A en entrées pour le modèle hydro-dynamique.

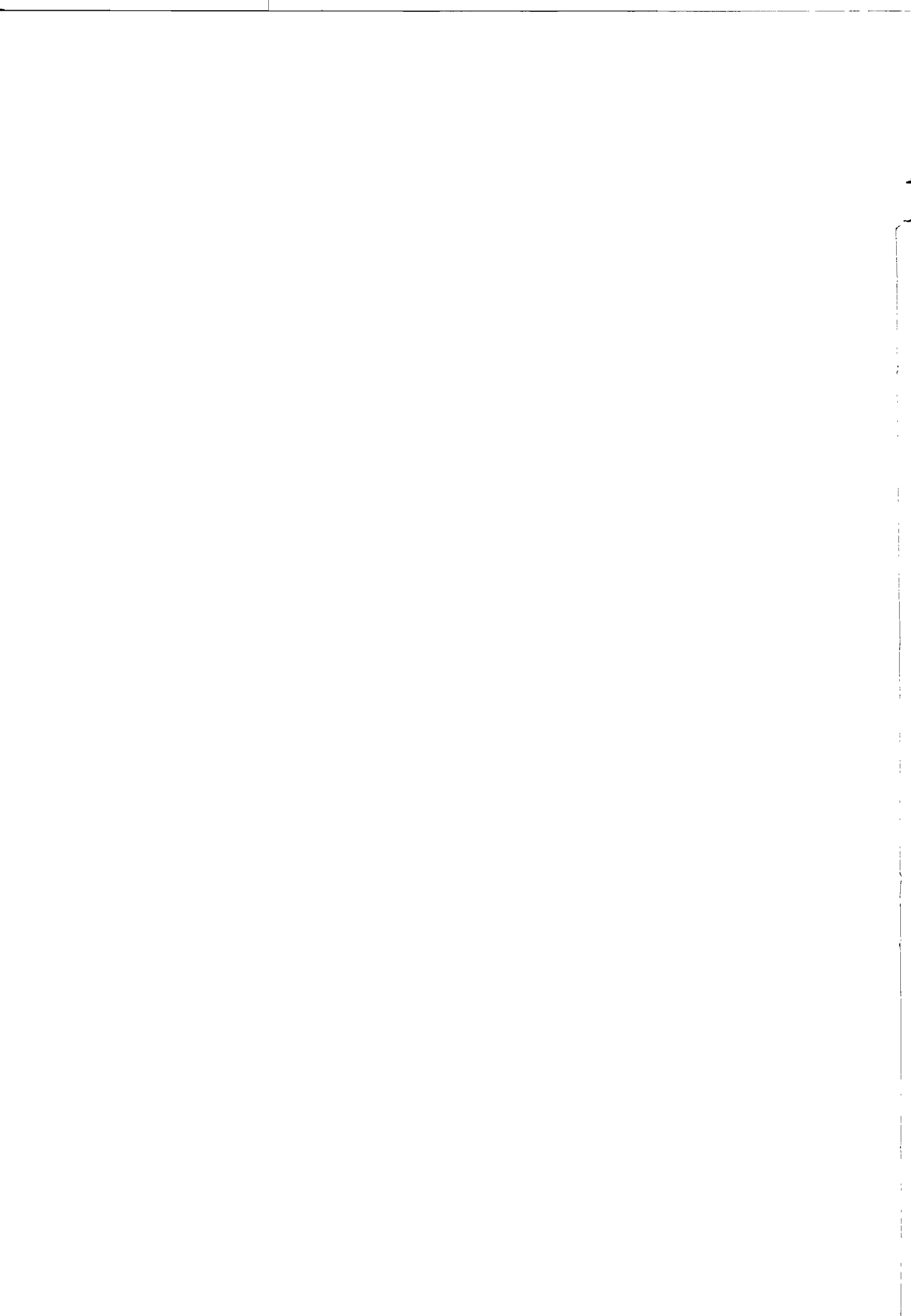
Il serait intéressant de suivre une telle démarche dans le cadre de l'étude qui démarre sur le bassin de l'Ouémé au Bénin [3]. Il faudra cependant attendre pour cela que des données appropriées aient été collectées sur le haut bassin qui couvre 10.000 km<sup>2</sup> environ : données pluviométriques haute résolution du même type que celles d'EPSAT-Niger et données d'écoulement à l'exutoire de plusieurs bassins dont la taille varie de quelques centaines de km<sup>2</sup> à plusieurs milliers de km<sup>2</sup>.

