

**ESTIMATION DES PLUIES AU SAHEL:
QUELQUES PROBLÈMES DE MODÉLISATION
ANALYSÉS À PARTIR DES DONNÉES DE
L'EXPÉRIENCE EPSAT-NIGER**

**Thèse de Philosophiae Doctor (Ph.D.)
Département de Génie civil**

**Abou AMANI
Juillet 1995**




**ÉCOLE
POLYTECHNIQUE
MONTRÉAL**
*Le Savoir
sans frontières*

AFFILIÉE À L'UNIVERSITÉ
DE MONTRÉAL

UNIVERSITÉ DE MONTRÉAL
ECOLE POLYTECHNIQUE DE MONTRÉAL

Cette thèse intitulée:

**ESTIMATION DES PLUIES AU SAHEL: QUELQUES PROBLÈMES DE
MODÉLISATION ANALYSÉS À PARTIR DES DONNÉES DE
L'EXPÉRIENCE EPSAT-NIGER**

présentée par: Abou AMANI

en vue de l'obtention du diplôme de: Philosophiae Doctor

a été dûment acceptée par le jury d'examen constitué de:

M. MARCHE, Claude, Ph.D., président

M. ROUSSELLE, Jean, Ph.D., membre et directeur de recherche

M. LEBEL, Thierry, Dr. Ing., membre et co-directeur de recherche

M. MOORE, Marc, Ph.D., membre

M. ZAWADZKI, Isztar, Ph.D., membre

A mon père, Amani Bagouari

et

A ma mère, Mariama Ibrahim

REMERCIEMENTS

Il me fait plaisir de remercier tout d'abord, M. Jean Rousselle, professeur à l'École Polytechnique de Montréal d'une part pour avoir accepté la direction de ce travail et d'autre part pour tous ses encouragements, son soutien moral et sa confiance. Ma grande reconnaissance va à M. Thierry Lebel, chargé de recherche à l'institut Français de recherche pour le développement en coopération (ORSTOM) pour avoir accepté la codirection de ce travail, ainsi que pour son encadrement scientifique, ses encouragements et son soutien moral. Je remercie François Trochu pour son soutien sur certains points théoriques de ce travail. Je tiens à remercier également M. Claude Marche, pour avoir accepté de présider mon jury de thèse, ainsi que les autres membres du jury M. Marc Moore et M. Isztar Zawadzki pour m'avoir fait l'honneur de juger ce travail.

Mes remerciements vont également à mes collègues de la section hydraulique et à tous les chercheurs, stagiaires et VSN avec qui j'ai eu à travailler, lors de mes différents séjours aux centres ORSTOM à Niamey, à Montpellier et à Grenoble. Enfin, mes remerciements vont au gouvernement du Canada pour m'avoir octroyé la bourse d'études à travers son programme de bourse de la francophonie.

SOMMAIRE

Le travail présenté dans ce mémoire de thèse, traite d'un certain nombre de problèmes reliés à l'estimation spatio-temporelle des pluies au Sahel. Il s'inscrit dans le cadre du programme EPSAT-NIGER qui vise à développer des algorithmes adaptés d'estimation des pluies au Sahel utilisant notamment des données de télédétection. Les trois années (1990, 1991 et 1992) de données spatio-temporelles de pluie collectées lors de cette expérience ont servi de base à cette étude. Notre objectif est d'une part d'évaluer l'efficacité des méthodes de la géostatistique pour l'estimation de la pluie aux pas de temps allant de la durée de l'événement à quelques minutes et d'autre part de proposer des nouvelles méthodologies permettant d'améliorer l'estimation des pluies sahéliennes ainsi que d'approfondir nos connaissances sur la climatologie des systèmes précipitants au Sahel.

Les études préliminaires entreprises sur l'utilisation des méthodes de la géostatistique montrent que ces méthodes ne sont pas dans tous les cas adaptées pour l'estimation des précipitations sahéliennes. Les problèmes rencontrés sont fonction du pas de temps de cumul. Au pas de temps saisonnier, la faible performance de ces méthodes est

principalement due à la structure de la saison des pluies qui est liée d'une part à la structure de base (intra-événement) et à l'agencement des événements dans le temps et l'espace. En effet les cumuls sahéliens proviennent principalement de trois systèmes pluvieux: convectifs isolé d'une part et systèmes convectifs de méso-échelle (SCM), d'autre part, présentant eux-mêmes deux degrés d'organisation, les SCM courants et les lignes de grains (SCM-LG).

Au pas de temps égal à celui de l'événement pluvieux, l'intermittence des champs déjà perceptible à cette échelle, nous ramène aussi à la question de classification des événements pluvieux pour une meilleure estimation soit par l'intermédiaire des méthodes de la géostatistique ou par les méthodes climatologiques ou globales. En effet les méthodes climatologiques supposent directement ou indirectement une certaine similarité au niveau des structures spatiales des événements.

Pour ces différentes raisons, des méthodes de classification des événements pluvieux ont été proposées et analysées sur la base des données collectées par un réseau pluviographique. Pour cela, il faut définir des paramètres et critères pertinents. Au total, neuf méthodes de classification caractérisant la structure spatiale et spatio-

temporelle des événements pluvieux, ont été proposées et analysées. Parmi ces méthodes, seulement les méthodes CA, SCA et DUPA ne font pas appel aux algorithmes de groupement (Clustering); par ailleurs seules les méthodes CA et CMRCA sont basées sur des données spatio-temporelles, les autres étant basées sur les données spatiales. L'apport de la classification a par la suite été évalué sur la base de deux modèles climatologiques: le krigeage climatologique et la méthode des seuils. Les résultats les plus intéressants ont été obtenus pour la combinaison de la méthode des seuils et de la méthode de classification basée d'un nouveau paramètre appelé UPA. L'importance de la classification pour l'estimation des champs pluvieux est sans équivoque.

Aux petits pas de temps, les problèmes sont: la forte intermittence des champs pluvieux et la présence manifeste d'une anisotropie au niveau des champs caractérisants les lignes de grains sahéliennes. A ces deux problèmes s'ajoutent ceux liés à la dynamique et à la dépendance temporelle entre les champs d'un même événement pluvieux. On ne peut espérer améliorer l'estimation aux faibles pas de temps qu'en considérant d'une certaine façon cette dépendance temporelle. Notre première tentative qui a consisté à généraliser le krigeage spatial dans le domaine espace temps n'a pas conduit à des très bons résultats comparés au krigeage spatial. Cette

situation est principalement due à une inadéquate interprétation de la dynamique du système pluvieux. C'est ainsi qu'une nouvelle approche basée sur la dynamique des champs pluvieux a été proposée.

Cette méthode dénommée krigeage dynamique est basée principalement sur l'advection des champs pluvieux. L'estimation s'effectue par rapport au repère mobile, contrairement aux méthodes classiques pour lesquelles l'estimation se produit par rapport au repère fixe. Cette nouvelle approche permet par construction, de résoudre non seulement le problème relatif à l'intermittence des champs, mais à celui relatif à l'anisotropie. Des données spatio-temporelles simulées pour des lignes de grains idéale et bruitée nous ont permis de la valider. Elle a par la suite été appliquée aux données réelles de lignes de grains où des excellents résultats ont été obtenus. Il est important de noter que cette bonne performance est principalement due à l'effet d'advection et non à la nature de l'interpolateur. Enfin une nouvelle méthode d'estimation de vitesse de l'orage en chacun des points du réseau a été proposée, mais n'a pas été validée faute de données appropriées.

ABSTRACT

The work presented in this Ph.D's thesis addresses a number of the problems associated with the space-time estimation of rainfall in the Sahel. This study was conducted in connection with the EPSAT-NIGER program, the objective of which is to develop algorithms specifically adapted to the estimation of rainfall in the Sahel, using remote-sensed data in particular. The space-time rainfall data collected during this experiment over a three-year period (1990, 1991 and 1992) served as the basis for the study. Our objective was threefold: to evaluate the effectiveness of geostatistical methods in estimating rainfall at time steps ranging from the duration of the event to a few minutes; to propose new methodologies which would result in improvements in the estimation of Sahelian rainfall; and to deepen of our understanding of the climatology of rainfall systems in the Sahel.

Preliminary studies undertaken on the use of geostatistical methods show that these methods are not in all cases adapted to the estimation of Sahelian rainfall. The problems encountered are a function of the time-steps over which the data are accumulated. On a seasonal level, these methods do not performs well, primarily because of the

structure of the rainy season, which is related on the one hand to the basic structure of the season itself (intra-event) and on the other to the way in which the events occur in time and space. Three precipitation systems contribute to Sahelian rainfall: isolated convective systems and meso-scale convective systems, the later presenting at two organizational levels, MSC currents and MSC squall lines (MSC-SL).

At time-steps equal to the duration of the rainfall event, the intermittence of the rainfall fields, which is already noticeable at this scale, leads as well to the question of classifying rainfall events in such a way as to yield better estimations, whether by geostatistical methods or by climatological or global methods. The climatological methods presuppose, directly or indirectly, a degree of similarity in terms of the spatial structures of the events.

For these reasons, methods of classification of rainfall events have been proposed and analyzed on the basis of data collected by rain-gauge networks. To use these methods, pertinent parameters and criteria must be defined. In total, nine classification methods characterizing the spatial structure and space-time structure of rainfall events have been proposed and analysed. Of these methods, only CA,

SCA and DUPA do not rely on the clustering algorithms; in addition, only CA and CMRCA are based on space-time data, the others being based on spatial data. The contribution of classification was then evaluated on the basis of two climatological models: the climatological kriging and the area threshold method. The most interesting results were obtained using a method which combines the threshold method with the classification method based on a new parameter called UPA. The importance of classification in estimating rainfall fields is beyond question.

In the case of short time-steps, the problems are as follows: the strong intermittence of rainfall fields and the clear presence of anisotropy associated with the rainfall fields which characterize the Sahelian squall lines. In addition, there are problems related to the dynamics and the time dependence between the fields of a single rainfall event. The only way to improve estimations at short time-steps is to consider this time dependence in a particular way. Our first attempt, which involved generalizing spatial kriging in space-time domain, did not yield very good results relative to spatial kriging. This was primarily due to inadequate interpretation of the dynamics of the precipitation system. As a result, a new approach, based on the dynamics of rainfall fields, was proposed.

This method, which has been named dynamic kriging, is based primarily on the advection of rainfall fields. The estimation is carried out relative to a mobile point of reference, as opposed to the classical methods by which the estimation is produced relative to a fixed point of reference. This new approach makes it possible to resolve, by construction, not only the problem of field intermittence, but the problem of anisotropy as well. We were able to validate the approach by using space-time data which had been simulated for ideal, noisy squall lines. The approach was then applied to real squall line data, and excellent results were obtained. It is important to note that this good performance was primarily due to the effect of advection, and not to the nature of the interpolator. Finally, a new method for estimating the speed of a storm at each point of the network was proposed, but was not validated due to inappropriate data.

TABLE DES MATIÈRES

	Page
Dédicace	iv
Remerciements	v
Sommaire	vi
Abstract	x
Table des matières	xiv
Liste des tableaux	xix
Liste des figures	xxvii
Liste des notations et symboles	xxxiv
Liste des annexes	xlvi
<u>CHAPITRE I</u> : INTRODUCTION GÉNÉRALE	1
1.1 Problématique de l'estimation des pluies au Sahel	1
1.2 Expérience EPSAT-NIGER	4
1.3 Objectifs et démarche de ce mémoire	5
<u>CHAPITRE II</u> : REVUE DE LITTÉRATURE	9
2.1 Introduction	9
2.2 Modèles d'estimation spatiale	11
2.2.1 Généralité sur la géostatistique	13

2.2.2	Problématique du krigeage pour l'estimation des précipitations	20
2.2.3	Méthodes de transfert indirect ou méthodes globales	22
2.3	Typologie des systèmes précipitants	26
2.4	Dépendance spatio-temporelle et dynamique de la pluie	29
2.4.1	Estimation spatio-temporelle	29
2.4.2	Dynamique du système précipitant	34
2.5	Conclusion	38

CHAPITRE III: TESTS ET PROBLÈMES RENCONTRÉS POUR

L'ESTIMATION DES PLUIES SAHÉLIENNES PAR LES MODÈLES DE LA GÉOSTATISTIQUE 40

3.1	Introduction	40
3.2	Caractéristiques générales des pluies sahéliennes .	41
3.3	Application de la géostatistique aux pluies sahéliennes	45
3.3.1	Considérations générales	45
3.3.2	Cumuls saisonniers et par événement	46
3.3.3	Pas de temps inférieur à l'événement	48
3.4	Conclusion	49

<u>CHAPITRE IV</u> : ESTIMATION SPATIALE A L'ECHELLE DE L'EVENEMENT: CLASSIFICATION ET PRISE EN COMPTE DE L'INTERMITTENCE	53
4.1 Introduction	53
4.2 Méthodes de classification sur la base des données sol	57
4.2.1 Méthode CDA	58
4.2.2 Méthode CCVA	59
4.2.3 Méthode CA	60
4.2.4 Méthode SCA	61
4.2.5 Méthode CUPAA	62
4.2.6 Méthode DUPA	63
4.2.7 Méthode CCVUPAA	64
4.2.8 Méthode CMRCA	65
4.2.9 Méthode CMRCA+DUPA	65
4.2.10 Synthèse	66
4.3 Application de la classification aux méthodes d'estimation	69
4.3.1 Application au krigeage climatologique	70
4.3.2 Application à la méthode des seuils	74
4.4 Autre intérêt de la méthode proposée: application à la climatologie sahélienne	77
4.5 Conclusion	78

<u>CHAPITRE V: ESTIMATION SPATIO-TEMPORELLE DES PLUIES AUX</u>	
FAIBLES PAS DE TEMPS: KRIGEAGE DYNAMIQUE .	82
5.1 Introduction	82
5.2 Krigeage spatio-temporel	86
5.3 krigeage dynamique	87
5.3.1 Prise en compte des corrélations temporelles	87
5.3.2 Du krigeage dynamique au krigeage dynamique lagrangien	91
5.3.3 Interpolation par krigeage dynamique lagrangien	96
5.3.4 Analyse structurale pour le cas sahélien . .	103
5.4 Mise en oeuvre du krigeage dynamique lagrangien 2D sur les lignes de grains sahéliennes	105
5.4.1 Procédure pour l'interpolation ponctuelle .	105
5.4.2 Estimation de la pluie moyenne (méthode simplifiée)	108
5.4.3 Procédure de validation	110
5.4.5 Création d'un étalon à l'aide d'un hyétogramme théorique de ligne de grains	112
5.5 Résultats et discussions	118
5.5.1 Données simulées	119
5.5.2 Données réelles (estimation ponctuelle) . .	124
5.5.3 Pluies moyennes	130

5.6 Conclusion 133

CHAPITRE VI: CONCLUSION GÉNÉRALE ET PERSPECTIVES DE
RECHERCHE 136

RÉFÉRENCES BIBLIOGRAPHIQUES 141

ANNEXES 161

LISTE DES TABLEAUX

Tableau 4.1: Synthèse et caractéristiques des méthodes de classification étudiées 67

Tableau 4.2: Synthèse des groupes obtenus pour les différentes méthodes (taille des groupes) 68

ANNEXE A1

TABLE 1: Parameters of the linear regression of UPA against $1/CV$ and $(1/CV)^2$. ($UPA = a_0 + a_1/CV + a_2/(CV)^2$). l_1 and l_2 are the estimates of l computed respectively as $l=1/a_1$ and $l=1/(2a_2)^{1/2}$ (see expression 19) 206

TABLE 2: Parameters and group size for each significant group obtained from each of the four methods for seasons 1990, 1991 and 1992 206

TABLE 3: Statistics and contributions of the different groups of method III to the total rainfall accumulation in the three seasons 1990, 1991 and 1992. For each group and for each year, we have, in the first row the non conditional statistics (average rainfall depth (mm), average standard of deviation (mm), and average coefficient of variation. In the second row, we have the corresponding conditional statistics, and the last row, we have the sample size, the percentage, represented by the group's events and the percentage of the group's events contribution for the total rainfall accumulation 207

TABLE 4: Parameters of the optimal ATM model for the different groups obtained by Methods I, II and III for the combined rainfall events of 1990, 1991 and 1992 (coefficient of determination R^2 , sample size N, optimum threshold Opt, slope of the regression St and the ordinate at the origin B) 208

TABLE 5: Optimal ATM model parameters and the validation errors for validation case I for the method III. (Optimal threshold Opt, coefficient of determination, slope of the regression St, the ordinate at the origin B, the error before classification RMSEb and the error after classification RMSEa) 209

TABLE 6: Calibration error of the 1992 mean areal rainfall events estimated by the ATM models before and after classification. (Optimal threshold Opt, coefficient of determination, slope of the regression St, the ordinate at the origin B, the error before classification RMSEb and the error after classification RMSEa). For method I, the events belonging the significant groups 1 and 2 are only considered, this is why the RMSEb is 2.67 and not 2.50 210

ANNEXE A2

TABLE 1: Some characteristics of the 27/07/90 and 20/08/91 rainfall events 250

TABLE 2: Percentage of reconstituted rainfall values with a ratio r less than one for the different 3D kriging models (K) and cumulative time-step durations for the 27/07/90 event 250

TABLE 3: Percentage of reconstituted rainfall values with a ratio r less than one for the different 3D kriging models (K) and cumulative time-step durations for the 20/08/91 event 251

TABLE 4: Percentage of reconstituted rainfall values of the two fields considered using 3D and 2D kriging for the cumulative time-step durations 5, 15, 30, 45 and 60 min for the 27/07/90 and 20/08/91 events 251

ANNEXE B1

Tableau B1.1: Valeurs des paramètres du variogramme (effet de pépite, portée et palier), du CV et de UPA pour les événements pluvieux de 1990 262

Tableau B1.2: Valeurs des paramètres du variogramme (effet de pépite, portée et palier), du CV et de UPA pour les événements pluvieux de 1991 264

- Tableau B1.3**: Valeurs des paramètres du variogramme (effet de pépite, portée et palier), du CV et de UPA pour les événements pluvieux de 1992 266
- Tableau B1.4**: Caractéristiques principales des événements pluvieux observés en 1990 (durée D, nombre de stations touchées N, cumul maximum C5, C10, C15, C30, C60 respectivement pour les temps de cumul de 5, 10, 15, 30 et 60 minutes) . . 268
- Tableau B1.5**: Caractéristiques principales des événements pluvieux observés en 1991 (durée D, nombre de stations touchées N, cumul maximum C5, C10, C15, C30, C60 respectivement pour les temps de cumul de 5, 10, 15, 30 et 60 minutes) . . 270
- Tableau B1.6**: Caractéristiques principales des événements pluvieux observés en 1992 (durée D, nombre de stations touchées N, cumul maximum C5, C10, C15, C30, C60 respectivement pour les temps de cumul de 5, 10, 15, 30 et 60 minutes) . . 272

Tableau B1.7: Groupes des événements pluvieux pour les neuf méthodes de classification étudiées pour l'année 1990. (M1= CDA, M2= CMRCA; M3= CA; M4=SCA; M5= CUPAA; M6=CCVA; M7=CCVUPAA; M8= DUPA et M9= CMRCA+CUPAA) 274

Tableau B1.8: Groupes des événements pluvieux pour les neuf méthodes de classification étudiées pour l'année 1991. (M1= CDA, M2= CMRCA; M3= CA; M4=SCA; M5= CUPAA; M6=CCVA; M7=CCVUPAA; M8= DUPA et M9= CMRCA+CUPAA) 276

Tableau B1.9: Groupes des événements pluvieux pour les neuf méthodes de classification étudiées pour l'année 1992. (M1= CDA, M2= CMRCA; M3= CA; M4=SCA; M5= CUPAA; M6=CCVA; M7=CCVUPAA; M8= DUPA et M9= CMRCA+CUPAA) 278

Tableau B1.10 Groupe (UPA) et statistiques des événements pluvieux de 1990 (moyennes, Moy; Ecart type, Ect; Probabilité de pluie non nulle, F0) 280