# Index des noms cités

- Abry, P. 26  
Adler, R. 26  
Afouda, A. 120  
Alabert, F. 32  
Albergel, J. 10  
Alé, G. 120  
Allard, D. 74  
Amani, A. 15, 17, 59, 78  
Argoul, F. 119  
Arnaud, Y. 11  
Arneodo, A. 119
- Bacry, E. 119  
Bastin, G. 39  
Bechler, N. 10  
Bell, T. 44, 119  
Bellin, A. 32  
Beven, K. 11, 22  
Blanc-Lapierre, A. 21  
Bondesson, F. 30  
Bonef, E. 13  
Bonifacio, R. 11  
Bouleau, N. 25, 26, 27, 47  
Bras, R. 119  
Braud, I. 17, 80
- Carbonnel, J. 9  
Carmona, R. 26, 87  
Carn, M. 11  
Cârsteanu, A. 20  
Casenave, A. 10  
Céron, J. 9  
Cérou, F. 26, 87  
Cressie, N. 51  
Creutin, J. 17, 39, 80
- D'Amato, N. 14, 15, 44, 59  
De Felice, P. 15  
Delclaux, F. 75  
Depraetere, C. 120
- Desbois, M. 13, 14, 44, 59  
Desconnets, J. 10, 120  
Devroye, L. 33  
Dhonneur, G. 9, 10  
Dietrich, C. 33  
Duqué, C. 39  
Duvel, J. 14, 59
- Eagleson, P. 67  
Elezgaray, J. 119  
Entekhabi, D. 67
- Faure, H. 9  
Faurès, J. 11, 22  
Fontaine, F. 9  
Fortet, R. 21  
Foufoula-Georgiou, E. 13, 20, 118
- Gac, J. 9  
Gasquet, C. 30  
Georgakakos, K. P. 118  
Gervers, M. 39  
Ghikhman, I. 21, 24  
Gnamien, S. 59  
Gneiting, T. 33  
Goodrich, D. 11, 22  
Grimes, D. 11  
Guan, X. 44, 47  
Guillot, B. 13, 44  
Guillot, G. 17, 53  
Gupta, V. 13, 20, 22, 119  
Guttorp, P. 51
- Hoepffner, M. 13, 44  
Holley, R. 119  
Hoopingarner, J. 9  
Hosking, J. 22  
Hu, L. 27  
Hubert, P. 9, 13, 44

- Isham, V. 87  
 Isichenko, M. 26  
  
 Jan, C. 32  
 Janicot, S. 9  
 Jeeyoung, K. 9  
 Jeulin, D. 32  
 Jobard, A. 15, 59  
 Johnson, K. 67  
  
 Kayiranga, T. 59  
 Kedem, B. 44, 47  
 Koroliouk, V. 29, 30, 43  
 Korou Motcho, H. 10  
 Kumar, P. 13  
  
 Lahuec, J. 11  
 Lall, U. 34  
 Lamotte, M. 10  
 Lantuéjoul, C. 26, 53  
 Laroche, C. 13  
 Laurent, H. 11, 14, 15, 59  
 Le Barbé, L. 15, 120  
 Le Barbé, L. 44, 120  
 Lebel, T. 10, 13, 14, 15, 16, 17, 39, 44, 59, 75, 78, 80, 120  
 Lecocq, J. 10, 16  
 Leduc, C. 10, 120  
 Lévy-Vehel, J. 22  
 Loireau, M. 10  
 Lorent, B. 39  
 Lovejoy, S. 119  
  
 Machado, L. 14, 59  
 Mandelbrot, B. B. 26, 94  
 Mantoglou, A. 32  
 Matheron, G. 22, 26, 29, 30, 32, 33, 65  
 Mattioli, J. 29  
 Meyer, Y. 22  
 Mietton, M. 10  
 Monestiez, P. 51  
 Muzy, J. 119  
  
 Nicholson, S. 9  
 Northrop, P. 87  
  
 Obeysekera, J. 118  
 Obled, C. 11, 22, 39  
 O'Connel, P. 22  
 Onof, C. 87  
 Over, T. 20, 119  
  
 Pardo-Igúzquiza, E. 11  
 Pavlopoulos, H. 44, 47  
 Picon, L. 59  
 Polcher, J. 15, 75  
 Portenko, N. 29, 30, 43  
 Pouyaud, B. 10  
 Pronzato, L. 21  
  
 Rasmussen, E. 77  
 Rasmussen, W. 77  
 Redelsperger, J. 10  
 Ribstein, P. 10, 120  
 Rivoirard, J. 30  
 Robin, J. 13  
 Rodriguez-Iturbe, I. 119  
 Rousselle, J. 15, 59  
 Roux, F. 10  
 Rubbin, Y. 32  
 Rudin, W. 19  
 Rutledge, S. 77  
  
 Salas, J. 118  
 Sampson, D. 51  
 Saporta, G. 43  
 Sauquet, E. 17  
 Sauvageot, H. 13, 44  
 Schaake, J. 12  
 Schertzer, D. 119  
 Schmitt, M. 29  
 Shah, S. 22  
 Sharma, A. 34  
 Shinozuka, M. 32  
 Short, D. 44, 47  
 Sircoulon, J. 9  
 Skorokhod, A. 21, 24, 29, 30, 43  
 Sorooshian, S. 11, 22  
 Switzer, S. 51

Tabios III, G. 118  
Tarboton, D. 34  
Taupin, J. 10, 11, 13, 15, 59, 120  
Teissier, Y. 119  
Thauvin, V. 70  
Toma, A. 15, 59  
Torresani, B. 32  
Turbine, A. 29, 30, 43  
  