Tableau B1.11 Groupe (UPA) et statistiques des événements pluvieux de 1991 (Moyenne, Moy; Ecart type, Ect; Probabilité de pluie non nulle, F0) 282

Tableau B1.12 Groupe (UPA) et statistiques des événements pluvieux de 1991 (Moyenne, Moy; Ecart type, Ect; Probabilité de pluie non nulle, F0) 284

Tableau B1.13 Paramètres des modèles de seuillage aux seuils optimum pour les groupes des méthodes de classification CMRCA, SCA, DUPA et CMRCA+DUPA (en colonne on a pour un groupe donné, le seuil optimum, le coefficient de détermination R^2 , l'ordonnée à l'origine A et la pente B) pour les échantillons, 1990, 1991, 1992, 1990+1991 et 1990+91+92 286

LISTE DES FIGURES

	Page
<u>Figure 3.1:</u> Illustration du grandient climatologique Nord-Sud au Sahel (isohyètes moyennes annuelles)	43
<u>Figure 3.2:</u> Répartition de la pluie annuelle sur le Réseau EPSAT-NIGER (Taupin et al., 1992).	44
<u>Figure 3.3:</u> Champs d'isohyètes de la ligne de grains du 20/08/91. (Taupin et al., 1991)	50
<u>Figure 4.1:</u> Exemple d'isohyètes de différents événements pluvieux observés durant EPSAT-Niger (Taupin et al., 1992)	55
<u>Figure 4.2:</u> Variogrammes climatologiques des événements pluvieux appartenant aux groupes 1 et 2 de la méthode CUPAA pour l'année 1991	73
<u>Figure 5.1:</u> Illustration des repères fixe et mobile et de la dynamique de champs pluvieux	84

Figure 5.2: Répartition spatiale d'un champ pluvieux, événement du 20/08/91. a) repère fixe et b) repère mobile 85

Figure 5.3: Illustration du décalage temporel sur la base de la fonction d'autocorrélation 89

Figure 5.4: Illustration de l'estimation dynamique dans un cas unidimensionnel. V vitesse du système, P_{ik} et q_{ik} pluies dans les repères fixe et mobile 92

Figure 5.5: Isochrones de début de pluie pour la ligne de grain du 20/08/91 97

Figure 5.6: Variogrammes des champs pour l'événement du 20/08/91, a) repère mobile b) variogrammes climatologiques 106

Figure 5.7: Modèle théorique d'hyétogramme pour la simulation de la ligne de grains 115

Figure 5.8: Maillage pour la simulation de la ligne de grains 115

- Figure 5.9:** Histogrammes du vecteur vitesse des orages au Sahel, a) module vitesse b) direction . . 116
- Figure 5.10:** Erreurs de validation ponctuelle pour la LGP a) PRMSE, b) PME et c) Emax 121
- Figure 5.11:** Erreurs de validation ponctuelle pour la LGB (CV = 20%). a) PRMSE, b) PME et c) Emax . 122
- Figure 5.12:** Erreurs de validation ponctuelle pour la LGB (CV = 50%). a) PRMSE, b) PME et c) Emax . 123
- Figure 5.13:** Erreurs de validation ponctuelle pour l'événement du 20/08/91 sur le réseau de base, a) PRMSE, b) PME et c) Emax 127
- Figure 5.14:** Erreurs de validation ponctuelle pour l'événement du 20/08/91 sur la cible. a) PRMSE, b) PME et c) Emax 128
- Figure 5.15:** Erreurs de validation (PRMSE et PME) pour la pluie moyenne pour l'événement du 20/08/91. a) $\Delta t = 10$ mn et b) $\Delta t = 15$ mn 132

ANNEXE A1

Figure 1: The EPSAT-NIGER recording raingauge network, 1992. 211

Figure 2: Event rainfall versus proportion of zero-rainfall for the events of 1992. 212

Figure 3: Illustration of the rainy area above a threshold (C_i , i^{th} threshold value and A_{ci} is the rainy area where it rains above C_i).213

Figure 4: Rainy area function of the threshold (C_M is the maximum rainfall depth and A_0 is the total area of the network) 214.

Figure 5: Classification of storm structure types: a) type 1b and 1a; c) type 2b and 2a; c) type 3b and 3a. 215

Figure 6: Illustration of the UPA parameter and spatial crossig approach. 216

Figure 7: Histogram of the frequency factor values (three years combined). 217

Figure 8: Histogram of the UPA values (the three years combined). 218

Figure 9: Variance of the average conditional rainfall depth for the events of 1991 and the three groups of the DUPA method against the threshold considered. a) variance b) average conditional rain depth. 219-220

Figure 10: Coefficient of determination of the ATM model for various thresholds for each year and combination of years. 221

Figure 11: Coefficient of determination versus threshold for the ATM models after classification. a) method I, b) method II and c) method III 222-4

Figure 12: Optimal ATM models for the three years combined for the groups of method III groups. a) total sample, b) groups 1, c) group 2 and d) group 3. 225-8

ANNEXE A2

Figure 1: EPSAT-NIGER raingauges basic network and target area in 1991. 253

Figure 2: Standardized Root Mean Square Error (SRMSE) versus the number of fields considered for the spatio-temporal kriging K of the target area and at cumulative time-step durations of 5, 10, 15, 20, 25 and 30 min for the 27/07/90 event; a) first approach, and b) second approach. 254

Figure 3: Standardized Root Mean Square Error (SRMSE) versus the number of fields considered for the spatio-temporal kriging K of the target area and at cumulative time-step durations of 5, 10, 15, 20, 25 and 30 min for the 20/08/91 event, a) first approach, and b) second approach. 256

Figure 4: Standardized Root Mean Square Error (SRMSE) versus the cumulative time-step duration for 2D and 3D kriging models for the basic network (27/07/90 event). 258

<u>Figure 5:</u>	Standardized Root Mean Square Error (SRMSE) versus the cumulative time-step duration for 2D and 3D kriging models for the basic network (20/08/91 event).	259
-------------------------	---	-----

ANNEXE B3

<u>Figure B3.1:</u>	Variogrammes des champs pour l'événement du 27/07/90, a) repère mobile; b) variogrammes climatologiques.	292
----------------------------	--	-----

<u>Figure B3.2:</u>	Variogrammes des champs pour l'événement du 30/08/92, a) repère mobile; b) variogrammes climatologiques.	293
----------------------------	--	-----

<u>Figure B3.3:</u>	Erreurs de validation ponctuelle pour l'événement du 27/07/90 sur le réseau de base. a) PRMSE, b) PME et c) Emax.	294
----------------------------	---	-----

<u>Figure B3.4:</u>	Erreurs de validation ponctuelle pour l'événement du 30/08/92 sur le réseau de base. a) PRMSE, b) PME et c) Emax.	295
----------------------------	---	-----

- Figure B3.5:** Erreurs de validation ponctuelle pour l'événement du 27/07/90 sur la cible. a) PRMSE, b) PME et c) Emax. 296
- Figure B3.6:** Erreurs de validation ponctuelle pour l'événement du 30/08/92 sur la cible. a) PRMSE, b) PME et c) Emax. 297
- Figure B3.7:** Erreurs de validation ponctuelle (PRMSE) sur le réseau de base pour le pas de temps de cumul de 15 min, a) 27/07/90, b) 20/08/91 et c) 30/08/92. 298
- Figure B3.8:** Erreurs de validation ponctuelle (PRMSE) sur la cible pour $\Delta t = 10$ min, a) 27/07/90, b) 20/08/91 et c) 30/08/92. 299
- Figure B3.9:** Erreurs de validation (PRMSE et PME) pour la pluie moyenne l'événement du 27/07/90. a) $\Delta t = 10$ min et b) $\Delta t = 15$ min. 300
- Figure B3.10:** Erreurs de validation (PRMSE et PME) pour la pluie moyenne l'événement du 30/08/92. a) $\Delta t = 10$ min et b) $\Delta t = 15$ min. 301

Figure B3.11: Erreurs de validation (PRMSE et PME) pour la pluie moyenne sur la cible, $\Delta t = 10$ min a) 27/07/90, b) 20/08/91 et c) 30/08/92 . . . 302

LISTE DES ABRÉVIATIONS ET SYMBOLESCHAPITRE I

DMN	Direction de la Météorologie du Niger
EPSAT-Niger	Estimation des Pluies par SATellite, Expérience NIGER
FIFE	First International Satellite Land Surface Climatology Project (ISLSCP) Field Experiment
FIT	Front Inter-Tropical
GATE	GARP Atlantic Tropical Experiment
GARP	Global Atmospheric Research Program
HAPEX-MOBILITY	Hydrological Atmospheric Pilot EXperiment
HAPEX-SAHEL	Hydrological and Atmospheric Pilot EXperiment in the Sahel
OMM	Organisation Météorologique Mondiale
ORSTOM	Institut Français de Recherche pour le développement en Coopération
ZCIT	Zone de Convergence Inter-Tropical

CHAPITRE II

$P(x,y)$	hauteur de pluie observée au point de coordonnées (x,y)
----------	--

$P(x', y')$	hauteur de pluie observée au point de coordonnées (x', y')
$P(x_i, y_i)$ ou P_i	hauteur de pluie observée au point i
(x_i, y_i)	coordonnées spatiales du point i ($i=1, \dots, N$)
	N nombre total de stations de mesure
(x_o, y_o)	coordonnées spatiales du point o d'estimation
P_o	hauteur de pluie au point o
\hat{P}_o	hauteur de pluie estimée au o
λ_i	pondération associée au point i
h	distance euclidienne
$\nu(h)$	fonction variogramme
$E[]$	opérateur d'espérance
KS	Krigeage Simple
KO	Krigeage Ordinaire
KU	Krigeage Universel
KG	Krigeage Généralisé
FIA	Fonction Intrinsèque Aléatoire
x_i ($i=1, \dots, N$)	ensemble de N réels
$\nu(x_i - x_j)$	valeur du variogramme à la distance $x_i - x_j$
$P(x, y, k)$	hauteur de pluie observée au point de coordonnées (x, y) pour le champ k
K	nombre total de champs pluvieux
$a(k)$	constante associée au champ k

$W(x, y)$	processus spatial invariant au point (x, y)
b	bruit blanc
σ_b^2	variance de bruit blanc b
$\nu(h, k)$	fonction de variogramme associée au champ k
$\nu_w(h, k)$	fonction de variogramme associée au processus b
g	nombre de groupes
n_g	taille de l'échantillon à classifier
2D	deux dimensions (processus spatial)
3D	trois dimensions (processus espace-temps)
ARIMA	Auto Regressive Integrated Moving Average
$K(h)$	fonction de covariance unidimensionnelle
$K(h, \tau)$	fonction de covariance bidimensionnelle
τ	distance temporelle
$P(x, y, t)$	processus spatio-temporel
(x, y, t)	coordonnées spatio-temporelles
$t(x, y)$	temps d'observation d'une caractéristique de l'hétérogramme au point de coordonnées
$P(x, t)$	processus spatio-temporel
$\rho(h, \tau)$	correlation spatio-temporelle

CHAPITRE III

K	nombre de champs pluvieux
2D	deux dimensions (processus spatial)
3D	trois dimensions (processus espace-temps)
x, y	coordonnées spatiales

CHAPITRE IV

a_0, a_1, a_2	constantes de regression
CA	Crossing Approach
CCVA	Clustering Coefficient of Variation Approach
CCVUPAA	Clustering Coefficient of Variation Under Profile Area Approach
CDA	Clustering direct Approach
CUPAA	Clustering Under Profile Area Approach
CV	Coefficient of Variation
CMRCA	Clustering Maximum Rainfall Cumul Approach
DUPA	Distribution Under Profile Area
RMSE	Root Mean Square Error
R^2	coefficient of determination
UPA	Under Profile Area

CHAPITRE V

A	l'aire du ou d'une partie du réseau
a	constante de dérive
b_i ($i=1, \dots, N$)	Coefficients du système de krigeage dual
dA	aire élémentaire
KD	Krigeage Dynamique
KF	Krigeage Fixe
MD	Moyenne Dynamique
MF	Moyenne Fixe
Ex()	fonction partie entière
EM	Erreur moyenne
Emax	Erreur maximale
$G(i, h, j, h')$	fonction de covariance entre les observations aux points (i, h) et (j, h')
k	numéro du champ pluvieux
k_o	numéro du champ pluvieux pour lequel l'estimation est effectuée
K	nombre de champs pluvieux
K_{o_j}	covariance entre les points o et j
K_{i_j}	covariance entre les points i et j
l	Facteur de fréquence
LG	Ligne de Grain
LGP	Ligne de Grain Parfaite
LGB	Ligne de Grain Bruitée

LGR	Ligne de Grain Réelle
N	nombre de stations pluviographiques
N_k	nombre de stations en fonctionnement au champ k
M	nombre de stations les plus proches du point d'estimation à l'intérieur d'un rayon R
m	moyenne
PME	Erreur moyenne en pourcentage
PRMSE	Erreur moyenne quadratique en pourcentage
$P(x_i, y_i, k)$	pluie observée au point i et au champ k
$P_k(x_i, y_i)$	pluie observée au point i et au champ k
P_{ik}	pluie observée au point i au champ k
$P_{i,k,e}$	pluie estimée au point i au champ k
$P_{ij}(k)$	pluie simulée au k ^{eme} champ en (i, j)
$Pb_{ij}(k)$	pluie bruitée simulée au k ^{eme} champ en (i, j)
$P_o(k)$	pluie au point o et au champ k
$P_o(t)$	pluie au point o et au temps t
$P_A(k_o)$	pluie moyenne sur la surface d'aire A
$q(x, y, h)$	hauteur de pluie au point (x, y) pour le champ pluvieux h dans le repère mobile
$q(x_i, y_i, t_i)$	hauteur de pluie au point i pour le temps t_i dans le repère mobile
$q(x_o, y_o, t_o)$	hauteur de pluie au point à estimer au point o et au temps t_o dans le repère mobile

q_j	hauteur de pluie au point j (x_j, y_j, t_j) dans le repère mobile
$q(k*\Delta t)$	serie temporelle des cumuls pour le modèle d'hyétogramme
(x_i, y_i)	coordonnées spatiales du point i
(x_o, y_o)	coordonnées spatiales du point o d'estimation
d_{oi}	distance métrique entre les points o et i
d_{ij}	distance entre le point de grille (i, j) et le point de référence (début de pluie)
d_j	distance séparant le point j et le point de début de pluie sur le réseau
DL	dimension du réseau dans le sens de déplacement de l'orage
d_k	distance parco
d_o	distance entre le point d'estimation o et le point de début de pluie sur le réseau
$C_{ij}(\tau)$	covariance entre les séries temporelles observées aux points i et j pour un décalage temporel de τ
τ	distance ou décalage temporel
τ_{max}	décalage temporel pour lequel la covariance maximale entre deux points
$\tau_{max}(ij)$	décalage temporel pour lequel la covariance $C_{ij}(\tau)$ est maximale

τ_{oi}	décalage temporel entre les observations aux points i et o (point d'estimation)
Δx_{ij}	différence des abscisses entre les points i et j
Δy_{ij}	différence des ordonnées entre les points i et j
$\Delta x_{\max}(\tau)$	abscisse du point de covariance maximale de la surface de corrélation associée au décalage τ
Δy_{\max}	ordonnées du point de covariance maximale de la surface de corrélation associée au décalage τ
Δt	pas de temps de cumul
t_j	temps mis par l'orage du point de début de pluie sur le réseau au point j
t_a	temps d'arrivée de la pluie au point a
t	temps pour lequel l'estimation est effectuée
U_x, U_y	composantes x et y du vecteur vitesse
U	module du vecteur vitesse
λ_i ($i=1, \dots, N$)	pondération associée à l'observation au point i
λ_{ik}	pondération associée à l'observation au point i pour le champ pluvieux k

λ_{ih}	pondération associée à l'observation au point i pour le champ pluvieux h dans le repère mobile
μ	multiplicateur de Lagrange
s	ecart type
T	temps mis par l'orage pour traverser entièrement la superficie A
T'	temps mis par l'orage pour arriver au début de la surface A
T_c	durée de la partie convective de l'orage sur le réseau de mesure
Z	nombre de champs à considérer pour l'estimation de $P_A(k_o)$
Z_1	nombre entier le plus proche de $T'/\Delta t$

ANNEXE A1

AC or AC_θ	percentage of the rainy area where the rainfall exceeds the threshold C for time step-duration θ
AC_j	percentage of rainy area where rainfall exceeds the threshold C for event j
a_j	parameter of the exponential function
AR	mean Areal Rainfall
ATM	Areal Time Integral

BC	area of rainfall below the threshold C
BC_j	BC value for a given event j
$b(c)$	ordinate at the origin for the regression equation between $m(t)$ and $F(t,C)$
C	threshold
C^*	standardized threshold by the maximum rainfall cumul
CM or CM_θ	maximum cumul rainfall for time-step duration θ for a given event
CM_j	value of CM for event j
CUPAA	Clustering Under Profile Area Approach (classification method)
CV	Coefficient of Variation
CV_j	value of CV for event j
CMRCA	Clustering Maximum Rainfall Cumul Approach (classification method)
$E[]$	Expectation operator
F_0	probability of zero rainfall
$F(t,C)$	fractional area where the rainfall recorded at time t exceeds the threshold C
$F(C,k)$	value of $F(t,C)$ for event k
ITCZ	InterTropical Convergence Zone
$I(x,y,t)$	indicator variable at point (x,y,t)
$I(i)$	value of $I(x,y,t)$ for a given point i

l_j	frequency factor of the maximum rainfall cumul for event j
MCS	Mesoscale Convective Systems (MCS)
m_j or $m(k)$	mean areal rainfall for event j or k
$m(t)$	mean intensity at time t
P_n	probability (empirical approach)
PGA	Percentage of Gauges Affected by rainfall
Q_n	quantile value
RMSE	Root Mean Square Error
$R\theta(i)$	rainfall cumul at gauge i for time-step duration θ
$R_T(x,y)$	value of $R\theta(i)$ for time-step duration equal the total duration of the event
R^2	coefficient of determination
$S(c)$	slope of the regression equation between $m(t)$ and $F(t,C)$
s_j	standard of deviation for event j
SCA	Spatial Crossing Approach (classification method)
UPA	Under profile Area
Z_j	rainfall value for event j
λ	frequency factor of the population
μ_T	mean of the rainfall process (population)
σ	standard of deviation (population)

ANNEXE A2

2D	deux dimensions (spatial)
3D	trois dimensions (espace temps)
a_p	constantes du monome $f_p(x,y,t)$
$b_p(p=1..,M)$	coefficients du krigeage dual
$f_p(x,y,t)$	monome
h_{ij}	distance euclidienne entre les points i et j
K	nombre total de champ
$K(h)$	fonction de covariance généralisée
$K(h, \tau)$	fonction de covariance bidimensionnelle
$l(q)$	nombre de monomes
M	nombre total d'observations spatio-temporelles
N	nombre de stations de mesure
θ	pas de temps de cumul
$\lambda_i (i=1, .., N)$	pondération associée à l'observation au point i
λ_{ik}	pondération associée à l'observation au point i pour le champ pluvieux k
$P(x,y,t)$	processus espace temps
$P(x_i, y_i, k)$	hauteur de pluie observée au point i pour le champ k
P_i	pluie observée au point i (x_i, y_i, t_i)
q	dégré du polynome de dérive

k_o	champ pour lequel l'estimation est effectuée
$m(x, y, t)$	dérive associée au processus $P(x, y, t)$
r_i	performance relative du krigeage 3D par rapport au krigeage 2D pour l'estimation au point i
SRMSE	erreur moyenne quadratique standardisée par la moyenne
V	module du vecteur vitesse de l'orage
$X=(x, y, t)$	coordonnées espace temps d'une observation

ANNEXE B2

h	distance euclidienne
i, j	points de mesure
$N(h), N_T(h)$	
$N_c(h)$ et $N_o(h)$	nombre de couple de points (i, j) appartenant à la classe de distance h pour les variogrammes conditionnel, intermittent et croisé et nul
$\gamma_I(h)$	variogramme intermittent
$\gamma_c(h)$	variogramme croisé
$\gamma(h)$	variogramme conditionnel
$R(i), R(j)$	pluies observées aux points i et j
r	portée du variogramme
σ_r	palier du variogramme conditionnel

ANNEXE B4

a_0, a_1, a_2	coefficient de dérive linéaire
b_j ($j=1, \dots, N$)	coefficients du système du krigeage dual
h	distance euclidienne
$K(h)$	fonction de covariance généralisée
$K'(h)$	fonction dérivée de $K(h)$
$m(x,y)$	dérive ou espérance de $t(x,y)$
N	nombre de stations pluviographiques
$t(x,y)$	temps d'observation d'une caractéristique de l'hyétogramme en (x,y)
t_i ($i=1, \dots, N$)	valeurs de $t(x,y)$ aux points de mesure
$V_x(x,y)$	composante en x du vecteur vitesse pour le point de coordonnées (x,y)
$V_y(x,y)$	composante en y du vecteur vitesse pour le point de coordonnées (x,y)
(x,y) ou X,Y	coordonnées spatiales

LISTE DES ANNEXES

Annexe A1 162

Typology of Rainfall Fields to Improve Rainfall Estimation in the Sahel by the Area Threshold Method (ATM) by A. Amani, T. Lebel, J. Rousselle et J.D. Taupin, soumis après corrections des réviseurs à Water Resource Research.