Valencia, R. 12  
Valentin, C. 10  
Viltard, A. 15  
von Storch, H. 12  
  
Wackernagel, H. 51  
Wallis, J. 94  
Walter, E. 21  
Waymire, E. 13, 20, 119  
Weather, H. 87  
Wendling, J. 11, 22  
Wilson, J. 32  
Witowski, P. 30  
Woolhiser, D. 11, 22  
  
Zorita, E. 12



# Bibliographie

- [1] P. ABRY. *Ondelettes et turbulences. Multirésolutions, algorithmes de décomposition, invariance d'échelle et signaux de pression*. Diderot, Paris, 1997.
- [2] R.J. ADLER. *The geometry of random fields*. Series in Probability and Mathematical Statistics. Wiley, 1984.
- [3] A. AFOUDA, G. ALÉ, C. DEPRAETERE, L. LE BARBÉ, et T. LEBEL. « Impact de la variabilité climatique sur les ressources en eau en Afrique de l'Ouest, un défi pour l'hydrologie tropicale ». Dans *Proceedings of the AISH conference on Water Ressources Variability in Africa during the 20-th century*. Abidjan, Côte d'Ivoire, 1998.
- [4] F. ALABERT. « The practice of fast conditional simulations through the LU decomposition of the covariance matrix ». *Mathematical Geology*, 19(5):369–385, 1987.
- [5] J. ALBERGEL, A. CASENAVE, P. RIBSTEIN, et C. VALENTIN. « *L'aridité : une contrainte au développement* », Chapitre Aridité climatique, aridité édaphique : étude des conditions de l'infiltrabilité en Afrique tropicale sèche, pages 123–130. Didactiques. ORSTOM, 1992.
- [6] D. ALLARD. « Connexité des ensembles aléatoires. Application à la simulation de réservoirs pétroliers hétérogènes. ». Thèse de Doctorat de l'École des Mines de Paris, 1993.
- [7] A. AMANI. « *Estimation des pluies au Sahel : quelques problèmes de modélisation analysés à partir des données EPSAT-NIGER* ». PhD thesis, École Polytechnique, Université de Montréal, 1995.
- [8] A. AMANI et T. LEBEL. « Lagrangian kriging for the estimation of sahelian rainfall at small time steps ». *Journal of Hydrology*, 192:125–157, 1997.
- [9] A. AMANI, T. LEBEL, J. ROUSSELLE, et J.D. TAUPIN. « Typology of rainfall fields to improve rainfall estimation in the Sahel by the area threshold method ». *Water Ressources Research*, 232(8):2473–2487, 1996.

- [10] Y. ARNAUD, J.D. TAUPIN, et H. LAURENT. « *Validation problems of rainfall estimation methods by satellite in tropical Africa* », Chapitre Validation de deux méthodes d'estimation des pluies par satellite à l'aide du réseau dense de pluviographes de l'expérience EPSAT-Niger. ORSTOM, Paris, 1996.
- [11] A. ARNEODO, F. ARGOUL, E. BACRY, J. ELEZGARAY, et J.F. MUZY. *Ondes-lettes, multifractales et turbulences. De l'ADN aux croissances cristallines*. Diderot, 1995.
- [12] G. BASTIN, B. LORENT, C. DUQUÉ, et M. GERVERS. « Optimal estimation of the average areal rainfall and optimal selection of rain gauge locations ». *Water Resources Research*, 20(4):463-470, 1984.
- [13] N. BECHLER, M. LAMOTTE, H. KOROU MOTCHO, et M. MIETTON. « Les étiages du fleuve Niger : risque de pénurie en eau pour l'agglomération de Niamey ». Dans IAHS, éditeur, *Actes du symposium de Rabat : Sustainability of Water Resources under Increasing Uncertainty*, volume 240, pages 67-74, 1997.
- [14] T.L. BELL. « A space-time stochastic model of rainfall for satellite remote-sensing studies ». *Journal of Geophysical Research*, 92(8):9631-9643, 1987.
- [15] A. BELLIN et Y. RUBBIN. « HYDRO-GEN: A spatially distributed random field generator for correlated properties ». *Stochastic Hydrology and Hydraulics*, 10:253-278, 1996.
- [16] A. BLANC-LAPIERRE et R. FORTET. *Théorie des fonctions aléatoires*. Masson, Paris, 1953.
- [17] F. BONDESSON. « A general result on infinite divisibility ». *Journal of Applied Probability*, 7(6):965-979, 1979.
- [18] N. BOULEAU. *Processus stochastiques et applications*. Hermann, Paris, 1988.
- [19] N. BOULEAU. « Splendeurs et misères des lois de valeurs extrêmes ». Centre de Mathématiques Appliquées de l'École Nationale des Ponts et Chaussées., 1998.
- [20] R.L. BRAS et I. RODRIGUEZ-ITURBE. « Rainfall generation: a non stationary time-varying multidimensional model ». *Water Resources Research*, 12(3):450-456, 1976.
- [21] R. CARMONA et F. CÉROU. « Transport by incompressible random velocity fields: simulations and mathematical conjectures in stochastic partial differential equations ». *Stochastic Partial Differential Equations preprint*, 1998.

- [22] M. CARN et J.P. LAHUEC. « Estimation des précipitations au Niger au cours de la saison des pluies 1986 à partir de l'imagerie IR Météosat. Bilan et critique des méthodes utilisées ». *Veille Climatique Satellitaire*, 17:40-48, 1987.
- [23] A. CÂRSTEANU et E. FOUFOULA-GEORGIU. « Assessing dependance among weights in a multiplicative cascade model of temporal rainfall ». *Journal of Geophysical Research*, 101(D21), 1996.
- [24] A. CÂRSTEANU et E. FOUFOULA-GEORGIU. « Non trivial scaling in the loss of prediction information with aggregation in hourly precipitation occurrences ». *Journal of Geophysical Research*, 102(D6):6631-6636, 1997.
- [25] Centre National de la Recherche Météorologique. *Désagrégation spatio-temporelle de la pluie en région sahélienne : modélisation, algorithme et application*, novembre 1997.
- [26] J.P. CÉRON. « L'influence des conditions météorologiques dans l'hémisphère Nord sur la pluie au Sahel. ». *La météorologie*, 8(2):54-61, 1993.
- [27] N.A.C CRESSIE. *Statistics for spatial data*. Probability and Mathematical statistics. Wiley, 1994.
- [28] N. D'AMATO et T. LEBEL. « On the characteristics of rainfall events in the Sahel, with a view to the analysis of climatic variability ». *International Journal of Climatology*, 18:955-974, 1998.
- [29] F. DELCLAUX, J. POLCHER, et T. LEBEL. « Rainfall event climatology in West-Africa : comparison from in situ and G.C.M outputs ». Dans *Proceedings of the 23rd general assembly of the European Geophysical Society*, Nice, mai 1998.
- [30] M. DESBOIS, T. KAYIRANGA, S. GNAMIEN, et L. PICON. « Characterization of some elements of the sahelian climate and their interannual variations for july 1983, 1984 and 1985 from the analysis of METEOSAT ISCCP data ». *Journal of Climate*, 1(9):867-904, 1988.
- [31] J.C. DESCONNETS, J.D. TAUPIN, T. LEBEL, et C. LEDUC. « Hydrology of the HAPEX-Sahel Central Super-Site : surface water drainage and aquifer recharge through the pool systems ». *Journal of Hydrology*, 188-199:155-178, 1997.
- [32] L. DEVROYE. *Non uniform random variate generation*. Springer-Verlag, New-York, 1994.
- [33] G. DHONNEUR. « L'équateur météorologique ». *La météorologie*, 6(6):113-122, septembre 1976.

- [34] G. DHONNEUR. « Les amas nuageux mobiles: principale composante de la météorologie du Sahel ». *La météorologie*, 4(27), décembre 1981.
- [35] C.R. DIETRICH. « A simple and efficient space domain implementation of the turning bands method ». *Water Ressources Research*, 31(1):147–156, 1995.
- [36] C.R. DIETRICH. « Reply to T. Gneiting ». *Water Ressources Research*, 32(11):3391–3396, 1996.
- [37] H. FAURE et J.Y. GAC. « Will the sahelian drought end in 1985? ». *Nature*, 291:475–478, 1981.
- [38] J.M. FAURÈS, D.C. GOODRICH, D.A. WOOLHISER, et S. SOROOSHIAN. « Impact of small-scale spatial rainfall variability on runoff modeling ». *Journal of Hydrology*, 173:309–326, 1995.
- [39] E. FOUFOULA-GEORGIU et K. P. GEORGAKAKOS. « *Recent Advances in the modeling of hydrologic systems* », Chapitre 3, pages 47–65. Kluwer Academic Publishers, Netherlands, 1991.
- [40] C. GASQUET et P. WITOMSKI. *Analyse de Fourier et applications. Filtrage, calcul numérique, ondelettes*. Masson, Paris, août 1990.
- [41] I. GHIKHMAN et A. SKOROKHOD. *Theory of stochastic processes*, volume 1-3. Springer-Verlag, 1974.
- [42] I. GHIKHMAN et A. SKOROKHOD. *Introduction à la théorie des processus aléatoires*. Mir, 1980.
- [43] T. GNEITING. « Comment on "A simple and efficient space domain implementation of the turning bands method" by C.R. Dietrich ». *Water Ressources Research*, 32(11):3391–3396, 1996.
- [44] D.I.F. GRIMES, E. PARDO-IGÚZQUIZA, et R. BONIFACIO. « Optimal areal rainfall estimation raingauges and satellite data ». *Journal of Hydrology (preprint)*, 1999.
- [45] G. GUILLOT. « Éléments pour l'estimation des lignes de grain d'Afrique de l'Ouest aux faibles pas de temps ». Mémoire de D.E.A de Géostatistique de l'École des Mines de Paris, juin 1995.
- [46] G. GUILLOT. « Approximation of sahelian rainfall fields with meta-gaussian random functions 1: model definition and methodology ». *Stochastic Hydrology and Hydraulics*, janvier 1999.
- [47] G. GUILLOT, A. AMANI, T. LEBEL, et J.D. CREUTIN. « Rainfall statistics at small time steps in the Sahel: a preliminary investigation into the validation of a disaggregation model ». Dans *Annales Geophysicae - Proceedings of*