Annexe A2 229

Squall Line Rainfall Estimation at Small Time Steps in the Sahel by Dual Spatio-temporal Kriging, By A. Amani, J. Rousselle et F. Trochu, soumis au Journal of Hydrological Sciences.

Annexe B1 261

Complément des résultats sur la classification

Annexe B2 287

Variogramme d'un champ intermittent

Annexe B3 291

Résultats complémentaires au krigeage dynamique

Annexe B4 303

Estimation des vitesses d'approche par krigeage dual

CHAPITRE I: INTRODUCTION GÉNÉRALE

1.1 PROBLÉMATIQUE DE L'ESTIMATION DES PLUIES AU SAHEL

La connaissance de la distribution spatiale et temporelle des précipitations sahéliennes à différentes échelles de temps et d'espace est d'une grande importance pour les pays sahéliens. En effet, plus de 90% de la population sahélienne est rurale; l'agriculture et l'élevage constituent leur principale, voire leur unique source de revenu. L'importance des pluies au Sahel a été mise en évidence par les épisodes de sécheresses intenses et désastreuses (1972, 1973 et 1984) de ces 25 dernières années observées dans cette région. Une importante littérature existe sur les études et les tentatives d'explication de sécheresses sahéliennes (Sircoulon, 1976; Faure et Gac, 1981; Hubert et Charbonnel, 1988a; Nicholson, 1980; 1981; 1986; Nicholson et al., 1988; Janicot, 1992a,b; Céron, 1993; Janicot et Fontaine, 1993 entre autres). Il ressort de toutes ces études que les mécanismes à la base de la sécheresse au Sahel sont multiples et complexes.

En dehors de la fragilité de l'écosystème sahélien, mise en évidence par les effets cumulatifs de récentes sécheresses, la zone sahélienne joue aussi un rôle important

dans la circulation générale atmosphérique (Dhonneur, 1985). En effet le Sahel appartient à la partie chaude du moteur thermique responsable de cette circulation générale atmosphérique. La quantité de vapeur d'eau contenue dans l'atmosphère et dans l'interface sol-atmosphère joue un rôle primordial dans le mécanisme de transfert d'énergie entre les différentes parties de l'atmosphère. Afin de caractériser l'influence de l'interface sol-atmosphère sur les échanges thermiques et hydriques au sein de l'atmosphère, la communauté scientifique internationale a entrepris des expériences pilotes sous différents climats (Hapex-Mobihy, Fife, Boreas). C'est le cas de l'expérience HAPEX-Sahel qui a eu lieu au Niger en 1992 (Goutorbe et al., 1994). La variabilité spatiale et temporelle des pluies joue un grand rôle dans de telles expériences, particulièrement pour ce qui est d'une zone semi-aride comme le Sahel (Goutorbe et al., 1994).

L'étude des précipitations au Sahel présente donc un intérêt certain pour les populations des pays sahéliens et pour la communauté scientifique internationale. L'obtention d'estimations fiables de pluie sur une large gamme d'échelles de temps et d'espace est nécessaire aussi bien aux climatologues, qu'aux météorologues, aux agronomes et aux hydrologues. Ceci est dû au fait que la distribution spatiale

et temporelle de précipitations est un facteur déterminant pour le bilan hydrique de ces régions.

La nature des réseaux pluviométriques nationaux des pays sahéliens, avec une densité moyenne inférieure à 1 pluviomètre au 6000 km², rend impossible toute étude de pluies sahéliennes aux faibles pas de temps et d'espace. Bien qu'il existe une certaine régularité des isohyètes des pluies moyennes inter-annuelles sur l'Afrique de l'Ouest, cette régularité disparaît pour une année donnée. Donc à priori, les techniques de télédétection peuvent être une bonne alternative pour estimer les pluies sahéliennes (Thiao, 1988; Ba et Deschamps, 1990). L'exploitation des données de télédétection pour l'estimation des pluies sahéliennes, nécessite un renforcement, localement soit-il des réseaux pluviométriques de pays sahéliens.

C'est à la fois pour mieux caractériser et modéliser les fluctuations pluviométriques au Sahel, et pour développer des algorithmes d'estimation adaptés au régime pluviométrique de cette région, que le réseau EPSAT a initié l'expérience EPSAT-NIGER (Lebel et al., 1992).

1.2 EXPÉRIENCE EPSAT-NIGER

EPSAT-NIGER est une expérience pilote pour l'estimation de précipitation au Sahel, qui a été mise sur pied par l'ORSTOM (Institut Français de Recherche en Coopération), en association avec la direction nationale de la météorologie du Niger (DMN), et le laboratoire d'Aérodynamique de Toulouse.

Grâce à l'utilisation conjointe d'un réseau dense de pluviographes à mémoire statique (une centaine d'appareils répartis sur 16 000 km²) et d'un radar météorologique bande C (5.4 cm de longueur d'onde), le dispositif expérimental d'EPSAT-NIGER a permis pour la première fois, de 1989 à 1993, d'étudier la répartition des précipitations en milieu sahélien aux petites échelles de temps et d'espace. La thématique de l'expérience est présentée de façon détaillée dans Cadet (1988). Ses objectifs principaux sont:

- étudier le potentiel réel des données satellitaires pour estimer les précipitations sahéliennes, plus particulièrement aux petites et moyennes échelles de temps et d'espace;
- déterminer les moyens d'étalonnage (radar ou pluviomètre) nécessaires pour cela et développer des algorithmes correspondants.

- modéliser les précipitations sahéliennes et approfondir notre connaissance sur la climatologie de cette région.

Des travaux de recherche sont nécessaires afin d'atteindre ces différents objectifs. Il faut, dans une première phase étudier et caractériser séparément les données collectées par chacun des capteurs (réseau sol, satellite et radar). Et puis dans un deuxième temps, procéder à une superposition des données des différents capteurs par des techniques appropriées. Les données de l'expérience peuvent également permettre d'explorer la possibilité d'un suivi en temps réel des précipitations sahéliennes à une échelle dépendant de la combinaison de capteurs utilisés. Les travaux de recherche relatifs à l'analyse, à la caractérisation et à la modélisation des données sol sont particulièrement déterminants dans la mesure où ces dernières constituent notre seule vérité terrain.

1.3 OBJECTIFS ET DÉMARCHE DE CE MÉMOIRE

Les travaux entrepris ici s'inscrivent dans le cadre général des recherches initiées par l'expérience EPSAT-NIGER. Nos travaux sont une suite logique de ceux de Thauvin (1992) qui s'était intéressée à l'estimation spatiale des pluies sahéliennes pour des échelles allant de l'événement à la

saison sur la base des données EPSAT-NIGER pour les années 89 et 90. Nos travaux préliminaires sur l'applicabilité de techniques de krigeage nous ont conduit à distinguer deux niveaux de pas de temps: pas de temps égaux à celui de l'événement et plus et pas de temps inférieurs à celui de l'événement. La synthèse des problèmes rencontrés est présentée au chapitre 3.

Nous nous intéresserons donc aux pas de temps inférieurs ou égal à celui de l'événement pour lesquels différentes approches de modélisation seront testées, en particulier celles liées à la géostatistique. A l'échelle de l'événement et plus, le principal problème est celui des structures différentes entre les événements pluvieux, différence de structures qui amoindrit l'efficacité des modèles climatologiques (ou modèles globaux). Les modèles climatologiques ou globaux prennent en compte pour la modélisation d'un événement pluvieux donné, les données provenant d'autres événements pluvieux. Une stratification des événements pluvieux est donc nécessaire, afin d'améliorer l'estimation fournie par ces modèles. La classification des événements pluvieux sahéliens permet également de caractériser une saison de pluies donnée, en analysant la proportion des différents événements enregistrés. Notons que la question de la classification des événements pluvieux a

été longtemps négligée dans la littérature bien que son importance ait été soulignée par un certain nombre de chercheurs dont entre autres Niemczynowicz et Dahlblom. (1984), Desbordes (1987) et Kottegoga et Kassim (1991). La classification permet de rendre plus efficace la méthode des seuils pour l'estimation de la pluie moyenne sur une grande échelle d'espace. Elle permet également de montrer que la composition de la saison des pluies en événements est un bon marqueur.

Aux faibles pas de temps, les techniques classiques de géostatistique doivent être modifiées pour pouvoir tenir compte de certaines particularités des champs pluviométriques à ces pas de temps. La première particularité est la forte dépendance existant entre les caractéristiques spatiale et temporelle de la pluie. L'autre particularité essentielle est la dynamique de pluie. Celle-ci doit donc être prise en ligne de compte directement ou indirectement dans les algorithmes d'estimation. Aussi, à ces échelles de temps, les pluies dues aux lignes de grains présentent une anisotropie manifeste dont il faut tenir compte. Plusieurs approches, qui vont des techniques de krigeage spatio-temporel avec prise en compte explicite ou non de la dynamique à la technique de krigeage dynamique qui est une nouvelle façon d'interpréter les données spatio-temporelles de pluie, ont été proposées afin

de tenir compte de ces deux importantes particularités du phénomène de pluie.

Le corps du texte est organisé comme suit. Une revue de littérature traitant de l'estimation de la pluie, est présentée au chapitre 2. Le chapitre 3 présente la synthèse des principaux problèmes rencontrés pour la mise en oeuvre des techniques classiques de géostatistique pour l'estimation des pluies sahéliennes aux différents pas de temps de cumul (saison, événement, ..jusqu'à 5 mn). Ce chapitre présente la synthèse de nos résultats et de ceux de l'équipe EPSAT-NIGER sur le test des techniques de géostatistique. Le chapitre 4 est consacré à une présentation sommaire des méthodes de classification proposées ainsi que de leur apport pour l'estimation de la pluie par les deux modèles globaux considérés et son apport pour l'étude de la climatologie sahélienne. Les résultats détaillés de cette partie sont présentés dans un article soumis pour publication au Water Resource Research. Afin de tenir compte aux faibles pas de temps, de la dynamique, de l'intermittence et de l'anisotropie des champs pluvieux dans l'estimation, une nouvelle méthode d'estimation, dénommée krigeage dynamique a été proposée et testée au chapitre 5. La conclusion générale présentant nos différentes contributions ainsi que les perspectives de recherches sont présentées au chapitre 6.

CHAPITRE II: REVUE DE LITTÉRATURE

2.1 INTRODUCTION

Le besoin de modéliser des données de précipitations dans le cadre de projets de ressources hydriques, fait qu'il existe une importante littérature sur l'analyse et l'étude des précipitations. En fonction des objectifs du projet, les modèles de précipitations peuvent être divisés en trois grandes familles (Georgakakos et Kavvas, 1987; Krajewski et Creutin, 1990): (1) les modèles d'estimation; (2) les modèles de simulation et (3) les modèles de prévision. Il est important de noter que tous ces modèles (prévision, simulation et estimation) ont la même base, celle qui consiste à modéliser tout d'abord le lien existant entre les données disponibles. Leur différence réside au niveau de la façon dont ce lien est par la suite interprété.

Les modèles d'estimation permettent de déterminer les valeurs manquantes (valeurs non observées). Les modèles de prévision, avec leur caractère chronologique, permettent une prédiction des valeurs futures (exemple des modèles Box-Jenkins unidimensionnels ou multivariés). Quant aux modèles de simulation, ils servent à produire des scénarios

d'observations ayant les mêmes caractéristiques statistiques que les données initiales.

Au lieu de distinguer les modèles de précipitations sur la base des objectifs, on rencontre plus généralement dans la littérature, une division des modèles des précipitations sur la base de la nature des données disponibles. Dans le cas des données de précipitations mesurées au sol, par opposition aux mesures de télédétection, on distingue trois familles de modèles: modèles ponctuels, spatiaux et spatio-temporels (Waymire et Gupta, 1981a,b; Waymire et Rodriguez-Iturbe, 1984; Smith et Karr, 1985a,b; Valdes et al., 1985; Desbordes, 1987; Georgakakos et Kavvas, 1987). Cette distinction a l'avantage de produire une discrimination systématique des modèles contrairement à la précédente. Les modèles temporels de précipitations ont fait l'objet de beaucoup d'études. La plupart de ces modèles sont de type stochastique (Waymire et Gupta, 1981a,b; Smith et Karr, 1985a; Rodriguez-Iturbe et al., 1987a,b; Rodriguez-Iturbe et al., 1988 entre autres). Ils sont utilisés pour caractériser soit la structure externe de la pluie (durée, hauteur et intensité moyenne) ou bien sa structure interne (durées de cellules, temps séparant deux cellules, intensités et hauteurs des cellules pluvieuses). Les modèles spatiaux caractérisent la structure spatiale de la pluie, alors que les modèles spatio-temporels sa structure

spatiale et temporelle (Smith et Karr, 1985a; Rodriguez-Iturbe et al, 1986; Cox et Isham, 1988). Enfin, signalons qu'il est possible de distinguer les modèles de pluie en termes d'aggrégation ou de désaggrégation (Valencia et Shaake, 1972; Meija et Rousselle, 1976).

Une revue de littérature détaillée sur les différents modèles de précipitations est présentée dans Amani (1992). Pour un souci de synthèse et de clarté, seule la synthèse de la revue de littérature relative aux modèles spatiaux et spatio-temporels pour fin d'estimation des précipitations et aux questions connexes est présentée ici.

2.2 MODÈLES D'ESTIMATION SPATIALE

L'idée d'estimer la pluie à partir des modèles conceptuels proposés par Waymire et Rodriguez-Iturbe. (1984); Rodriguez-Iturbe et al. (1986;1988); Eagleson et al. (1987), est très séduisante. Mais le manque des données précises sur les caractéristiques des cellules élémentaires constituant les systèmes pluvieux, rend ces modèles pratiquement inexploitable et hasardeux d'utilisation pour l'estimation de la pluie au sol. Ainsi les modèles non conceptuels d'estimation de pluie à partir des observations sur un réseau de pluviomètres auront toujours leur place dans l'étude des

précipitations. Ils sont les seuls considérés par la suite.

On cherche à construire une fonction $P(x,y)$ d'interpolation sur l'échantillon des mesures spatiales de pluie (x_i, y_i, P_i) avec $i=1, \dots, N$ (N étant le nombre de stations de mesure) et (x_i, y_i) étant les coordonnées spatiales de la station i . Il existe dans la littérature plusieurs façons d'aborder cette question: de façon déterministe ou stochastique. L'idée principale est la même pour la plupart de ces méthodes. Elle consiste à estimer la précipitation en un point de coordonnées (x_o, y_o) ou la précipitation moyenne sur une surface, comme étant égale à une combinaison linéaire des valeurs observées aux points de mesure. C'est à dire que si p_i $1 \leq i \leq N$ est la mesure à la station i et P_o , l'estimation au point o , on a

$$\hat{P}_o = \sum_{i=1}^N \lambda_i P_i \quad (2.1)$$

avec λ_i étant la pondération associée à la station i . Les méthodes diffèrent principalement au niveau de la façon de déterminer les poids λ_i . L'avantage des méthodes stochastiques d'estimation sur les modèles déterministes est qu'elles donnent dans la plupart des cas (cas des réseaux de mesure à faible densité) une meilleure estimation que ces derniers. Une variance d'estimation est associée aux modèles stochastiques. Les méthodes déterministes classiques

d'estimation de pluie (méthodes de Thiessen, des isohyètes, de Cressman, de quatre cadrants, splines, ..etc) étant très bien connues de tous, elles ne sont pas présentées ici (Amani, 1992). Seules les méthodes stochastiques (krigeages ou géostatistiques) et les modèles globaux sont considérés.

2.2.1 Généralité sur la géostatistique

Les méthodes stochastiques, avec leur fondement probabiliste, permettent de tenir compte dans l'estimation de la structure spatiale propre au champ et en plus donnent un baromètre quantitatif sur la qualité de l'estimation (variance d'estimation). Ces techniques, basées sur le concept de variables aléatoires ont été développées depuis un quart de siècle. Elles sont principalement l'oeuvre de Gandin (1965) et Mathéron (1970). Ce sont les techniques de la géostatistique (Krigesages) (Matheron, 1970) ou les techniques d'interpolation optimale (Gandin, 1966; Tabios et Salas, 1985). La technique du krigage a été développée pour la première fois dans le domaine minier, pour l'estimation de la teneur du minerai à partir des données de sondage de prospection en un certain nombre de points. Tandis que ces techniques sous l'appellation interpolation optimale, ont été développées dans le domaine de la météorologie-climatologie. Depuis, ces méthodes ont fait l'objet de nombreuses

applications dans les domaines les plus variés. Dans le domaine des ressources hydriques, on peut citer entre autres les travaux de Gambolati et Volpi (1979); Ahmed et Marsily (1987); Ribeiro et al. (1989) et Rouhani et al. (1989) pour l'étude des nappes phréatiques, Villeneuve et al. (1979); Bogardi et al. (1985); Shamsi et al. (1988) et Kassim et Kottegoda (1991) pour le design de réseaux hydrométrique et pluviométrique, Chua et Bras (1982); Creutin et Obled (1982); Laborde et Lempereur. (1986); Meylan (1986); Lebel et Laborde (1988) et Marand et Zumstein (1990) pour l'estimation des pluies et Andrieu (1986) et Delrieu et al. (1988) pour l'estimation des précipitations à partir des données de pluie et radar.

Les mesures $P(x_i, y_i)$ sont prises comme étant la réalisation d'une fonction aléatoire $P(x, y)$. Dans le cadre du krigeage et de l'interpolation optimale, seuls les deux premiers moments de la variable aléatoire $P(x, y)$ sont exploités. On raisonne ainsi dans un espace de variable aléatoire. De ce fait, l'équation (2.1) est interprétée comme étant l'équation d'un estimateur égal à la combinaison linéaire des variables aléatoires. Les techniques de krigeage sont basées sur les deux conditions d'optimalité de l'estimateur donné en (1): l'estimateur doit être non biaisé et doit avoir l'erreur quadratique minimale. L'application de

ces deux conditions sur cet estimateur conduit au système linéaire de krigeage dont la résolution nécessite la modélisation de la fonction de covariance spatiale (Journel et Huijbregts, 1978).