*the 23rd general assembly of the European Geophysical Society, Hydrology, Oceans, Atmosphere and non linear Geophysics, Nice, mai 1998. European Geophysical Society.*

- [48] G. GUILLOT et T. LEBEL. « Disaggregation of sahelian meso-scale convective system rainfields : further developments and validation ». *submitted to Journal of Geophysical Research*, 1998.
- [49] G. GUILLOT et T. LEBEL. « A review of small scale properties of sahelian rainfields : analysis and downscale modelling ». Dans *Proceedings of the Sixth International Conference on Precipitations*. AGU-AMS, juin 1998.
- [50] G. GUILLOT et T. LEBEL. « Approximation of sahelian rainfall fields with meta-gaussian random functions 2 : parameter estimation and comparison to data ». *Stochastic Hydrology and Hydraulics*, janvier 1999.
- [51] V.K. GUPTA. « Lectures notes on scale invariance and its applications in hydrologic science ». Graduate program in Hydrology, University of Colorado, Boulder, Colorado, 1996.
- [52] V.K. GUPTA et E.C. WAYMIRE. « A statistical analysis of mesoscale rainfall as a random cascade ». *Journal of Applied Meteorology*, 32:251-267, 1993.
- [53] R. HOLLEY et E.C. WAYMIRE. « Multifractal dimensions and scaling exponents for strongly bounded random cascades ». *Annales of Applied Probability*, 2:819-845, 1992.
- [54] L.Y. HU. « Mise en oeuvre du modèle gamma pour l'estimation des distributions spatiales ». Thèse de Doctorat de l'École des Mines de Paris, 1988.
- [55] P. HUBERT et J.P. CARBONNEL. « Approche statistique de l'aridification de l'Afrique de l'Ouest ». *Journal of Hydrology*, 110:349-367, 1988.
- [56] M.B ISICHENKO. « Percolation, statistical topography and transport in random media ». *Revue of Modern Physics*, 64:961-1043, 1992.
- [57] S. JANICOT. « Spatio-temporal variability of west african rainfall. Part 1 : regionalization and typings. ». *Journal of Climate*, 5(5):489-497, 1992.
- [58] S. JANICOT. « Spatio-temporal variability of west african rainfall. Part 2 : Associated surface and air mass characteristics ». *Journal of Climate*, 5(5):499-511, 1992.
- [59] S. JANICOT et F. FONTAINE. « L'évolution des idées sur la variabilité inter-annuelle récente des précipitations en Afrique de l'Ouest ». *La météorologie*, 8(1):25-53, 1993.

- [60] D. JEULIN. « Modèles de structures aléatoires ». Note du Centre de Géostatistique de l'École des Mines de Paris, mars 1994.
- [61] K.D. JOHNSON, D. ENTEKHABI, et P.S. EAGLESON. « The implementation and validation of improved land surface hydrology in an atmospheric general circulation model ». Rapport Technique 334, Ralph M. Parsons Laboratory, Massachusetts Institute of Technology, 1991.
- [62] B. KEDEM, H. PAVLOPOULOS, X. GUAN, et D.A. SHORT. « A probability distribution model for rain rate ». *Journal of Applied Meteorology*, 33:1486–1493, 1994.
- [63] V. KOROLIOUK, N. PORTENKO, A. SKOROKHOD, et A. TOURBINE. *Aide-mémoire de probabilités et de statistique mathématique*. Mir, Moscou, 1983.
- [64] P. KUMAR et E. FOUFOULA-GEORGIU. « A new look at rainfall fluctuations and scaling properties of spatial rainfall using orthogonal wavelets ». *Journal of Applied Meteorology*, 32:209–221, 1993.
- [65] C. LANTUÉJOUL. « Ergodicity and integral range ». *Journal of Microscopy*, 161(3):387–404, 1991.
- [66] C. LANTUÉJOUL. « Non conditional simulation of stationary isotropic multi-gaussian random functions ». Dans P.A. DOWD et M. ARMSTRONG, éditeurs, *Geostatistical Simulations*, pages 147–177, Dordrecht, 1994. Kluwer Academic Publishers.
- [67] H. LAURENT, N. D'AMATO, et T. LEBEL. « How important is the contribution of the mesoscale convective complexes to the sahelian rainfall? ». *Physics and Chemistry of the Earth*, 1997.
- [68] H. LAURENT, A. JOBARD, et A. TOMA. « Validation of the satellite and ground based estimates of precipitation over the Sahel ». *Atmospheric Research*, 47-48:651–670, 1998.
- [69] L. LE BARBÉ et T. LEBEL. « Rainfall monitoring of the HAPEX-Sahel region during the years 1950-1990 ». *Journal of Hydrology*, 188-189:43–73, 1997.
- [70] T. LEBEL. « Rapport de campagne EPSAT-Niger 1991 ». Rapport technique de l'ORSTOM, 1991.
- [71] T. LEBEL, G. BASTIN, C. OBLED, et J.D. CREUTIN. « On the accuracy of areal rainfall estimation: A case study ». *Water Resources Research*, 23(11):2123–2134, 1987.
- [72] T. LEBEL, I. BRAUD, et J.D. CREUTIN. « A space-time rainfall disaggregation model adapted to sahelian meso-scale convective complexes ». *Water Resources Research*, 7(34):1711–1726, 1998.