Pour pouvoir modéliser la fonction de covariance à partir d'une seule réalisation spatiale, des hypothèses supplémentaires sont nécessaires. Ainsi, on suppose que le processus spatial est stationnaire d'ordre 2, c'est-à-dire que la fonction de covariance ne dépend que de la distance vectorielle entre les couples de points. Dans certains cas (cas où le processus ne possède pas de variance finie), il est préférable de relaxer cette hypothèse en une hypothèse moins forte: l'hypothèse intrinsèque. Cette dernière, suppose que la variance des accroissements ou l'espérance des carrés des accroissements (cas où le processus est stationnaire d'ordre 1, espérance constante) n'est fonction que de la position relative des stations. C'est de cette hypothèse que découle la notion de variogramme introduite par Mathéron (1970). Le variogramme est donné par

$$E[(P(x,y) - P(x',y'))^2] = 2\gamma(h) \quad (2.2)$$

avec $P(x,y)$ et $P(x',y')$ étant les observations aux points de coordonnées (x,y) et (x',y') ; h est la distance vectorielle

séparant les deux points et $\gamma(h)$ est la fonction de variogramme. On vérifie très bien que l'hypothèse de stationnarité d'ordre 2 implique l'hypothèse intrinsèque. Mais l'inverse n'est pas vrai.

Aussi dans la majorité de cas, on suppose que le processus est isotrope, c'est-à-dire que la fonction de covariance n'est fonction que de la distance euclidienne (indépendante de la direction). Dans le cas contraire, on dit que le processus est anisotrope. On distingue généralement deux sortes d'anisotropie: anisotropies géométrique et zonale. Un processus à anisotropie géométrique se ramène par simple transformation de coordonnées à un processus isotrope, de telle sorte que l'anisotropie géométrique est facilement intégrable au niveau de la mise en oeuvre du krigeage (Journel et Huijbregts, 1978; Journel et Froideveaux, 1982; Meyers et Journel, 1990); Le terme d'anisotropie zonale n'est pas accepté par la majorité des géostatisticiens à cause de la confusion que ça peut engendrer. C'est ainsi que Zimmerman (1993) a proposé une autre classification de l'anisotropie. Il distingue les anisotropies sur la portée, le palier et sur l'effet de pépite. Pour chaque type d'anisotropie, on suppose que les autres caractéristiques du variogramme ne changent pas avec la direction.

Sur la base du moment d'ordre 1 du processus spatial (stationnarité ou non stationnarité; moyenne connue ou non), on distingue plusieurs types de krigeage. Les techniques du krigeage simple (KS) et ordinaire (KO) s'appliquent à un processus stationnaire, alors que les techniques du krigeage universel (KU) et de fonctions aléatoires intrinsèques (F.I.A) ou krigeage généralisé (KG) s'applique à un processus non stationnaire. Ces dernières se réduisent au krigeage ordinaire dans le cas où le processus est stationnaire. Des fonctions de dérive polynomiale sont considérées au niveau des krigeages universel et généralisé. Dans le cas du krigeage généralisé, le choix de k (degré de la dérive) détermine aussi la fonction théorique de covariance généralisée à utiliser au niveau de l'étape de résolution du système de krigeage. Alors que dans le cas du krigeage universel, une fonction théorique de covariance ou de variogramme est à choisir parmi plusieurs modèles théoriques. Ce choix peut se faire sur la base de l'analyse spatiale directe des observations ou des résidus, après avoir filtré la dérive. Le choix de la forme de la dérive est très important et il doit si possible toujours être fonction de la physique du phénomène (Gambolati et Vopipi, 1979).

Deux étapes sont nécessaires pour la mise en oeuvre du krigeage. La première étape consiste à modéliser la fonction

de covariance ou de variogramme du processus spatial. Cette étape est appelée analyse structurale. La deuxième étape consiste à résoudre le système de krigeage obtenu par la minimisation de la variance d'estimation sous les contraintes de la condition de non biais de l'estimateur. L'estimateur de variogramme proposé par Mathéron (1970) est le plus utilisé dans la littérature. C'est un estimateur non biaisé. Il est donné par l'équation 1 de l'annexe 4.3. Le variogramme expérimental est par la suite modélisé par le biais de modèle théorique de variogramme. Les techniques de lissage utilisées à cette fin sont: les moindres carrés, le maximum de vraisemblance (Kitanidis, 1988; Zimmerman et Zimmerman., 1991) et la méthode d'erreur moyenne quadratique d'interpolation (Lebel et Bastin, 1985).

Afin d'améliorer la modélisation du variogramme dans le cas des données spatiales fragmentaires, certains auteurs ont proposé de travailler dans la phase de l'analyse structurale non pas sur le variogramme, mais sur la fonction intégrale du variogramme ou variogramme cumulatif (Sen, 1989; 1992; Delay et Marsily, 1994). L'estimateur de cette nouvelle fonction a généralement une structure régulière plus facilement modélisable que celle du variogramme (Delay et Marsily, 1994). Par contre, la forme très similaire de cette fonction pour les différents variogrammes théoriques, rend

difficile le choix a priori d'un variogramme théorique donné. C'est une très grosse limitation de cette approche. Une fois la fonction modélisée, on obtient le variogramme modélisé par une simple dérivation.

En dehors de la division des techniques géostatistiques entre les méthodes linéaires et non linéaires (pour les méthodes non linéaires voir Rendu. (1980); Yates et al. (1986a,b)), krigeage stationnaire et non stationnaire, signalons qu'il existe une autre division importante entre les méthodes, suivant l'interprétation qui résulte de la mise en oeuvre de l'étape d'analyse structurale: krigeages classique, bayésien et flou. Au niveau des krigeages classiques, on suppose que la fonction de structure est bien connue même si elle est estimée. Le krigeage bayésien (Omre, 1987; Handcock et Stein, 1993; Handcock et Wallis, 1994) et le krigeage flou (Bardossy et al., 1990a; 1990b) permettent de tenir compte de l'imprécision liée à la connaissance de la fonction de structure, imprécision due aux données fragmentaires ne permettant pas une bonne analyse structurale.

2.2.2 Problématique du krigeage pour l'estimation des précipitations

L'hypothèse de continuité étant à la base des techniques de krigeage, leur application pour l'estimation des précipitations doit nécessairement tenir compte du pas de temps de cumul. Le phénomène de pluie est de toute évidence un phénomène discontinu. En fonction de l'échelle temporelle de cumul, la précipitation peut être considérée comme un processus continu ou bien comme un processus discontinu. Pour l'étude de cumul saisonnier de pluie, on peut admettre la continuité spatiale du processus de pluie. Alors que pour de très petits pas de temps (échelle de temps inférieure ou égale à la durée de l'événement pluvieux), la continuité spatiale de la précipitation n'est pas facile à admettre (Desbordes, 1987; Bacancourt, 1990; Bacancourt et al., 1992). Dans la plupart des applications de ces techniques pour l'estimation de la précipitation, où de bons résultats ont été obtenus, les pas de temps de cumul étaient supérieurs ou égaux au pas de temps de l'événement pluvieux (Creutin et Obled, 1982; Bastin et al., 1984; Tabios et Salas, 1985). Pour des pas de temps inférieurs à la durée de l'événement pluvieux, l'intermittence est présente et l'utilisation de ces techniques n'est pas sans créer des problèmes de modélisation de la fonction de structure (Bacancourt, 1990).

Dans le cas d'un processus intermittent, Bacancourt (1990) propose de décomposer la variable aléatoire précipitation en un produit de deux variables aléatoires indépendantes: une variable aléatoire indicatrice de valeur 0 ou 1 (0 s'il n'a pas plu au point i et 1 s'il a plu) et une variable aléatoire conditionnelle donnant la hauteur de la pluie là où il a plu. L'approche consiste par la suite, en supposant les deux variables aléatoires indépendantes, à les kriger de façon indépendante (Bacancourt et al., 1992; Pawlowsky et al., 1993). L'interpolation des indicatrices par krigeage se nomme dans la littérature krigeage des indicatrices (Bacancourt et al., 1992; Dowd, 1992; Solow, 1993). Cette approche soulève un autre problème qui est celui de la délimitation de la surface pluvieuse par le krigeage des indicatrices à partir des données d'un réseau pluviométrique (Pawlowsky et al., 1993; Solow, 1993). L'utilisation des données de télédétection (radar) peut contribuer à diminuer l'erreur associée à cette délimitation.

Dans le cas où plusieurs réalisations spatiales sont disponibles (cas de plusieurs champs de pluie), on a recours aux modèles d'estimation spatio-temporelle. Les données disponibles peuvent être un ensemble des champs pluvieux caractérisant les cumuls événementiels pour une saison de pluie donnée. Dans ce cas le transfert d'information se fera

généralement de façon indirecte entre les champs. Les méthodes correspondantes sont dites globales. On distingue également parmi les méthodes globales, des méthodes déterministes et stochastiques. Les modèles de transfert direct sont généralement considérés pour un transfert entre les champs caractérisant un même événement pluvieux. Ils sont traités à la section 2.4.

2.2.3 Méthodes de transfert indirect ou méthodes globales

Au niveau des modèles globaux ou de transfert indirect, les données provenant de plusieurs événements pluvieux sont considérées, afin d'estimer une structure moyenne d'ensemble. Il est indirectement supposé au niveau de ce genre de modèles que ces événements pluvieux (caractéristiques à modéliser) appartiennent à une même population statistique. Ces modèles sont aussi appelés modèles climatologiques. Pour l'estimation de la pluie, trois modèles climatologiques sont généralement utilisés. Il s'agit du krigeage climatologique introduit par Bastin et al. (1984) et Lebel et Bastin (1985), de la méthode de seuillage introduite par Doneaud et al. (1981; 1984) et de la méthode de fonction Lagrangienne proposée par Braud et al. (1992) pour l'estimation de la pluie aux faibles pas de temps.

Pour ce qui est du krigeage climatologique, il est supposé que les champs possèdent une structure spatiale invariante d'un champ à un autre. Soit L_1 processus spatiaux $R(i,l)$ avec $(l=1, \dots, L_1)$, observés en N points $(i=1, \dots, N)$. Dans l'approche du krigeage climatologique, on suppose que

$$R(i,l) = a(l)W(i) + b \quad (2.3)$$

où $W(i)$ est un processus spatial invariant d'un champ à un autre, $a(l)$ est une constante fonction du champ l et b un processus aléatoire d'espérance nulle et de variance σ_b^2 . Avec cette décomposition, on démontre très facilement sous certaines hypothèses d'indépendance que le variogramme du processus $R(i,l)$ pour une valeur l donnée est aussi décomposable, et est donné par (Bastin et al., 1984; Lebel et Laborde, 1988) :

$$\gamma_R(h,l) = a(l)^2 \gamma_W(h) + \sigma_b^2 \quad (2.4)$$

où $\gamma_W(h)$ est le variogramme de $W(i)$. Il est appelé variogramme climatologique. C'est au niveau de la modélisation de cette structure que les données de plusieurs événements pluvieux peuvent être utilisées.

La méthode de seuillage est utilisée pour l'estimation de la pluie moyenne. Elle découle de l'observation faite par

Doneaud et al. (1981; 1984), selon laquelle pour une zone d'étude donnée de grande étendue, la pluie moyenne d'un événement pluvieux est linéairement corrélée au pourcentage de la superficie couverte par une pluie d'intensité supérieure ou égale à un seuil donné. Cette corrélation croît avec la valeur du seuil jusqu'à atteindre un maximum et puis décroît. La modélisation de cette linéarité est faite en considérant tous les événements observés. Après l'explication théorique de cette linéarité donnée par Kedem et al. (1990a) et Atlas et al. (1990), cette méthode a fait l'objet d'un certain nombre d'études et d'applications (Kedem et al., 1990b; Rosenfield et al., 1990; Braud et al., 1993; Short et al., 1993; Sauvageot, 1994). Kedem et Pavlopoulos (1991) ont présenté une approche pour le choix du seuil optimum. Tout récemment, Morrisey et al. (1994) ont eu à présenter une application de cette méthode pour l'estimation de la pluie dans le temps. Aussi Morrisey (1994) a pu montrer que les résultats de cette méthode sont fonction de la résolution des données disponibles (cas des données de télédétection). De toutes ces applications, l'hypothèse que les événements pluvieux sont caractérisés par la même loi de distribution a été considérée. Ce qui est n'est pas toujours vrai, surtout pour le Sahel où les pluies proviennent de différents types de systèmes météorologiques (convectives isolé et organisé et ligne de grain). Une classification judicieuse des événements

pluvieux ne pourrait-elle pas améliorer l'utilisation de cette méthode ?

La méthode de fonction lagrangienne (Braud et al., 1992) est aussi utilisée pour l'estimation de la pluie moyenne pour des pas de temps petits. La base de la méthode provient du fait qu'aux faibles pas de temps, les champs pluvieux possèdent une structure de cellule, c'est-à-dire que la hauteur de pluie (ou l'intensité) croît des bords au centre du champ comme dans le cas d'une cellule pluvieuse. La méthode consiste à modéliser cette structure en intégrant tous les champs disponibles (hauteur de pluie versus distance par rapport aux bords de la surface pluvieuse). Une fois cette structure modélisée, on peut estimer par la suite pour chaque champ sa pluie moyenne. Ici aussi, il est supposé de façon indirecte, tout comme dans le cas du krigeage climatologique que les champs possèdent des structures spatiales semblables. Ce qui n'est toujours pas le cas.

Ainsi, toutes ces trois méthodes cachent la notion de classification du moment où elles sont basées sur l'existence d'une certaine similarité structurale entre les différentes observations spatiales. On peut donc se poser la question de savoir si la classification des champs sur la base des paramètres pertinents peut-elle contribuer à améliorer

l'estimation par les modèles globaux ?

2.3 TYPOLOGIE DES SYSTÈMES PRÉCIPITANTS

Au cours de la construction des modèles de précipitations, l'hydrologue se trouve parfois avec une importante quantité de données spatio-temporelles. Dans la plupart des modèles, l'hydrologue combine toutes ces données de pluie provenant des différents événements sans se soucier parfois de leur appartenance météorologique. En d'autres termes tous les événements sont supposés provenir d'une même population, ce qui est parfois loin d'être le cas. L'exemple typique de ce genre de modèle est donné par les courbes "Intensité-Durée-Fréquence (I-D-F)" où toutes les données provenant de différents événements pluvieux sont combinées. Aussi, les modèles globaux qui sont le krigeage climatologique, les méthodes de seuil et de fonction lagrangienne sont dans cette catégorie.

La prise en compte d'une typologie dans la modélisation de la précipitation a fait l'objet de peu de recherche. Ce n'est que tout récemment que son importance a été démontrée par certains chercheurs (Bastin et al., 1984; Desbordes, 1987; Kottegoda et Kassim, 1991). Le besoin de classifier les événements pluvieux existe aussi au niveau des

méthodes d'estimation de pluie par satellite où une distinction entre les pluies convectives et stratiformes peut contribuer à améliorer la qualité de l'estimation (Barret et Martin, 1981). La typologie de systèmes précipitants nécessite la définition des critères rigoureux. Ces critères vont dépendre de la nature du modèle, de la nature des données disponibles et aussi de l'objectif de l'étude. C'est ainsi que les critères pour les modèles temporels seront différents de ceux des modèles spatiaux. Par exemple, pour les modèles temporels d'événement pluvieux, ces critères peuvent être: l'intensité maximale pour une durée donnée, la durée de la pluie, la hauteur totale de la pluie, .etc.

Dans le cas des modèles spatiaux, il faut définir des critères qui tiennent compte principalement du caractère spatial des précipitations. Une fois les critères bien définis, on peut utiliser n'importe quelle méthode de groupement (clustering) sur ces derniers. Mais l'important ici, est de définir des bons critères. Une stratification des systèmes précipitants basée sur des paramètres météorologiques (champs de pression et de vitesse) dans le cas où leur mesure existe, est la plus objective. La classification météorologique a été utilisée dans la plupart des techniques de transfert de variations climatiques de grande échelle aux échelles régionale et locale (techniques

de "Down scaling"). La classification est utile pour la simulation des impacts de la variation climatique sur certaines variables telles que les précipitations (Hay et al., 1991; 1992; Bardossy et Plate, 1992; Wilson et Lettenmaier, 1992). Dans la plupart des cas, les données météorologiques sont inexistantes, d'où la nécessité d'avoir recours aux données sol.

Kottegoda et Kassim (1991) ont proposé une technique de classification utilisant les propriétés des "crossing". Cette classification repose sur la structure (forme) de la courbe adimensionnelle "pourcentage du cumul de la précipitation en fonction du pourcentage de la durée de l'orage" ou plus connue sous l'appellation courbe de masse. La Figure 5 de l'annexe 4.1 donne les trois premières formes possibles de cette courbe. Un orage possédant une forme d'ordre supérieur à 3, est un orage intermittent en termes de cellules pluvieuses. La méthode telle que présentée dans Kottégoda et Kassim (1991) se limite à la classification des événements ponctuels. Mais par contre, elle peut être modifiée afin de pouvoir classifier de façon systématique des événements pluvieux observés par un réseau de mesure par le biais du hyétogramme moyen (Amani et al., 1993a).

Dans ses travaux, Thauvin (1992) a eu à classifier les événements pluvieux pour fin de modélisation du variogramme climatologique. Cette classification est basée sur le pourcentage de stations touchées par un événement pluvieux. A notre avis, ce paramètre n'est pas très pertinent (le pourcentage de stations touchées ne peut caractériser à lui seul la structure spatiale d'un événement pluvieux). Un bon critère de classification de données pluviométriques régionales, doit tenir compte de la distribution spatiale des événements pluvieux. Les paramètres de classification doivent nécessairement caractériser la structure spatiale et/ou spatio-temporelle de la pluie. Signalons qu'on distingue dans la littérature deux familles de méthodes de classification: méthodes hiérarchiques et non hiérarchiques. La différence principale entre ces familles réside au niveau de la façon dont se produit les groupements (Anderberg, 1973).

2.4 DÉPENDANCE SPATIO-TEMPORELLE ET DYNAMIQUE DE LA PLUIE

2.4.1 Estimation spatio-temporelle

L'étude spatio-temporelle des hauteurs de pluie cumulée pour des faibles pas de temps au cours d'un événement pluvieux, montre qu'il existe bel et bien une dépendance entre les données spatiales observées en des pas de temps

différents. Dans cette situation, le fait d'estimer un champ en n' utilisant uniquement que les données spatiales correspondant à ce dernier néglige une bonne partie de l'information. Ainsi on peut penser que l'introduction des données spatiales des autres champs pour l'estimation d'un champ donné contribuera à une utilisation optimale des données disponibles et à améliorer éventuellement l'estimation.

Au niveau des méthodes de transfert direct, un poids est associé de façon explicite à toutes observations spatiales appartenant aux différents champs, contrairement aux méthodes de transfert indirect d'informations temporelles. De cette façon, le transfert d'informations se produit directement à travers la fonction de covariance spatio-temporelle. Nous pouvons diviser les méthodes de transfert direct en trois familles: cokrigeages, krigeages spatio-temporels 3D et modèles ARIMA multisites (STARIMA). Contrairement à la première famille où le transfert se fait par le biais des variogrammes croisés dans l'espace 2D, au niveau de la deuxième famille, on considère le temps comme une troisième dimension. Même si les méthodes de la troisième famille sont développées à des fins de prévisions, elles constituent un très bon exemple de transfert spatio-temporel d'informations et de plus peuvent être utilisées

pour compléter des données manquantes (Bras et Rodriguez-Iturbe, 1985).

. Cokrigeages

La technique du cokrigeage est une technique de transfert d'informations entre deux ou plusieurs processus ayant une certaine dépendance entre eux, en vue d'améliorer l'estimation de l'un des processus. Le transfert d'informations se fait par le biais des variogrammes croisés des différents processus. De même la stationnarité des moments croisés d'ordre 2 est considérée. Le cokrigeage, à cause de son intérêt pratique, a fait l'objet de beaucoup de travaux (Meyers, 1981; Creutin et al., 1989; Seo et al., 1990a,b). L'estimation de la pluie à partir des données sol et des données de télédétection (images radar et satellitaires) constitue les applications les plus intéressantes du cokrigeage en hydrologie (Andrieu, 1986; Delrieu et al., 1988; Creutin et al., 1989; Seo et al., 1990a,b). Tout comme au niveau du krigeage, il existe également plusieurs types de cokrigeages linéaires.