- [73] T. LEBEL, L. LE BARBÉ, et N. D'AMATO. « Rainfall monitoring during HAPEX-Sahel: 1. General conditions and climatology ». *Journal of Hydrology*, 188-189:74–96, 1997. Hapex-Sahel special issue.
- [74] T. LEBEL, H. SAUVAGEOT, M. HOEPPFNER, M. DESBOIS, B. GUILLOT, et P. HUBERT. « Estimation des Précipitations au Sahel: l'expérience Epsat-Niger ». *Hydrologie Continentale*, 2:133–143, 1991.
- [75] J. LECOCQ. « Validation d'un radar météorologique bande-C pour l'étude des précipitations sahéniennes ». Thèse de doctorat de l'Université Montpellier 2, 1997.
- [76] C. LEDUC et M. LOIREAU. « Fluctuations piézométriques et évolution du couvert végétal en zone sahénienne (sud-ouest du Niger) ». Dans IAHS, éditeur, *Actes du symposium de Rabat: Sustainability of Water Resources under Increasing Uncertainty*, volume 240, 1997.
- [77] J. LÉVY-VEHEL. « Régularité ponctuelle, analyse multifractale et traitements des signaux ». Cours de l'école d'été du C.N.R.S: Méthodes d'ondelettes pour l'analyse de fonctions fractales et multifractales, E.N.S Cachan, octobre 1997.
- [78] L.A. MACHADO, M. DESBOIS, et J.P. DUVEL. « Structural characteristics of deep convective systems over tropical Africa ». *Monthly Weather Review*, 120:392–406, 1992.
- [79] B. B. MANDELBROT. *Les objets fractals*. Champs. Flammarion, Paris, 1995.
- [80] B. B. MANDELBROT. *Fractales, hasard et finance*. Champs. Flammarion, Paris, 1998.
- [81] B. B. MANDELBROT et J.R. WALLIS. « Noah, Joseph, and operational hydrology ». *Water Resources Research*, 4(5):909–918, 1968.
- [82] B. B. MANDELBROT et J.R. WALLIS. « Computer experiments with fractional Gaussian noises, 1. Averages and variances ». *Water Resources Research*, 5(1):228–241, 1969.
- [83] B. B. MANDELBROT et J.R. WALLIS. « Computer experiments with fractional Gaussian noises, 2. Rescaled ranges and spectra ». *Water Resources Research*, 5(1):242–259, 1969.
- [84] A. MANTOGLIOU et J. WILSON. « The turning bands method for simulation of random fields using line generation by a spectral method ». *Water Resources Research*, 18(5):1379–1394, 1982.
- [85] G. MATHERON. *Les variables régionalisées et leur estimation*. Masson, Paris, 1965.

- [86] G. MATHERON. « La théorie des variables régionalisées, et ses applications ». Cahier du Centre de Morphologie Mathématique de l'École des Mines de Paris, 1970. Fascicule 5.
- [87] G. MATHERON. « Leçons sur les fonctions aléatoires d'ordre deux ». Cours du Centre de Géostatistique de l'École des Mines de Paris, 1972.
- [88] G. MATHERON. « The intrinsic random functions and their applications ». *Advances in Applied Probability*, 5:439–468, 1973.
- [89] G. MATHERON. « L'estimation globale des réserves récupérables ». Note du Centre de Géostatistique de l'École des Mines de Paris, septembre 1978.
- [90] G. MATHERON. « La destructuration des hautes teneurs et le krigeage des indicatrices ». Note du centre de Géostatistique N-761, École des Mines de Paris, juin 1982.
- [91] G. MATHERON. « Suffit-il pour une covariance d'être de type positif? ». Note du Centre de Géostatistique de l'École des Mines de Paris, avril 1986.
- [92] G. MATHERON. « The internal consistency of models in geostatistics ». Dans P.A. DOWD et M. ARMSTRONG, éditeurs, *Geostatistics*, volume 1, pages 21–38, Dordrecht, 1989. Kluwer Academic Publishers.
- [93] G. MATHERON. « Une conjecture sur la covariance d'un ensemble aléatoire ». Dans *Cahiers de Géostatistique*, volume 3, pages 107–113, Fontainebleau, 1993. Centre de Géostatistique de l'École des Mines de Paris.
- [94] Y. MEYER. « Exposants de Hölder ponctuels et analyse multifractale ». Cours de l'école d'été du C.N.R.S : Méthodes d'ondelettes pour l'analyse de fonctions fractales et multifractales, E.N.S Cachan, octobre 1997.
- [95] P. MONESTIEZ et S. SWITZER. « Semiparametric estimation of nonstationary spatial covariance models by multidimensional scaling ». Rapport Technique 165, Department of Statistics, Department of Applied Earth Science, Stanford University, California, novembre 1991.
- [96] S.E. NICHOLSON. « The nature of rainfall fluctuations in Subtropical West Africa ». *Monthly Weather Review*, 108:473–487, 1980.
- [97] S.E. NICHOLSON. « Rainfall and atmospheric circulation during drought periods and wetter years in West Africa ». *Monthly Weather Review*, 109:2191–2208, 1981.
- [98] S.E. NICHOLSON. « The spatial coherence of african rainfall anomalies : inter-hemispheric teleconnections ». *Journal of Climate and Applied Meteorology*, 25:1365–1381, 1986.