Bien que le cokrigeage permette l'estimation des processus ayant une certaine dépendance, il a des limites pratiques. La mise en oeuvre du cokrigeage nécessite la modélisation de $L(L-1)/2$ fonctions de structure pour L champs

donnés. Aussi on n'a aucune garantie sur la positivité de la variance d'estimation. Dans la même famille du cokrigeage, signalons l'existence du krigeage multivarié ou vectoriel (Rouhani et Wackernagel, 1990; Bourgault et Marcotte, 1991) où une seule fonction de structure généralisée est construite pour l'ensemble des champs. On démontre facilement sous certaines hypothèses que le krigeage vectoriel se réduit au krigeage climatologique. Tout comme les modèles STARIMA, le cokrigeage ne peut être utilisé que pour l'estimation pour de pas de temps discrets. Dans le cas contraire, il faut avoir recours au krigeage spatio-temporel.

. Krigeages spatio-temporels

De toutes les techniques de transfert d'informations spatio-temporelles, les techniques de krigeage spatio-temporel semblent être les plus directes du moment où l'interaction entre le temps et l'espace est considérée dans un espace 3D (x, y, t) . Comme le krigeage peut être généralisé dans un espace à n dimensions, le krigeage 2D (x, y) peut être étendu dans le domaine espace temps, assurant ainsi le transfert d'informations spatio-temporelles. Tout comme le krigeage 2D, le krigeage spatio-temporel est caractérisé par une fonction de structure, cette fois-ci spatio-temporelle. Dépendant de l'interprétation de cette fonction de structure spatio-temporelle, on peut diviser en deux grandes familles

les krigeages spatio-temporels: i) famille des krigeages 3D caractérisée par une fonction de structure variogramme $\gamma(h)$ avec h distance euclidienne et ii) famille des krigeages 3D caractérisée par une fonction de structure $\gamma(h, \tau)$ où h et τ sont des distances euclidiennes dans l'espace et dans le temps.

Dans l'approche de krigeage 3D caractérisée par une fonction de structure $K(h)$, le temps est considéré comme une troisième dimension au même titre que les 2 dimensions spatiales x et y . Les données spatio-temporelles sont supposées être une réalisation d'un processus aléatoire à trois dimensions $R(x,y,t)$. Il suffit de généraliser le krigeage 2D en 3D sans se soucier de l'importance des différentes dimensions. Pour les équations du krigeage spatio-temporel, on se référera à l'article présenté à l'annexe 5.1. Pour ce qui est de la modélisation de la pluie, cette technique de transfert direct n'est pas sans créer certains problèmes d'ordre physique.

L'application de cette technique pour la pluie, suppose que le phénomène de pluie a le même comportement dans l'espace et dans le temps. Ce qui n'est pas toujours le cas (Zawadzki, 1973b). Aussi la distance euclidienne 3D dans l'espace et le temps n'a aucun sens physique. Des approches

pour palier à ces deux contraintes ont été proposées dans l'article présenté à l'annexe 5.1. Une façon directe de contourner les deux contraintes du krigeage spatio-temporel 3D avec $K(h)$ pour l'estimation de la pluie, est de considérer que la fonction de structure, covariance ou variogramme, est une fonction explicite de la distance métrique h et de la distance temporelle τ (Rouhani et Hall., 1989; Rouhani et Meyers, 1990; Rouhani et Wakernagel, 1990). La pratique consistant à décomposer la fonction de structure $K(h, \tau)$ sous forme d'une somme ou d'un produit de deux fonctions temporelle et spatiale (Rodriguez-Iturbe et Mejia, 1974a,b; Rouhani et Hall, 1989) n'est pas sans créer des problèmes d'ordre numérique. Dans leur travaux, Rouhani et Meyers (1990) et Meyers et Journel (1990) ont fait ressortir par des exemples, le fait que cette pratique peut conduire dans certains cas à un système de krigeage singulier ou à une instabilité du système de krigeage à cause du mauvais conditionnement de la matrice du système.

2.4.2 Dynamique du système précipitant

Aux faibles pas de temps, le phénomène de pluie est un processus dynamique. Celui-ci est caractérisé par le déplacement en fonction du temps du système de pluie à travers la zone d'étude, ici le réseau de pluviographes.

Cette dynamique est mise en évidence en temps réel beaucoup plus dans le cas des données de télédétection (données radar). L'intérêt de la dynamique pour la prévision des pluies aux faibles pas de temps à partir des données radar a été démontré tout récemment par les études de Bellon et Zawadzki. (1993); Brémaud et Pointin. (1993) et Zawadzki et al. (1994). La caractéristique de la dynamique, qui est la vitesse, s'obtient beaucoup plus facilement à partir des données radar. Aussi dans le cas où, l'observation radar n'est pas disponible, les données spatio-temporelles de pluie collectées par un réseau de pluviographes peuvent servir à estimer la vitesse du système précipitant.

. Estimation de vitesse à partir des données sol

Il existe deux familles de méthodes pour l'estimation de la vitesse d'un système pluvieux à partir des données d'un réseau pluviographique. La première famille de méthodes se base sur l'analyse spatiale des temps d'observation au niveau des différentes stations, d'une des caractéristiques du hyétogramme, début de la pluie, intensité maximale ou centre de gravité de la pluie (Johnson et Bras, 1979; Niemczynbicz et Dahlblom, 1984; Diskin, 1987; 1990). Pour plus de détails sur ces méthodes, voir annexe 5.3. La deuxième famille des méthodes proposée par Marshall (1980) est quant à elle basée sur l'analyse de la structure spatio-temporelle du champ de

pluie aux différents pas de temps. A chaque décalage temporel, une surface de corrélation lui est associée. Celle-ci est donnée par la fonction de corrélation croisée dans le plan de coordonnées spatiales (distances entre les points). Chaque surface de corrélation est caractérisée par le point correspondant à la corrélation maximale. A cause de la dynamique de la pluie, les positions des points à corrélation maximale se déplacent avec les décalages temporels, tout en restant sur une même droite (Marshall, 1980). La vitesse se déduit directement à partir des caractéristiques de cette droite. Cette notion nous sera très utile pour la mise en oeuvre de la méthode du krigeage dynamique (chapitre 5).

. Dynamique de pluie et hypothèse de Taylor

L'énoncé de l'hypothèse de Taylor date des études entreprises par ce dernier en 1938 sur les écoulements turbulents. Il suppose que, pour un champ turbulent ayant une vitesse de déplacement d'ensemble supérieure à celle d'un tourbillon, les changements dans le temps en un point fixe sont produits par le passage d'une configuration spatiale qui ne subit pas de modification essentielle au cours de son déplacement. Cette hypothèse formulée pour les précipitations pour la première fois par Zawadzki (1973b) est la suivante: les événements pluvieux sont statistiquement "stationnaires" dans la mesure où on peut déduire les caractéristiques de la

variabilité spatiale à partir des caractéristiques de la variabilité temporelle en un point, simplement en convertissant le pas de temps en pas d'espace par le biais de la vitesse de déplacement du système pluvieux. Zawadzki (1973b) a montré à partir des données radar de pluie de la région de Montréal que cette hypothèse est limitative du moment où elle dépend de la durée de vie du système. Cette hypothèse est vérifiée pour un pas de temps inférieur à 35 min. La formulation mathématique de cette hypothèse est la suivante (Gupta et Waymire, 1987): soit $p(x,t)$ un processus spatialement homogène et temporellement stationnaire, avec t le temps et x l'espace. Ce dernier satisfait l'hypothèse de Taylor si la corrélation spatio-temporelle

$$\rho(h,\tau) = \text{corr}[p(x,t), p(x+h, t+\tau)] \quad (2.5)$$

vérifie la relation suivante:

$$\rho(0,\tau) = \rho(-u\tau, 0) \quad (2.6)$$

avec u vecteur vitesse et h et τ étant des distances spatiale et temporelle. Dans le cas de la précipitation u n'est rien d'autre que la vitesse du système pluvieux.

D'après Zawadzki. (1987), il existe une certaine contradiction entre la Théorie de l'hypothèse de Taylor qui est admise pour la construction de certains modèles

concepts de précipitations (Waymire et Rodriguez-Iturbe (1984); Cox et Isham. (1988) et Rodriguez-Iturbe et al., 1986; 1987b; 1988) et l'utilisation grandissante de modèles fractaux pour la caractérisation de la variabilité des précipitations (Hubert et Carbonnell, 1988b, 1989; Beer, 1990; Gupta et Waymire., 1990; Moussa, 1991).

2.5 CONCLUSION

Il existe une abondante littérature pour ce qui est des méthodes de la géostatistique. L'estimation de précipitations par ces méthodes a été dans la plupart des cas effectuée pour des grands pas de temps. Ce qui faisait que ces méthodes donnaient des bons résultats comparés aux méthodes déterministes. Par contre dès à partir de l'échelle de l'événement pluvieux, l'apparition de l'intermittence au niveau des champs de pluie aux faibles pas de temps nous oblige à nous poser des questions quant à la pertinence de l'application des méthodes géostatistiques. Que valent ces méthodes à des petites échelles de temps ?

La généralisation du krigeage dans l'espace temps est très séduisante, mais permet-elle d'améliorer l'estimation aux faibles pas de temps ? Aussi à ces échelles de temps, la dynamique du système pluvieux a une place de choix. Il existe

peu de travaux qui se sont penchés sur ces questions. Les méthodes globales d'estimation développées pour palier d'une certaine façon au problème d'intermittence, souffrent sans aucun doute d'un manque de stratification des événements pluvieux pouvant contribuer à l'amélioration de l'estimation. Aussi peu de travaux se sont intéressés à la question de classification de événements pluvieux et surtout à sa pertinence pour l'estimation des précipitations. Nos travaux tentent de donner des éléments de réponses à chacune de ces questions. Sur la base des données EPSAT-NIGER, l'applicabilité des méthodes de la géostatistique sera évaluée aux différents pas de temps. Par la suite, des nouvelles méthodologies seront proposées d'une part pour l'estimation à l'échelle de l'événement par les modèles globaux par la prise en compte de la classification des événements et d'autre part pour l'estimation aux faibles pas de temps en tenant compte de la dynamique.

CHAPITRE III: TESTS ET PROBLÈMES RENCONTRÉS POUR
L'ESTIMATION DES PLUIES PAR LES MODÈLES
GÉOSTATISTIQUES

3.1 INTRODUCTION

Les modèles issus de la géostatistique développés initialement pour l'estimation des réserves minérales (Matheron, 1970; Journel et Huijbregts, 1978) ont vite été considérés en hydrologie comme étant des outils pouvant remplacer les techniques déterministes jusqu'alors utilisées pour l'estimation spatiale de la pluie (Delhomme, 1978; Creutin et Obled, 1982; Chua et Bras, 1982; Lebel et al., 1987, entre autres). Les différentes applications de ces techniques pour l'estimation de la pluie moyenne ou ponctuelle ont montré de façon générale, la supériorité de ces techniques par rapport aux techniques déterministes pour des pas de temps supérieurs à ceux de l'événement (Creutin et Obled, 1982; Tabios et Salas, 1985). Cette supériorité s'amointrit à mesure que la densité du réseau augmente (Lebel et al., 1987).

Grâce aux données de l'expérience EPSAT-NIGER, nous sommes maintenant en mesure de tester sur la région sahélienne les techniques issues de la géostatistique pour

l'estimation de la pluie pour des pas de temps allant de la saison aux très petits pas de temps de cumul. De toute évidence, les caractéristiques de la pluie sahélienne vont beaucoup influencer la performance des modèles issus de la géostatistique. Nous essayerons par la suite de synthétiser sur la base de nos travaux préliminaires et ceux des autres membres de l'équipe EPSAT-NIGER les principales limitations concernant l'utilisation directe des techniques classiques de la géostatistique pour l'estimation de la pluie au Sahel (Thauvin, 1992; Taupin et al., 1993; Lebel et al., 1995a; Lebel et Lebarbé, 1995).

3.2 CARACTÉRISTIQUES GÉNÉRALES DES PLUIES SAHÉLIENNES

Au Sahel, trois types de systèmes pluvieux sont communément distingués: systèmes convectifs isolés, systèmes convectifs organisés et lignes de grains. Ces deux derniers types peuvent être regroupés sous l'appellation de systèmes convectifs de méso-échelle (SCM). La contribution de chacun de ces systèmes au cumul total annuel n'est pas très bien connue et varie probablement avec la latitude considérée. Il est par contre reconnu que les lignes de grain même si elles sont très faibles en nombre contribuent pour beaucoup au cumul annuel. L'identification d'un événement pluvieux à partir de sa signature au sol est une question qui a suscité

peu d'études jusqu'à aujourd'hui, bien qu'elle soit indispensable pour améliorer notre connaissance du régime pluviométrique sahélien.

Avant l'expérience EPSAT-NIGER, notre connaissance de la répartition spatiale de la pluie au sahel se limitait principalement à la parfaite organisation des isohyètes des moyennes interannuelles sur le Sahel avec un gradient Nord-Sud sensiblement constant de 1mm/km (Figure 3.1). Cette caractéristique de pluie sahélienne est loin d'être valable pour des échelles spatiales inférieures ou égales à celle du réseau EPSAT-NIGER que ça soit sur une échelle annuelle ou inférieure. Les figures 3.2a,b et c montrent que ce gradient n'est pas présent sur les cumuls annuels, par contre la figure 3.2d qui représente la distribution des cumuls moyens annuel sur les trois années présente un certain gradient Nord-Sud. Aussi il est connu qu'aux petites échelles de temps, la pluie au Sahel est aléatoire ou discontinue et très variable spatialement. L'étude de cette variabilité n'a jamais été entreprise faute des données appropriées. Les systèmes de type ligne de grains se déplacent de façon régulière en général d'Est en Ouest avec des vitesses moyennes comprises entre 40 et 70 km/h. Les données EPSAT-NIGER nous permettent maintenant de le confirmer (un exemple est donné en figure 3.3).

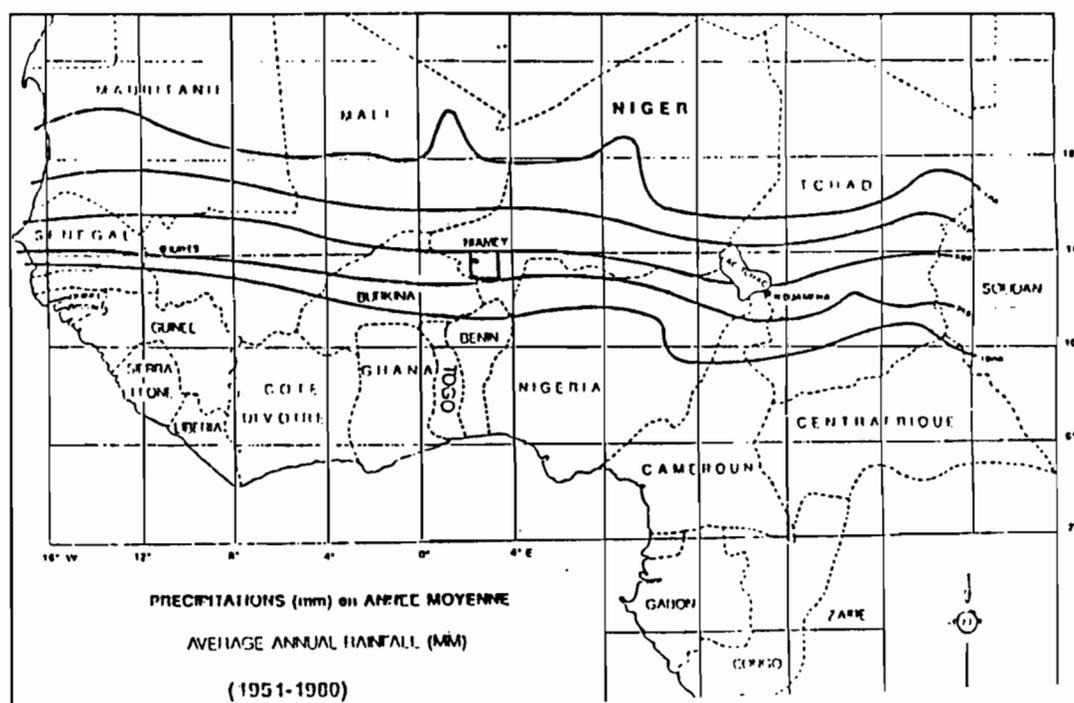


Figure 3.1: Illustration du gradient climatologique Nord-Sud au Sahel (isohyètes moyennes annuelles)

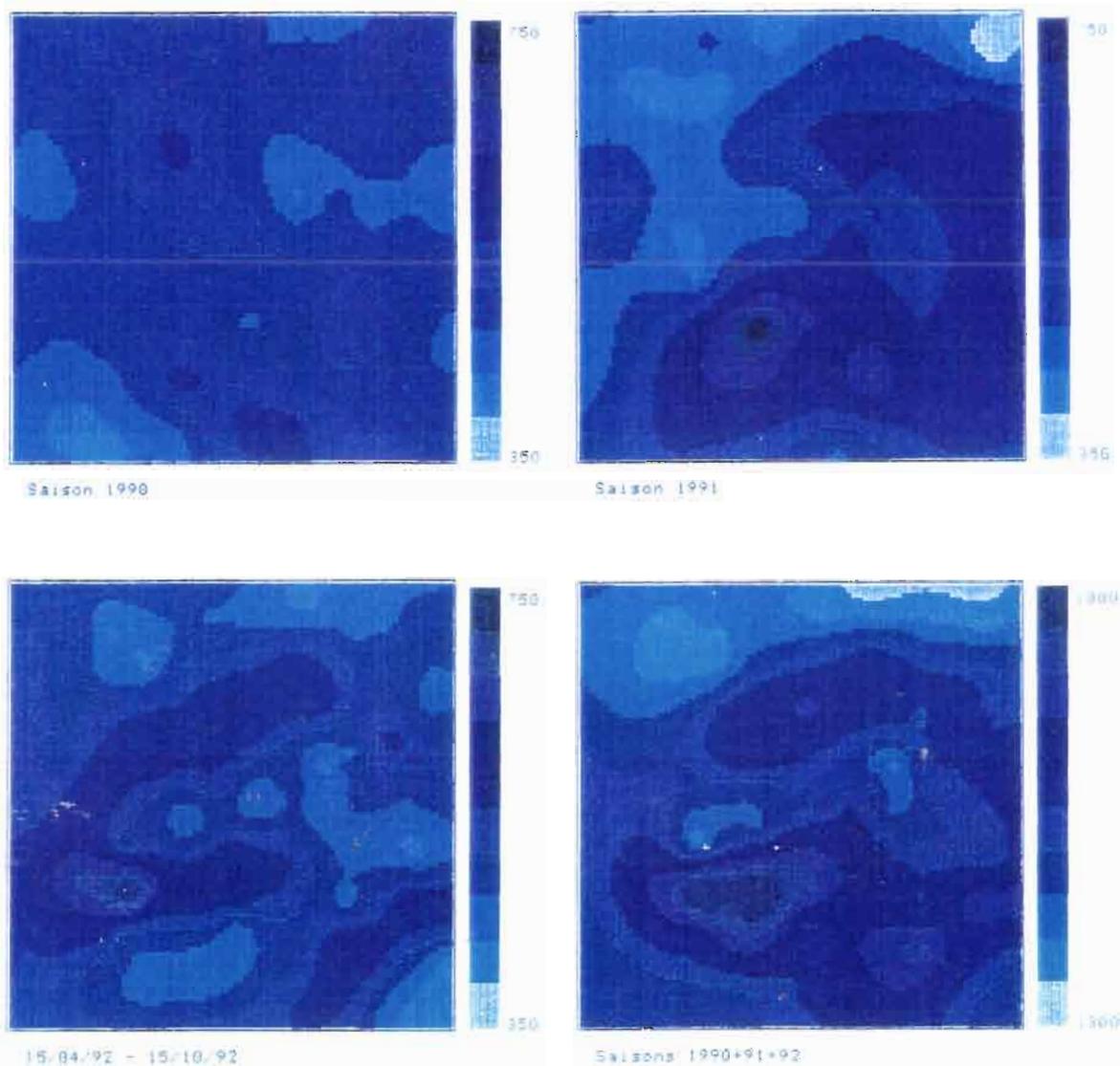


Figure 3.2: Répartition de la pluie annuelle sur le réseau EPSAT-Niger (Taupin et al., 1992)

3.3 APPLICATION DE LA GÉOSTATISTIQUE AUX PLUIES SAHÉLIENNES

3.3.1 Considérations générales

Les techniques de la géostatistique étant bien adaptées pour des processus stationnaires et continus, elles ont été dans un premier temps considérées pour l'estimation spatiale des cumuls saisonniers et événementiels. Il est vite apparu que l'événement pluvieux constitue sans aucun doute la pierre angulaire de l'estimation des cumuls saisonniers. En effet les champs décennaires, mensuels et saisonniers ne sont que des variantes d'un même phénomène aléatoire: la superposition de n événements pluvieux. Pour une décennie d'août n peut être plus grand que pour le mois de juin par exemple. Ainsi étudier le champ pluvieux saisonnier, revient à étudier la superposition de n événements pluvieux où n est suffisamment grand. Les champs décennaires et mensuels ne sont, vus sous cet angle, que des sous-cas avec n plus petit.

Pour des pas de temps de cumul inférieurs à celui de l'événement, on change de processus aléatoire et a priori les techniques de la géostatistique ne sont plus bien adaptées pour trois principales raisons: (1) la forte intermittence des champs pluvieux, celle-ci augmentant avec la diminution du pas de temps de cumul; (2) l'existence d'une

interdépendance du phénomène dans l'espace et le temps; celle-ci mérite considération; (3) la dynamique des champs pluvieux, qui implique une dépendance temporelle évidente des champs pluvieux d'un pas de temps à l'autre.

Malgré tout, on montrera dans cette thèse que moyennement quelques modifications, on peut encore avoir recours à une partie de l'approche géostatistique pour estimer et interpoler les champs pluvieux sahéliens lorsque le pas de temps de cumul est inférieur à celui de l'événement pluvieux.

3.3.2 Cumuls saisonniers et par événement

Les cumuls saisonniers et par événement collectés lors de l'expérience EPSAT-NIGER, ont fait l'objet d'un certain nombre d'études visant à caractériser leur structure spatiale pour une estimation par la géostatistique. Les premiers travaux qui sont ceux de Thauvin et Lebel (1991) et Thauvin (1992) montrent que la portée du variogramme saisonnier est du même ordre de grandeur que celle du variogramme climatologique des événements pluvieux de grande extension spatiale (ceux qui touchent plus de 70% des stations en fonctionnement).

Pour les années 89, 90 et 91, Taupin et al. (1993) montrent que le variogramme saisonnier présente effectivement une portée pour une valeur de l'ordre de 30 km, ce qui est confirmé par Lebel et al. (1995b) pour les données des saisons 92 et 93. Il s'agit bien d'un caractère stable et significatif. Par ailleurs, Lebel et al. (1995a) montrent qu'en sélectionnant de façon plus exigeante les événements pluvieux, en ne gardant par exemple que le tiers des événements pluvieux qui touchent plus de 90% des stations, la portée du variogramme du cumul de ces événements n'augmente pas, par comparaison avec ce qu'elle est pour le cumul de tous les événements. Ce n'est donc pas le mélange d'événements plus ou moins intermittents qui est la cause essentielle de la faible portée des cumuls saisonniers.

Ce résultat peut sembler paradoxal eu égard aux travaux de Lebel (1984) pour les pluies des Cévènes où la portée des champs est fonction du pas de temps de cumul et augmente avec ce dernier. Mais il n'en est rien. Lebel et Lebarbé (1995) montrent qu'effectivement il existe un lien étroit entre la portée moyenne du champ par événement et celle du champ saisonnier. Le champ saisonnier présente une structure gigogne dont la partie supérieure est difficile à étudier du fait de la faiblesse des distances couvertes par le réseau EPSAT-NIGER.

Sur la base de ces différents travaux, on peut conclure que la géostatistique permet d'une certaine façon de bien rendre compte de la structure spatiale des cumuls saisonniers et par événement. Il reste malgré tout quelques questions à étudier sur ce terrain: (1) quelle est l'influence de l'intermittence sur l'inférence du variogramme et quels peuvent être l'intérêt et l'apport de la stratification des événements pluvieux à cette fin ? Nous avons étudié cette question dans cette thèse (voir chapitre 4 et annexe A1). L'autre question est de savoir quel est le réseau optimal pour échantillonner correctement le champ de précipitations sahéliennes. Cette question reste à étudier.

3.3.3 Pas de temps inférieur à l'événement

Aux pas de temps de cumul inférieur à l'événement, la forte intermittence des champs pluvieux, rend difficile l'application des techniques de la géostatistique. Cette situation est due au fait qu'à l'échelle de temps inférieur à celle de l'événement pluvieux, une autre caractéristique importante est présente: la dynamique du système pluvieux. Aussi, à ces deux caractéristiques, vient se greffer pour ce qui des lignes de grain sahéliennes, une anisotropie manifeste des champs pluvieux. Ces trois importantes caractéristiques des champs aux faibles de temps de cumul

sont illustrées à la figure 3.3. A cause de la dynamique des champs pluvieux, une quatrième caractéristique est à considérer: il s'agit de l'interdépendance temporelle des champs d'un pas de temps à un autre. Ainsi, l'utilisation des techniques classiques de la géostatistique sans aucune adaptation ne peuvent rendre compte de la structure spatio-temporelle des champs pluvieux à ces échelles de temps. Des nouvelles techniques sont proposées pour tenir compte dans l'estimation de ces différentes caractéristiques: krigeage spatio-temporel et krigeage dynamique lagrangien.

3.4 CONCLUSION

La géostatistique fournit un cadre facile à mettre en oeuvre pour interpoler les champs pluviométriques, mais ses hypothèses en limitent a priori l'utilisation. Au Sahel, la très forte intermittence des précipitations semblait être un facteur limitant important. Malgré cela Lebel et al. (1995b) ont montré que tout au moins pour les échelles de temps de l'événement et au-dessus, l'approche géostatistique permettrait tout à la fois: (1) d'inférer des variogrammes utilisables pour l'interpolation; (2) de retrouver dans ces variogrammes la trace des structures sous-jacentes aux champs de pluie saisonnier.

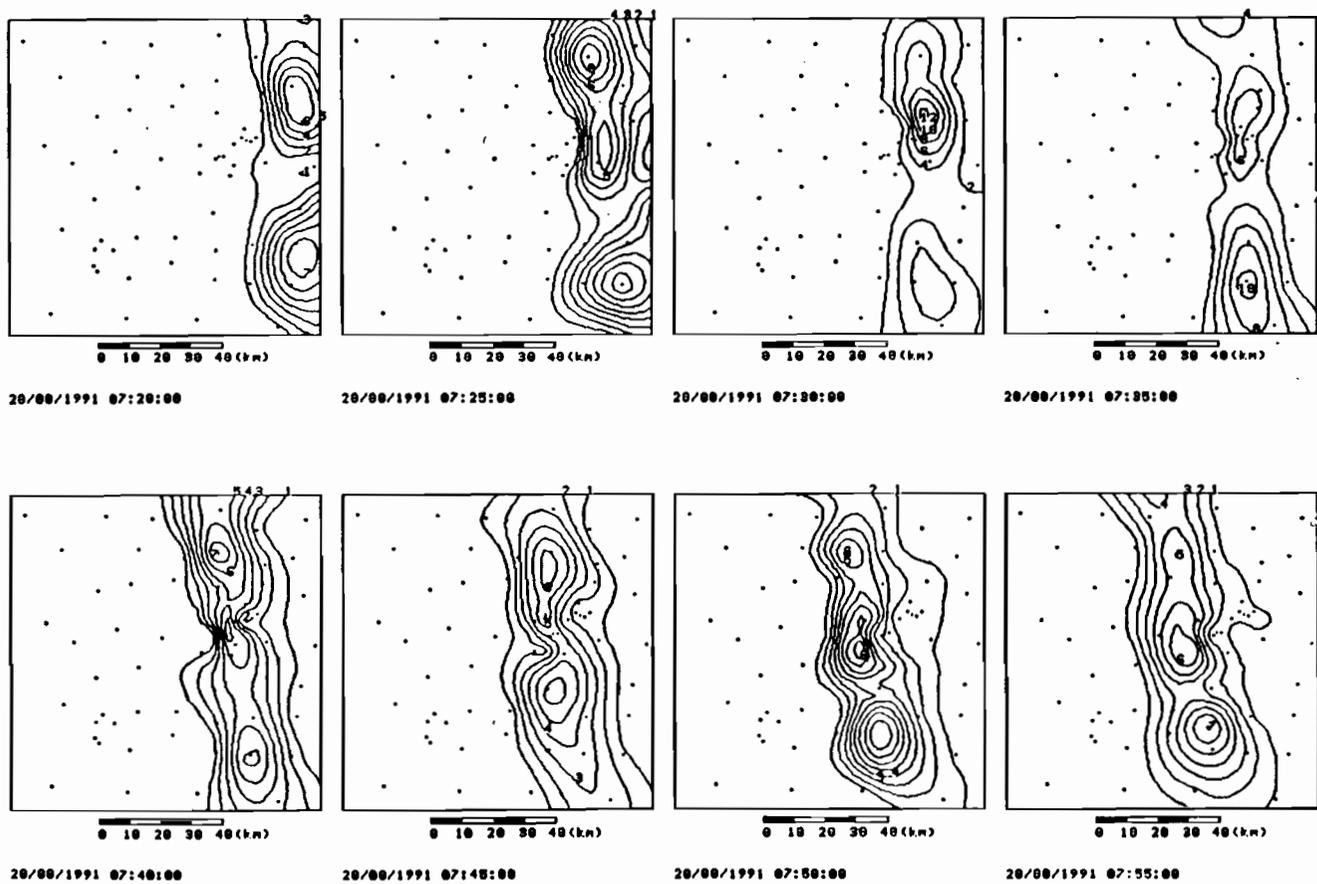


Figure 3.3: Champs d'isohyètes de la ligne de grains du 20/08/91
(Taupin et al., 1991).

Ces propriétés ont été exploitées par Thauvin (1992) et Lebel et LeBarbé (1995) pour calculer la précision d'estimation des pluies à l'échelle de la saison et de l'événement à partir de réseaux sol de différentes densités. Nous nous intéresserons donc dans la suite de cette thèse à deux points permettant de préciser ou d'étendre le champ d'application des techniques géostatistiques en matière d'estimation des pluies sahéliennes.

Tout d'abord, nous chercherons au chapitre 4 s'il existe des classes d'événements présentant des caractéristiques et distributions spatiales différentes d'un groupe à l'autre. Dans l'affirmative nous étudierons l'influence d'une telle classification sur les performances des modèles d'estimation.

Au chapitre 5, nous nous intéresserons aux pas de temps inférieurs à l'événement, non traités dans les travaux cités précédemment. A ces pas de temps, les champs pluviométriques ont des propriétés qui rendent en principe difficile, pour ne pas dire impossible, la mise en oeuvre des techniques géostatistiques:

- intermittence plus forte
- forte dynamique temporelle créant une interdépendance des champs pluvieux d'un pas de temps à l'autre

- non stationarité
- anisotropie.

Plusieurs algorithmes de krigeage spatio-temporel et de krigeage dynamique seront comparés pour prendre en compte ces propriétés dans l'estimation aux faibles pas de temps.

CHAPITRE IV: ESTIMATION SPATIALE A L'ÉCHELLE DE L'ÉVÉNEMENT:
CLASSIFICATION ET PRISE EN COMPTE DE
L'INTERMITTENCE

4.1 INTRODUCTION

On a vu au chapitre 3 que même à l'échelle de la saison, où les hypothèses de stationarité qui sous-tendent les méthodes de la géostatistique sont les mieux respectées, l'identification des fonctions de structure et l'interpolation par krigeage qui en résulte présente des difficultés. Néanmoins Lebel et LeBarbé (1995) ont montré que le variogramme saisonnier pouvait se déduire analytiquement du variogramme des cumuls par événement, ce qui place au premier plan la nécessité de modéliser correctement la structure spatiale des champs de précipitation à cette échelle de l'événement.

Or tous les événements observés au cours d'une saison des pluies ne sont pas météorologiquement identiques. Desbois et al. (1988) ont ainsi distingué trois grands types de systèmes convectifs: les convections isolées, les systèmes convectifs organisés et les lignes de grains. D'un autre côté, les cartes d'isohyètes calculées à partir des données EPSAT-NIGER se distinguent clairement (figure 4.1) en

fonction de leur degré d'intermittence. Une double question se pose donc:

- Peut-on établir un lien entre les caractéristiques spatio-temporelles des champs pluvieux et leur "signature" météorologique et en déduire ainsi une classification pertinente pour la mise en oeuvre d'algorithmes d'estimation, de type krigeage ou autre ?

-Comment prendre en compte l'intermittence dans ces algorithmes, sachant que très souvent ils présupposent une certaine stationnarité qu'exclut en principe un point d'accumulation de la distribution sur les valeurs nulles ?

On ne prétend pas véritablement traiter ici le premier point car l'approche objective consistant à classifier les événements pluvieux suivant leur environnement météorologique (locale, de méso-échelle ou synoptique) n'est pas actuellement à notre portée pour ce qui est du Sahel. Un travail dédié exclusivement à cette question serait nécessaire. On s'est donc limité à produire des méthodes de classification se basant sur des données pluviométriques. Ce n'est qu'une fois la classification opérée qu'on essayera d'interpréter en terme de caractéristiques météorologiques probables chaque groupe ainsi constitué.

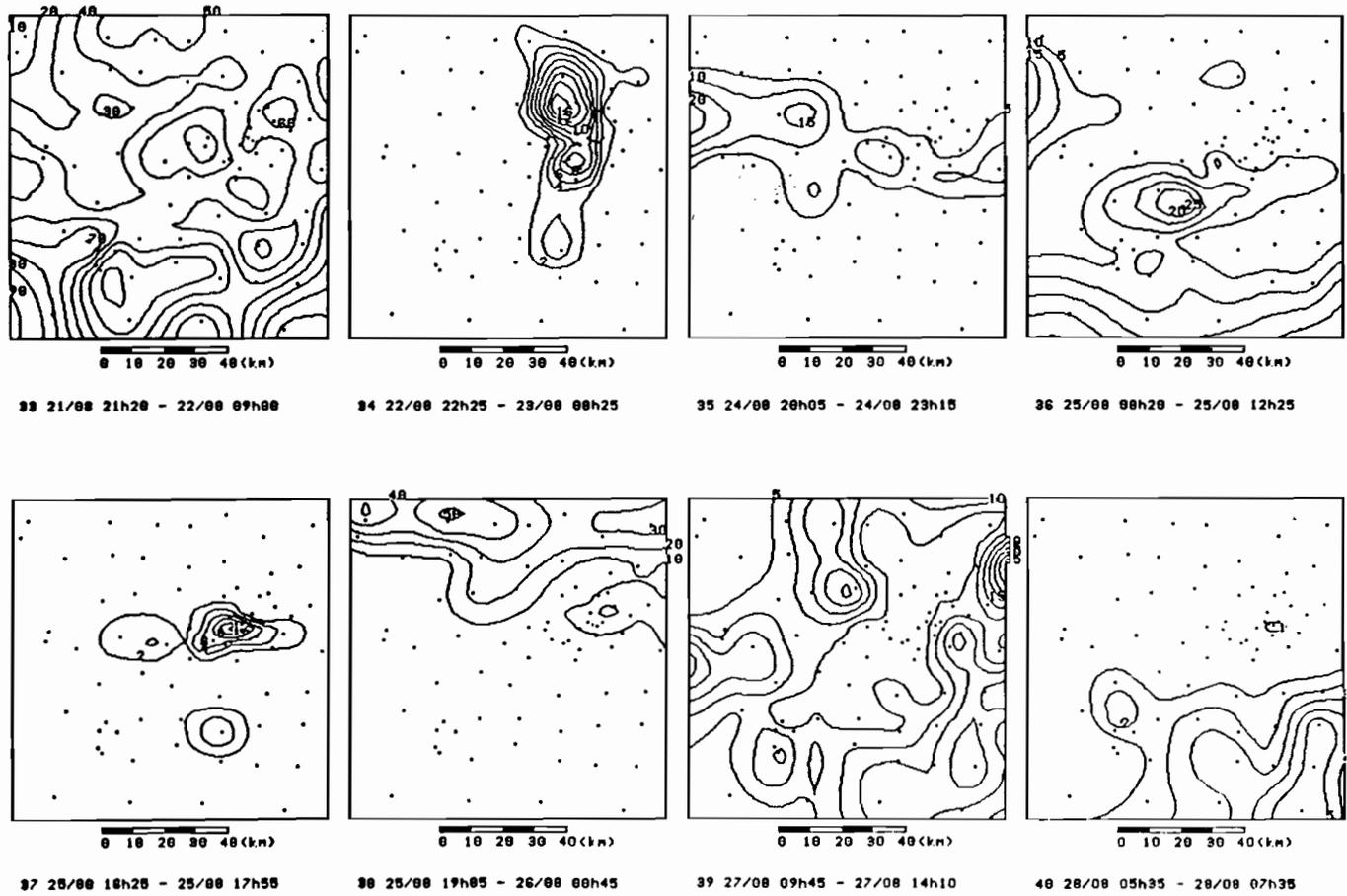


Figure 4.1: Exemple d'isohyètes de différents événements pluvieux observés durant EPSAT-NIGER (Taupin et al., 1992)

En ce qui concerne le rôle joué par l'intermittence spatiale (assimilée ici à la fréquence de pluie nulle dans le champ spatial ou à tout autre critère de distribution spatiale à définir) l'objectif ne sera pas de trouver une nouvelle formulation la prenant en compte explicitement. Plutôt, il s'agira d'étudier comment une classification intégrant implicitement un critère d'intermittence peut conduire à déterminer des paramètres de modèle d'estimation différents selon les groupes produits.

On examinera également, quel est l'apport de la classification proposée à la caractérisation du régime pluviométrique sahélien, tant sur le plan de la climatologie générale qu'au niveau d'une saison des pluies particulières.

Neuf méthodes de classification seront analysées et comparées, présentant des liens entre elles. Certaines de ces méthodes seront par la suite utilisées pour tester l'influence de la classification sur le calcul des paramètres de deux méthodes d'estimation: le krigeage climatologique et la méthode des seuils. Seule la synthèse de nos résultats et conclusions est présentée dans ce chapitre. Pour plus de détails on se rapportera à l'article Amani et al. (1995a) dont le texte est donné en annexe A1, ainsi qu'aux annexes B1 et B2 où figurent les résultats complets et certains

développement théoriques complémentaires.

4.2 MÉTHODES DE CLASSIFICATION SUR LA BASE DES DONNÉES SOL

Bien que l'importance de la classification des événements pluvieux pour la modélisation des précipitations à partir des données de télédétection (parties stratiforme et convective) ou de données pluviographiques ait été soulevée par certains chercheurs, celle-ci n'a fait l'objet à notre connaissance d'aucune étude systématique. Ceci est peut être dû à la difficulté de produire une méthode de classification des événements pluvieux valable non seulement pour les différentes régions climatiques mais aussi pour tout modèle de pluie.

Notre objectif est de définir des paramètres pertinents caractérisant la structure spatiale ou spatio-temporelle de la pluie servant de base à la classification et non pas de produire des algorithmes numériques de classification en dehors de ceux bien connus de "clustering". Dans certains cas, la technique de clustering sera appliquée sur les paramètres définis pour produire des groupes. Certaines méthodes peuvent ne pas avoir recours à la technique de clustering, elles sont dites systématiques.

On distingue deux approches pour définir, pour un modèle donné, des paramètres pertinents. La première consiste à les définir directement à partir de la théorie qui soutient le modèle en question. Cette approche est qualifiée de directe. L'autre façon consiste à définir des paramètres que l'on juge a priori caractéristiques de la structure à modéliser. C'est une approche heuristique dans laquelle, le choix des paramètres se pratique par essai-erreur.