- [99] S.E. NICHOLSON, K. JEEYOUNG, et J. HOOPINGARNER. « Atlas of african rainfall and its interannual variability. », 1988.
- [100] P. NORTHROP, C. ONOF, V. ISHAM, et H. WEATHER. « A Poisson cluster spatio-temporal rainfall model ». Dans *Proceedings of the Sixth international Conference on Precipitations, Hawaii*, 1998.
- [101] J.T.B. OBEYSEKERA, G.Q. TABIOS III, et J.D. SALAS. « On parameter estimation of temporal rainfall models ». *Water Ressources Research*, 23(10):1837–1850, 1987.
- [102] C. OBLED, J. WENDLING, et K. BEVEN. « The sensitivity of hydrological models to spatial rainfall patterns: an evaluation using observed data ». *Journal of Hydrology*, 159:305–333, 1994.
- [103] T.M. OVER et V.K. GUPTA. « Statistical analysis of mesoscale rainfall: dependence of a random cascade generator on large-scale forcing ». *Journal of Applied Meteorology*, 33:1526–1542, 1994.
- [104] J. POLCHER. « Sensitivity of tropical convection to land surface processes ». *Journal of Atmospheric Sciences*, 52:3143–3161, 1995.
- [105] B. POUYAUD. « Variabilité spatiale et temporelle des bilans hydriques de quelques bassins versants d’Afrique de l’Ouest en liaison avec les changements climatiques ». Dans IAHS, éditeur, *Actes du symposium de Vancouver: Climate change and climatic variability*, volume 168, pages 447–461, 1987.
- [106] E.N. RASMUSSEN et S.A. RUTLEDGE. « Evolution of a quasi-two-dimensional squall lines part 1: kinematic and reflectivity structure ». *Journal of Atmospheric Sciences*, 50(16):2584–2606, 1993.
- [107] W.O. RASMUSSEN. « Infiltration-advance for radial spreading ». *Water Ressources Research*, 30(4):929–937, avril 1994.
- [108] W.O. RASMUSSEN. « Computation and plotting of infiltration-advance isochrones for any source ». *Advances in Water Ressources*, 17:353–362, 1995.
- [109] W.O. RASMUSSEN. « Stochastic sketching of infiltration-advance isochrones ». *Stochastic Hydrology and Hydraulics*, 1995.
- [110] J.L. REDELSPERGER. « *Les climats subtropicaux et leur évolution: de l’observation spatiale à la modélisation* », Chapitre Les lignes de grain tropicales, pages 233–259. Cépaduès, 1993.
- [111] P. RIBSTEIN. « Loi des fuites ». *Cahier ORSTOM*, 20(2):117–145, 1983.
- [112] J. RIVOIRARD. *Introduction to disjunctive kriging and non-linear geostatistics*. Oxford University Press, Oxford, 1994.

- [113] F. ROUX. *Les orages : météorologie des grains, de la grêle et des éclairs*. Payot, 362 pages, 1991.
- [114] W. RUDIN. *Analyse réelle et complexe*. Masson, 1992.
- [115] D. SAMPSON et P. GUTTORP. « Nonparametric estimation of nonstationary spatial covariance function ». *Journal of the American Statistical Association*, 87(147):108–119, mars 1992.
- [116] G. SAPORTA. *Probabilités, analyse des données et statistique*. Technip, 1990.
- [117] E. SAUQUET. « Analyse statistique d'un modèle de désagrégation spatio-temporelle de la pluie adapté aux lignes de grain sahéliennes ». Mémoire de D.E.A de Mécanique des Milieux Géophysiques et Environnement, septembre 1995.
- [118] E. SAUQUET. « Construction d'un modèle de désagrégation spatio-temporelle de la pluie appliqué aux lignes de grains sahéliennes ». Projet de fin d'étude de l'École Nationale Supérieure d'Hydraulique et de Mécanique de Grenoble, juillet 1995.
- [119] M. SCHMITT et J. MATTIOLI. *Morphologie mathématique*. Collection logique, mathématiques, informatique. Masson, Paris, 1993.
- [120] S.M.S SHAH, P.E. O'CONNEL, et J.M.R. HOSKING. « Modelling the effects of spatial variability in rainfall on catchment response, 1. Formulation and calibration of a stochastic rainfall field model ». *Journal of Hydrology*, 175:67–88, 1996.
- [121] M. SHINOZUKA. « Simulation of multivariate and multidimensional random processes ». *Journal of the Acoustical Society of America*, 17(2):357–367, 1970.
- [122] M. SHINOZUKA et C.M. JAN. « Digital simulation of random processes and its applications ». *Journal of Sound and Vibration*, 25(1):11–128, 1972.
- [123] J. SIRCOULON. « Les données hydro-pluviométriques de la sécheresse récente en Afrique inter-tropicale. Comparaison avec les sécheresses de 1913 et 1940. ». *Cahier ORSTOM, hydrol.*, 13(2), 1976.
- [124] J. SIRCOULON. « Climate, water and development ». Dans *Proceedings of the Second World Climate Conference*, page 582 pages. Cambridge University Press, 1991.
- [125] D.G. TARBOTON, A. SHARMA, et U. LALL. « Disaggregation procedures for stochastic hydrology based on non parametric density estimation ». *Water Resources Research*, 34(1):107–119, janvier 1998.