Parmi les 9 méthodes de classification analysées ici, six utilisent la technique de "clustering" et les trois autres sont systématiques. Deux méthodes sont considérées comme étant des méthodes directes; les sept autres sont heuristiques. Par ailleurs trois des méthodes nécessitent l'utilisation de données à des pas de temps inférieurs à l'événement. Seules les méthodes et la synthèse des caractéristiques des groupes obtenus par chacune d'elles sont présentées ici. Les valeurs des paramètres utilisés dans les différentes méthodes de classification sont présentées pour les 3 années à l'annexe B1 (Tableaux 1,2,3,4,5 et 6).

4.2.1 Méthode CDA

La méthode de classification CDA (Clustering Direct Approach), est basée sur l'application de la technique de

"clustering" aux paramètres du variogramme: l'effet de pépite, le palier et la portée. Cette méthode a le désavantage de n'être applicable qu'après avoir modélisé les variogrammes de tous les événements pluvieux. Bien qu'elle soit associée à des paramètres caractérisant la structure spatiale d'un champ pluvieux, cette méthode ne produit pas une classification discriminante comme le montre les tableaux 7,8 et 9 de l'annexe B1. De plus, les événements pluvieux ne possédant pas de structure à palier, doivent être exclus de l'échantillon de base. L'analyse des variogrammes climatologiques des groupes correspondant n'a montré aucune différence notable entre les structures des groupes. Cette méthode n'a à notre avis aucun intérêt pratique.

4.2.2 Méthode CCVA

La méthode de classification CCVA (Clustering Coefficient of Variation Approach) est une méthode directe développée pour améliorer la robustesse du krigeage climatologique. Elle est basée sur l'application de la technique de "clustering" au coefficient de variation (CV) spatial des hauteurs de pluie cumulées sur l'événement. En effet, les champs à utiliser pour mettre en oeuvre le krigeage climatologique doivent être issus d'un même processus caractérisé par son coefficient de variation (Lebel

et Bastin, 1985). Cette méthode de classification est très discriminante. Elle donne pour chacune des années deux groupes significatifs chacun de taille sensiblement égale à la moitié de l'échantillon (Tableaux 7,8 et 9 de l'annexe B1). L'un des groupes est composé des événements pluvieux de faible CV, tandis que l'autre groupe est composé des événements possédant des fortes valeurs de CV. Il est démontré dans l'article figurant en annexe A1 que cette méthode est liée à la méthode CUPAA présentée plus loin.

4.2.3 Méthode CA

La méthode de classification CA (Crossing Approach) est la généralisation dans l'espace et le temps de la méthode de classification proposée par Kottégoda et Kassim (1991) pour la classification des hyétogrammes ponctuels de pluie. Elle est basée sur l'application de la technique de "crossing" au hyétogramme moyen de l'événement pluvieux. La technique de "crossing" consiste pour un histogramme donné à construire un hyétogramme cumulé standardisé sur les deux axes et à comparer sa position et sa forme par rapport à la bissectrice du système de coordonnées. Suivant la position et la forme du hyétogramme cumulé standardisé par rapport à la bissectrice, un ordre et une classe sont associés à l'événement pluvieux correspondant. La figure 5 de l'annexe

A1 fournit un exemple de calcul de l'ordre et de la classe. Cette méthode est de type systématique.

Pour classifier les événements pluvieux suivant cette approche, il suffit d'appliquer à tour de rôle la technique de "crossing" sur le hyétogramme moyen de chaque événement. La méthode bien que très séduisante n'a pas été très discriminante pour les données EPSAT-NIGER étudiées. En effet, pour chacune des trois années, plus de 70% des événements pluvieux appartiennent au même groupe 1b (Tableaux 7, 8 et 9, annexe B1). Ceci correspond au fait que les pluies sahéliennes débutent généralement par de fortes intensités. Enfin cette méthode n'est applicable qu'avec des données pluviographiques. Une autre de ses limitations est que la classification obtenue est fonction du pas de temps de base des données pluviographiques. En effet, il a été observé dans Amani et al. (1993a,b) que plus le pas temps de cumul est grand, plus grand est le nombre d'événements appartenant au même groupe 1b.

4.2.4 Méthode SCA

La méthode de classification SCA (Spatial Crossing Approach), est une extension de la méthode CA, cette fois-ci uniquement à partir des données pluviométriques. Nous avons

proposée cette méthode pour classifier, sur la base de la même idée utilisée au niveau de la méthode CA des événements pluvieux observés par un réseau de pluviomètres. Pour ce faire, une fonction de structure spatiale donnant le pourcentage de la surface pluvieuse en fonction du seuil a été définie. La technique de "crossing" est appliquée sur la fonction standardisée complémentaire. Les détails sur cette technique sont présentés dans l'article à l'annexe A1. Tout comme la méthode CA, la méthode SCA s'est avérée non discriminante, avec plus de 70% des événements pluvieux classifiés de la même façon (Tableaux 7, 8 et 9, annexe B1). Elle donne par contre une importante information sur le caractère intermittent des événements pluvieux observés par le réseau EPSAT-NIGER.

4.2.5 Méthode CUPAA

La méthode de classification CUPAA (Clustering Under Profile Area Profile) est basée sur l'application de la technique de "clustering" au nouveau paramètre UPA (Under Profile Area). Ce paramètre est déduit du profil spatial défini dans la méthode SCA. Il est donné par l'intégrale du profil spatial. Rappelons que ce profil spatial provient de la courbe donnant la surface pluvieuse en fonction du seuil. Pour plus de détails sur cette méthode on se reportera à

l'article de l'annexe A1. Tout comme la méthode CCVA, Cette méthode est très discriminante (Tableaux 7, 8 et 9, annexe B1).

Les groupes obtenus possèdent des propriétés très intéressantes et similaires à celles des groupes de la méthode CCVA. L'analyse des tableaux 7, 8 et 9 de l'annexe B1 montre que ces deux méthodes classifient sensiblement de la même manière les événements pluvieux. Une analyse de contingence montre que ces méthodes classifient de la même façon 70%, 65% et 80% des événements pluvieux respectivement en 1990, 1991 et 1992. Cette similitude s'explique assez bien à partir du lien analytique qui a été démontré entre les paramètres UPA et CV (voir annexe A1).

4.2.6 Méthode DUPA

Cette méthode est elle aussi basée sur le paramètre UPA. Elle diffère de la méthode précédente, par la manière dont sont formés les groupes. L'analyse de la fonction de densité (histogramme) des valeurs prises par le paramètre UPA sur les 133 événements pluvieux de l'étude (Figure 8 de l'annexe A1) montre qu'elle présente à toute fin pratique trois parties pour chacune desquelles, la fréquence de UPA est sensiblement constante. Les valeurs frontières de UPA

sont sensiblement égales à 0.66 et 0.81. On peut donc classer directement les événements pluvieux selon que la valeur de UPA appartienne à l'un ou l'autre de ces 3 groupes. Les événements appartenant au groupe 1 ($0.66 > \text{UPA}$) peuvent être associés aux lignes de grains. Ceux du groupe 3 ($\text{UPA} > 0.81$) sont composés principalement des événements très intermittents. On peut les associer aux événements provenant de systèmes convectifs isolés. Enfin les événements du deuxième groupe sont composés de champs d'intermittence intermédiaire et peuvent être associés aux systèmes convectifs de méso-échelle organisés mais n'ayant pas encore évolués en véritables lignes de grains.

4.2.7 Méthode CCVUPAA

Elle est la combinaison des méthodes CUPAA et CCVA, la technique de clustering étant appliquée à la fois sur le paramètre UPA et sur le coefficient de variation (CV). Elle produit une classification proches de ces deux méthodes, ce qui est logique, compte tenu du lien théorique qui existe entre le paramètre UPA et le CV (Tableaux 7, 8 et 9, annexe B1).

4.2.8 Méthode CMRCA

La méthode de classification CMRCA (Clustering Maximum Rainfall Cumul Approach) est basée sur l'application de la technique de "clustering" aux cumuls de pluie maximum enregistrés sur différents pas de temps. Les paramètres de cette méthode sont les hauteurs de pluie maximales observées pour un événement durant un pas de temps de cumul de 5, 10, 15, 30 et 60 minutes. Ces hauteurs de pluie caractérisent d'une certaine façon la structure spatiale et temporelle d'un événement pluvieux. Bien que cette classification conduise à beaucoup de groupes, il ressort pour chaque année, deux groupes significatifs de caractéristiques très différentes (Tableaux 7, 8 et 9 de l'annexe B1). L'un des groupes est composé des événements ayant une faible pluviosité, tandis que l'autre groupe est composé des événements de forte pluviosité.

4.2.9 Méthode CMRCA+DUPA

Cette dernière méthode est la combinaison des méthodes DUPA et CMRCA. Elle a été considérée afin d'étudier l'importance de la pluviosité (CMRCA) et de l'intermittence (UPA) au niveau de la classification des événements pluvieux.

4.2.10 Synthèse

Le tableau 4.1 présente la synthèse des méthodes tout en donnant la nature de la méthode ("clustering" ou systématique; direct ou heuristique), la nature des données pluviométriques nécessaires (au pas de temps de l'événement uniquement ou à des pas de temps inférieurs également). Le tableau 4.2 quant à lui présente la taille des groupes significatifs obtenus pour chacune des méthodes.

L'analyse de la classification obtenue d'une année à une autre montre que les méthodes de classification proposées même si elles ne sont pas toutes discriminantes, ont une certaine stabilité.

Une étude théorique du paramètre UPA nous a permis de montrer qu'il existe une relation analytique entre le paramètre UPA et le CV (voir article à l'annexe A1).

Tableau 4.1: Synthèse et caractéristiques des méthodes de classification étudiées.

Méthodes	Données	Approche	Clustering
CDA	spatiales	directe	oui
CCVA	spatiales	directe	oui
CA	espace temps	heuristique	non
SCA	spatiales	heuristique	non
CUPAA	spatiales	heuristique	oui
DUPA	spatiales	heuristique	non
CCVUPAA	spatiales	heuristique	oui
CMRCA	espace temps	heuristique	oui
CMRCA+DUPA	espace temps	heuristique	oui

Cette relation est donnée par

$$UPA = a_0 + \frac{a_1}{CV} + \frac{a_2}{CV^2} \quad (4.1)$$

où a_0 , a_1 et a_2 sont des coefficients de régression. Le tableau 1 de l'annexe A1, montre que les coefficients de détermination de la régression $UPA = f(1/CV, 1/CV_2)$ sont très significatifs et que les valeurs des constantes a_0 , a_1 et a_2 varient peu d'une année à l'autre. Ceci explique bien la similitude observée au niveau des deux méthodes de classification. Malgré ces similitudes, il faut insister ici sur la particularité de la méthode de classification basée sur UPA.

Tableau 4.2: Synthèse des groupes obtenus pour les différentes méthodes (taille des groupes)

Méthodes	Groupe	Taille		
		1990	1991	1992
CDA	1	18	14	33
	2	8	14	5
CCVA	1	19	20	16
	2	6	26	23
	3	9	-	-
CA	1b	25	24	38
	2a	7	14	7
	3b	-	5	-
SCA	1b	32	39	35
	2a	5	8	7
	2b	-	-	5
CUPAA	1	19	21	14
	2	12	25	21
	3	6	-	7
DUPA	1	6	10	5
	2	13	21	21
	3	18	16	23
CCVUPAA	1	13	21	37
	2	6	-	-
	3	15	25	10
CMRCA	1	15	15	24
	2	19	15	22
CMRCA+DUPA	1	27	23	29
	2	10	20	12

Le paramètre UPA est une mesure de la distribution spatiale, qui n'est ni la dispersion moyenne (cas du paramètre CV) ni l'intermittence au sens du nombre de valeurs

nulles. C'est une synthèse entre les deux, ce qui explique les liaisons trouvées (Tableaux 10, 11 et 12, annexe B1).

4.3 APPLICATION DE LA CLASSIFICATION AUX METHODES D'ESTIMATION

Le caractère climatologique des modèles d'estimation globaux fait que la classification des événements pluvieux peut a priori contribuer à améliorer leurs performances. Dans le cas des pluies sahéliennes caractérisées par trois types de systèmes précipitants ayant des structures spatiales très différentes, on a à gagner en considérant la stratification des événements pluvieux. La meilleure manière pour un modèle donné, d'améliorer l'estimation par la prise en compte de la classification, est que celle-ci soit basée sur des paramètres liés d'une certaine façon au modèle en question. Ainsi, le choix d'une méthode de classification se fera en fonction de la nature du modèle d'estimation qu'on se propose d'utiliser.

Pour cette application qui vise à évaluer l'apport de la classification pour l'estimation des pluies sahéliennes, deux modèles globaux d'estimation sont considérés. Il s'agit du krigeage climatologique pour l'estimation ponctuelle et de la méthode des seuils pour l'estimation de la pluie moyenne.

Le choix de ces méthodes est motivé par leur simplicité et leur intérêt climatologique. Sur la base de ces modèles, cinq méthodes de classification parmi les neuf analysées sont considérées dans cette application. Il s'agit des méthodes CMRCA, SCA, DUPA, CCVA et la combinaison de CMRCA et DUPA. Les méthodes CUPAA, CCVUPAA étant de la même famille (donnant des résultats semblables) que la méthode CCVA, n'ont pas été considérées. La méthode DA bien qu'elle soit basée sur les paramètres du variogramme, a été écartée parcequ'elle n'est pas très discriminante. Enfin, la méthode CA ayant un caractère temporel est jugée non appropriée pour un modèle spatial.

Pour l'évaluation de l'apport de la classification, la procédure classique de reconstitution d'une partie de l'échantillon est utilisée. L'erreur de reconstitution RMSE (Root Mean Square Error) est notre critère d'évaluation. Les valeurs du RMSE avant et après classification sont calculées et comparées.

4.3.1 Application au krigeage climatologique

Le krigeage climatologique requiert une homogénéité statistique des champs pluvieux pour le calcul du variogramme climatologique. Ce qui justifie en soi le principe de la

classification pour ce modèle. Le choix de la méthode de classification CUPAA est de ce fait raisonnable. Elle est la seule considérée ici. Cette méthode permet de distinguer deux échantillons de taille sensiblement égale en 1991 et en 1992, l'un regroupant des champs de faible valeur de UPA, l'autre ceux de forte valeur de UPA. Les variogrammes climatologiques de ces deux groupes, présentés à la figure 4.2, sont assez nettement différents, ce qui semblerait montrer la pertinence de la classification fondée sur UPA pour la mise en oeuvre du krigeage climatologique.

La différence essentielle entre les deux variogrammes réside dans le comportement aux grandes distances. Dans un cas c'est un comportement de champs plutôt continus (groupe de champs de faible valeur de UPA), dans l'autre de champs plutôt intermittents (groupe de champs de forte valeur de UPA). De ce fait, on ne peut pas appliquer aux champs intermittents (ici a priori les champs de forte valeur de UPA) qui violent l'hypothèse de base du krigeage à savoir la continuité spatiale. Le comportement non borné du variogramme climatologique de ces champs est expliqué de façon théorique à l'annexe B2.

A priori, pour le groupe caractérisé des champs de faible dispersion spatiale, on devrait s'attendre à des

meilleurs résultats du krigeage climatologique. Or les validations effectuées n'ont rien apporté de significatif, ni sur un groupe ni sur l'autre.

Il est important de souligner que le krigeage climatologique a été testé ici dans sa version ponctuelle et que son formalisme ne permet pas de prendre en compte de fortes discontinuités locales (sauf par effet de pépète). Or ce qui distingue les événements les uns des autres ce sont ces échelles de discontinuité. Si la classification peut apporter quelque chose pour des méthodes d'estimation ponctuelle c'est probablement en développant des catégories de modèle qui seraient différentes selon les groupes considérés.

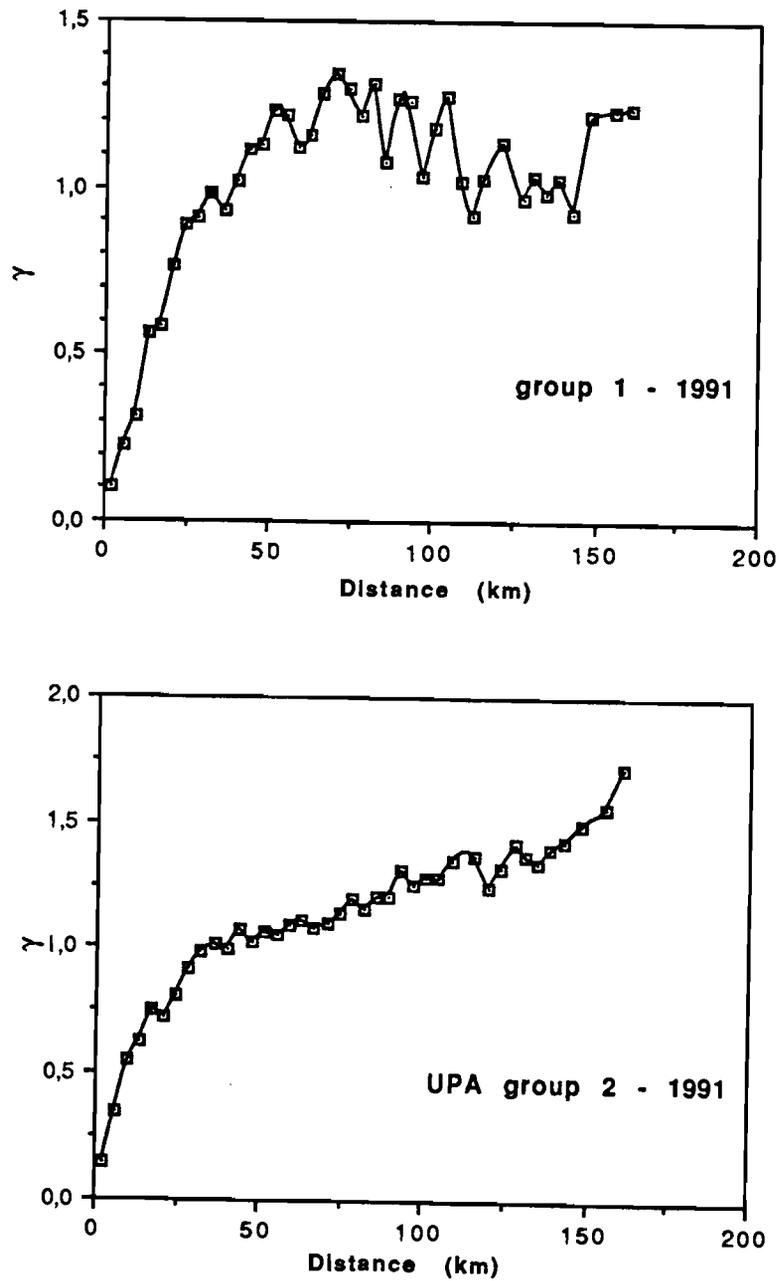


Figure 4.2: Variogrammes climatologiques des événements pluvieux appartenant aux groupes 1 et 2 de la méthode de classification CUPAA pour l'année 1991.

4.3.2 Application à la méthode des seuils

Elle provient de l'observation faite par Doneaud et al. (1981; 1984), selon laquelle pour une zone d'étude donnée de grande étendue (quelques milliers de Km²), la pluie moyenne d'un événement pluvieux est linéairement corrélée au pourcentage de la superficie couverte par une pluie d'intensité supérieure ou égale à un seuil donné. Il existe un seuil optimum pour lequel cette corrélation est maximale.

Cette méthode est considérée ici à cause de son intérêt pour l'estimation des pluies par les données de télédétection, plus particulièrement le satellite. Elle sera par contre appliquée ici comme l'ont fait Short et al. (1993) et Braud et al. (1993) à des données pluviographiques.

Ce choix avait tout d'abord pour raison de vérifier si la méthode des seuils peut être utilisée pour estimer la pluie moyenne et non l'intensité moyenne (il existe une différence entre les deux du moment où les événements pluvieux n'ont pas la même durée). Aussi, l'analyse des statistiques (moyennes conditionnelle et non conditionnelle) des groupes obtenus par les différentes méthodes de classification laissait entrevoir un intérêt a priori de la classification pour la méthode des seuils.