- [126] J.D. TAUPIN, E. BONEF, et J. ROBIN. « Rapport de campagne EPSAT-Niger 1994 ». Rapport technique de l'ORSTOM, 1994.
- [127] J.D. TAUPIN, C. LAROCHE, et J. ROBIN. « Rapport de campagne EPSAT-Niger 1995 ». Rapport technique de l'ORSTOM, 1995.
- [128] J.D. TAUPIN et T. LEBEL. « Rapport de campagne EPSAT-Niger 1992 ». Rapport technique de l'ORSTOM, avril 1992.
- [129] J.D. TAUPIN et T. LEBEL. « Rapport de campagne EPSAT-Niger 1993 ». Rapport technique de l'ORSTOM, 1993.
- [130] Y. TEISSIER, S. LOVEJOY, et D. SCHERTZER. « Universal multifractals: theory and observations from rain and clouds ». *Journal of Applied Meteorology*, 32:223–250, 1993.
- [131] V. THAUVIN. « Étude de la répartition spatiale des précipitations en milieu sahélien à l'aide du réseau dense de pluviographes EPSAT-Niger ». Thèse de doctorat de l'Université Montpellier 2, 1992.
- [132] B. TORRESANI. « Analyse de processus localement stationnaires ». Notes de cours du D.E.A de Physique Théorique de l'Université Aix-Marseille 1, janvier 1997.
- [133] R. VALENCIA et J.C. SCHAAKE. « Disaggregation processes in stochastic hydrology ». *Water Resources Research*, 9(3):580–585, 1973.
- [134] A.H. VILTARD, H. LAURENT, et P. De FELICE. « Lower tropospheric cyclonic vortices in Northern Africa in summer ». *Monthly Weather Review*, 118:818–823, 1990.
- [135] H. WACKERNAGEL. *Multivariate geostatistics: an introduction with applications*. Springer-Verlag, Berlin, 1995.
- [136] E. WALTER et L. PRONZATO. *Identification de modèles paramétriques à partir de données*. Masson, Paris, 1994.
- [137] D.A. WOOLHISER et D.C. GOODRICH. « Effects of storm rainfall intensity patterns on surface runoff ». *Water Resources Research*, 102:335–354, 1988.
- [138] E. ZORITA et H. von STORCH. « A survey of statistical downscaling ». *Journal of Climate*, 1998.

## THÈSE DE DOCTORAT

**Titre.** MODÉLISATION STATISTIQUE DES CHAMPS DE PLUIE SAHÉLIENS APPLICATION À LEUR DÉSAGRÉGATION SPATIALE ET TEMPORELLE

**Auteur :** Gilles Guillot

**Établissement :** Laboratoire d'étude des Transferts en Hydrologie et Environnement

### Résumé :

La résolution couramment atteinte par les données pluviométriques dérivées des mesures satellitaires ou des sorties de Modèles de Circulation Générale Atmosphérique (M C G A) sur la zone sahélienne, est de l'ordre de  $1^{\circ} \times 1^{\circ}$ , sur des pas de temps correspondant à la durée de vie des systèmes convectifs. Ces échelles sont incompatibles avec les échelles de la modélisation hydrologique, particulièrement dans cette région, où les pluies principalement d'origine convective, présentent une forte variabilité spatiale et temporelle.

Le travail présenté a consisté à développer et valider un modèle, permettant d'utiliser l'information de grande échelle fournie par un satellite ou un M C G A, pour simuler de champs de pluie à une plus haute résolution.

Les bases de ce modèle proviennent de l'analyse du jeu de données de l'expérience EPSAT-Niger. Il s'appuie sur trois composantes : (i) une représentation géostatistique des champs spatiaux au pas de temps de l'événement, (ii) un modèle de la cinématique des systèmes convectifs, (iii) un modèle statistique à quatre paramètres des hyétogrammes.

Ce modèle est développé dans un cadre mathématique qui maintient une cohérence entre les propriétés observées à différentes échelles. De ce fait, il peut être utilisé pour réaliser des simulations conditionnées par des valeurs pluviométriques de grande échelle.

### Mots-clés :

précipitations, Sahel, géostatistique, systèmes convectifs, modèle aléatoire, désagrégation, changement d'échelle, couplage modèle hydrologique/modèle atmosphérique