Au total, quatre des méthodes de classification proposées sont considérées pour évaluer l'apport de la classification pour l'estimation de la pluie moyenne événementielle sur le réseau EPSAT-NIGER par la méthode des seuils. Ces quatre méthodes sont: CMRCA, SCA, DUPA et la combinaison de CMRCA et DUPA. Les résultats détaillés de cette étude sont présentés dans l'article à l'annexe A1. On se contentera donc de faire ressortir ici les points importants.

1) Nonobstant la classification, la méthode de seuil s'applique très bien à l'estimation des moyennes des cumuls. Conformément à la théorie, le coefficient de détermination du modèle des seuils croît bien avec la valeur du seuil jusqu'à atteindre un maximum et puis décroît. Les pentes des droites de régression sont du même ordre de grandeurs que les moyennes conditionnelles.

2) le point le plus important de la classification est le fait que pour les quatre méthodes considérées et plus particulièrement pour la méthode basée sur le paramètre UPA, le seuil optimum est différent d'un groupe à un autre. Ceci est la conséquence du fait que les groupes ont des caractéristiques statistiques différentes. Il serait donc inconsistant, pour ce qui est des données étudiées ici,

d'appliquer sans classification préalable la méthode des seuils, puisque l'hypothèse qui est faite au niveau de cette méthode est que l'échantillon est caractérisé par une seule loi statistique. Ce résultat se doit d'être considéré pour toute application de la méthode de seuil, dans la mesure où les régimes pluviométriques pour lesquels elle est appliquée sont composés d'événements caractérisés par des distributions spatiales non homogènes.

3) L'apport de la classification sur l'amélioration de l'estimation par la méthode de seuil est sans équivoque que ça soit en phase de calibration ou de validation. Cette évaluation s'est faite aux seuils optimums. Pour chaque groupe, le modèle correspondant au seuil optimum est uniquement considéré. Les résultats de nos deux scénarios de validation montrent bien l'intérêt de la classification pour cette méthode. Une méthode de classification s'est particulièrement illustrée. Il s'agit de la méthode basée sur le paramètre UPA. En effet pour les deux scénarios de validation les pourcentages de réduction de l'erreur de validation avant et après classification sont de 33% et 25%. L'apport de la classification est le plus important pour les champs de faible UPA.

4.4 Autre intérêt de la méthode proposée: application à la climatologie.

L'analyse des statistiques présentées au tableau 3 de l'annexe A1 pour le cas de la méthode de classification DUPA montre un intérêt certain de la classification pour l'étude de la climatologie au Sahel. En effet, on observe que d'une année sur l'autre, en dehors de l'année 1992, les moyennes et les coefficients de variation sont assez stables pour les trois groupes obtenus par cette méthode. Aussi, on observe que le rapport entre la moyenne non conditionnelle et la moyenne conditionnelle est stable d'une année sur l'autre pour chacun des groupes. Ce rapport noté F_0 caractérise d'une certaine façon le degré d'intermittence des événements pluvieux. Ainsi sur la base de ce rapport, on peut a priori essayer d'identifier ces trois groupes avec les trois types de systèmes météorologiques observés au Sahel. De toute évidence des données météorologiques sont nécessaires pour confirmer ou non le bien fondé de ce rapprochement.

Par contre, il existe des statistiques qui sont quant à elles variables d'une année à l'autre pour un groupe donné. Il s'agit de la proportion des événements du groupe, ainsi que de leur contribution en pourcentage au cumul annuel. En outre, on observe que le rapport entre ces deux paramètres

est sensiblement stable d'une année à l'autre. Ce résultat est très intéressant pour ce qui est de la climatologie du Sahel. Ainsi, le nombre total d'événements pluvieux est le principal paramètre explicatif du cumul annuel pluviométrique au Sahel.

Bien que la proportion des événements des différents groupes ne peut pas expliquer le déficit ou l'excédent saisonnier, elle a par contre à l'évidence une portée climatique et des conséquences hydrologiques certaines au Sahel (Leroux, 1988). Une saison de faible proportion en événements de groupe 1 et/ou 2 peut être climatiquement et hydrologiquement anormale même si elle est une saison normale ou excédentaire en terme de cumul annuel. Cette situation s'explique bien par la nature différente de la réponse hydrologique à la pluviométrie.

4.5 CONCLUSION

Bien que presque exclusivement de nature convective, les systèmes précipitants de la région sahélienne présentent des signatures diverses, que ce soit au niveau de l'imagerie infra-rouge satellitaire ou au niveau des champs de précipitations observés au sol. Il est donc naturel d'en établir une classification permettant à la fois de mieux les

décrire et d'améliorer l'efficacité des algorithmes d'estimation qui présupposent généralement une certaine homogénéité des échantillons auxquels ils sont appliqués. Plusieurs méthodes de classification ont donc été proposées dans ce travail avec ce double objectif. Elles se basent uniquement sur des données sol: ce sont en effet les plus facilement disponibles en Afrique de l'Ouest et celles qui fournissent la donnée la plus couramment utilisée par les algorithmes d'estimation. Les résultats que nous avons obtenus concernent les méthodes de classification proprement dites, d'une part, et leur application à certains algorithmes d'interpolation et à la climatologie, d'autre part.

Pour ce qui est des méthodes de classification, il est apparu que la stratification la plus nette était obtenue en utilisant des critères liées à la structure spatiale des événements pluvieux, plutôt qu'à leur structure temporelle ou aux intensités observées. Une méthode basée sur un paramètre statistique unique, dénommé UPA (Under Profile Area), lié à la fois à la variabilité moyenne d'ensemble (CV spatial) et à la probabilité d'apparition des valeurs les plus fortes, est apparue comme particulièrement attractive.

Dans le domaine de l'application de la classification pour la mise en oeuvre d'algorithmes d'estimation, nous avons

choisi de travailler sur deux algorithmes relativement robustes et simples à mettre en oeuvre: le krigeage climatologique pour reconstituer des valeurs ponctuelles, et la méthode des seuils utilisée notamment pour estimer des valeurs moyennes sur des surfaces assez étendues (plusieurs centaines ou milliers de km² à partir des données de télédétection).

En ce qui concerne le krigeage climatologique, aucun résultat réellement significatif n'a été obtenu. Il semble bien que le problème posé soit plus ici celui de la validité de la méthode d'interpolation que celui de la classification, mais des investigations plus approfondies seraient nécessaires pour en juger valablement. En ce qui concerne la méthode des seuils, une amélioration spectaculaire de ses résultats a été enregistrée lorsqu'on l'applique après classification, c'est-à-dire en calant pour chaque groupe identifié une relation différente entre la pluie moyenne et la surface touchée au-dessus du seuil optimal. La valeur même de ce seuil optimal change également d'un groupe à l'autre. La classification basée sur le critère UPA est celle qui fournit la plus forte amélioration.

Enfin, dans le domaine de la climatologie, la classification UPA s'est aussi avérée prometteuse. En effet

les groupes obtenus présentent des caractéristiques statistiques (moyennes, CV, F0) stables d'une année sur l'autre. Par contre, la proportion d'événements d'un groupe donné, varie d'une année sur l'autre, ce qui pourrait expliquer une grande partie de la variabilité de la réponse hydrologique au forçage pluviométrique et servir de paramètre diagnostic dans les études des fluctuations des régimes pluviométriques.

Une étude approfondie des caractéristiques des trois groupes obtenus par la méthode de classification DUPA est une suite logique et intéressante surtout pour ce qui est de la climatologie sahélienne. Un travail de thèse est en cours dans ce sens dont un des volets est l'étude des distributions statistiques des cumuls pour les trois groupes d'événements obtenus par DUPA (D'Amato). Aussi, la méthode de classification DUPA peut très bien être appliquée pour stratifier des données de télédétection.

CHAPITRE V: ESTIMATION SPATIO-TEMPORELLE DES PLUIES AUX
FAIBLES PAS DE TEMPS: KRIGEAGE DYNAMIQUE

5.1 INTRODUCTION

Les cumuls successifs de pluie en cinq (5) minutes montrées à la figure 3.3 pour la ligne de grains du 20 août 1991 font apparaître un déplacement régulier de ce système précipitant. Une étude systématique menée par Nathalie D'Amato (communication personnelle reprise dans Lebel et al., 1995b), confirme que les systèmes convectifs de méso-échelle sahéliens se déplacent de façon généralement régulière. Zawadzki (1973b), Johnson et Bras (1979) et Niemczynowicz et al. (1984) ont remarqué un tel comportement pour d'autres régions climatiques et Zawadzki (1973b) a suggéré que pour de tels systèmes l'hypothèse de Taylor était vérifiée pour des pas de temps inférieurs à 35 min: les corrélations spatiale et temporelle sont alors liées par le biais du vecteur vitesse. Les algorithmes d'interpolation utilisant cette propriété ne sont toutefois pas toujours aisés à mettre en oeuvre, car il faut identifier à la fois les fonctions de corrélation et le vecteur vitesse. Or les images de la figure 3.3 soulignent combien les champs pluvieux en 5 minutes sont à la fois intermittents et anisotropes. C'est pourquoi nous

explorerons ici deux approches permettant de prendre en compte la dynamique du SCM dans l'interpolation des champs précipitants aux pas de temps inférieurs à 15 minutes.

La première approche est eulerienne, elle repose sur un repère fixe lié au réseau de mesure (annexe 5.1). La deuxième est lagrangienne et s'appuie sur un repère mobile lié au système précipitant. Les deux repères sont portés en figure 5.1. A titre d'illustration préliminaire, on montre en figure 5.2 comment deux champs eulériens de la ligne de grains du 20/08/91 se transforment dans le repère lagrangien. Le calcul de ces champs lagrangiens sera expliqué dans la section 5.2.2 ci-dessous.

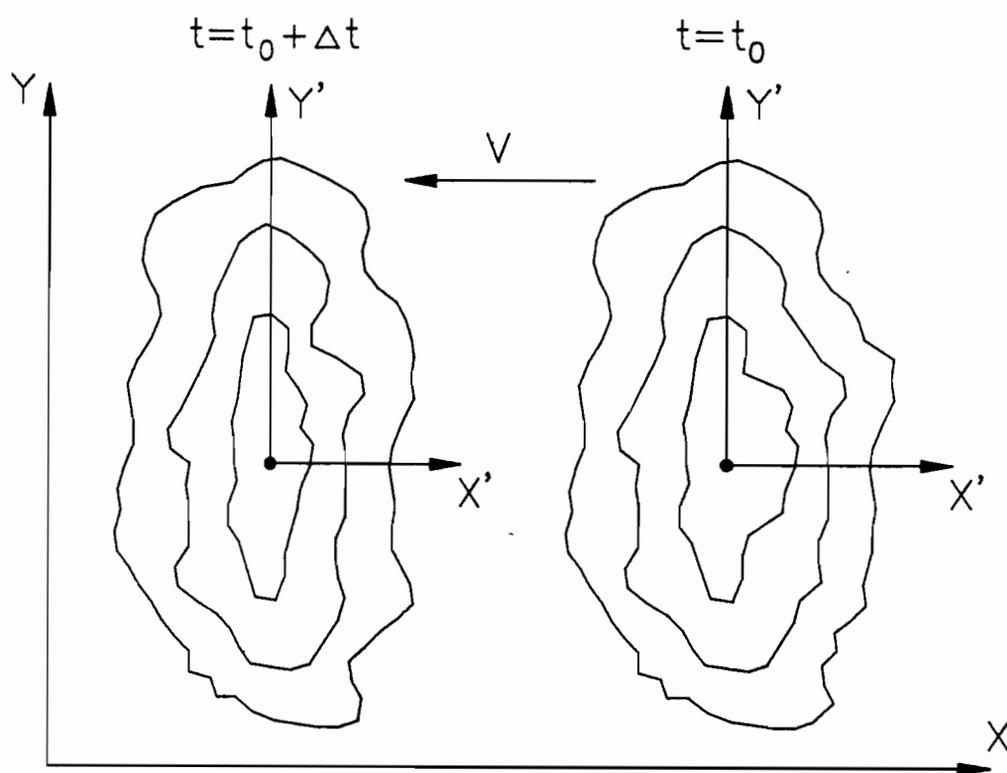


Figure 5.1: Illustration des repères fixe et mobile et de la dynamique de champs pluvieux

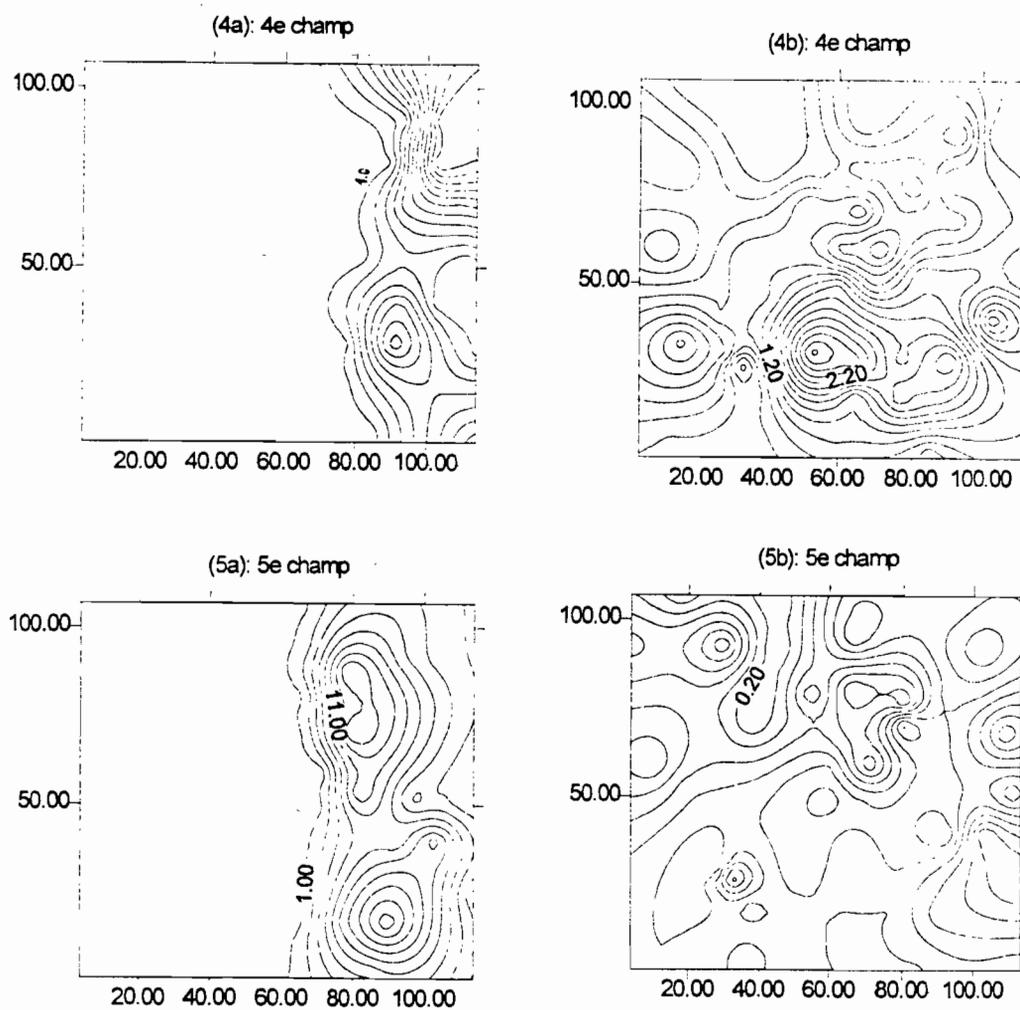


Figure 5.2: Répartition spatiale d'un champ pluvieux, événement du 20/08/91; a) repère fixe, b) repère mobile

5.2 KRIGEAGE SPATIO-TEMPOREL

Notre première tentative de prise en compte de la dépendance temporelle qui existe entre les champs pour l'estimation aux faibles pas de temps à consister à étendre le krigeage 2D dans l'espace et le temps. Ce qui nous donne la technique du krigeage spatio-temporel mise en application dans l'article Amani et al .(1995b) présenté à l'annexe A2. Afin de s'affranchir du problème relatif au calcul de la distance euclidienne (mélange du métrique et du temporel), deux méthodologies ont été considérées. La première consistait à standardiser les axes, tandis que la deuxième a consisté à convertir le temps en distance par le biais de la vitesse de déplacement du système pluvieux, cette conversion étant possible grâce à l'hypothèse de Taylor (Zawadzki, 1973b).

L'analyse des résultats de cette application présentés au niveau de l'article Amani et al. (1995b) à l'annexe A2, montre que la simple généralisation du krigeage dans l'espace temps, même avec l'utilisation de la vitesse permettant de convertir le temps en distance, ne contribue pas à améliorer de façon significative l'estimation 2D. Cette situation est probablement due à une mauvaise prise en compte de la dynamique des champs pluvieux.

5.3 KRIGEAGE DYNAMIQUE

5.3.1 Prise en compte des corrélations temporelles?

Les résultats de la généralisation brutale du krigeage 2D dans 3D n'ayant pas été à la hauteur de nos espérances, nous avons cherché d'autres façons de prendre en ligne de compte les corrélations temporelles entre les champs. L'une d'elle serait de considérer la fonction de covariance comme étant fonction d'une distance temporelle et d'une distance métrique. Mais ceci soulève le problème du choix de la fonction de covariance admissible. L'autre façon qui est considérée à consister à exploiter la dynamique des champs pluvieux. Elle permet une prise en compte plus efficace de l'intermittence, de la dépendance temporelle des champs et de leur anisotropie.

Soit $P(x_i, y_i, k)$ avec $(i=1, \dots, N)$ et $(k=1, \dots, K)$ la série de données spatio-temporelles observées dans le repère fixe lors du passage d'un événement pluvieux sur le réseau d'observation composé de N pluviographes pour un pas de temps de cumul Δt donné. K représente ici le nombre total de champs. L'analyse de la fonction de covariance croisée $C_{ij}(\tau)$ pour un décalage temporel τ entre les séries temporelles observées aux stations i et j montre que la valeur maximale

de $C_{ij}(\tau)$ est observée pour un décalage temporel différent de zéro et est fonction de la distance séparant les deux stations. Réciproquement, pour un décalage temporel τ donné, il existe un incrément spatial $(\Delta x_{\max}(\tau), \Delta y_{\max}(\tau))$ pour lequel $C_{ij}(\tau)$ est maximal. Ces incréments sont liés à la vitesse de déplacement du système pluvieux (Marshall, 1980) par

$$\Delta x_{\max}(\tau) = U_x \tau \text{ et } \Delta y_{\max}(\tau) = U_y \tau \quad (5.1)$$

où U_x et U_y sont les composantes sur les axes x et y de la vitesse du système pluvieux. Pour deux stations i et j , on a les relations suivantes:

$$\Delta x_{ij} = U_x \tau_{\max} \text{ et } \Delta y_{ij} = U_y \tau_{\max} \quad (5.2)$$

avec $\Delta x_{ij} = (x_i - x_j)$, $\Delta y_{ij} = (y_i - y_j)$ et τ_{\max} le décalage temporel pour lequel la corrélation $C_{ij}(\tau)$ est maximale. Ainsi à partir de l'étude de la corrélation spatio-temporelle, il est en théorie possible de déterminer la vitesse de déplacement du système pluvieux. La figure 5.3 illustre le décalage temporel entre deux séries temporelles données.

L'idée du krigeage dynamique (KD) repose sur ce lien qui existe entre la vitesse et les positions relatives des stations de mesure. Si $C_{ij}(\tau)$ est maximale pour le décalage temporel τ_{\max} , alors les séries temporelles $P(x_i, y_i, k)$ avec $k=1, \dots, K-\tau_{\max}$ et $P(x_j, y_j, k)$ avec $k=\tau_{\max}, \dots, K$ auront une covariance croisée maximale, pour un décalage temporel nul.

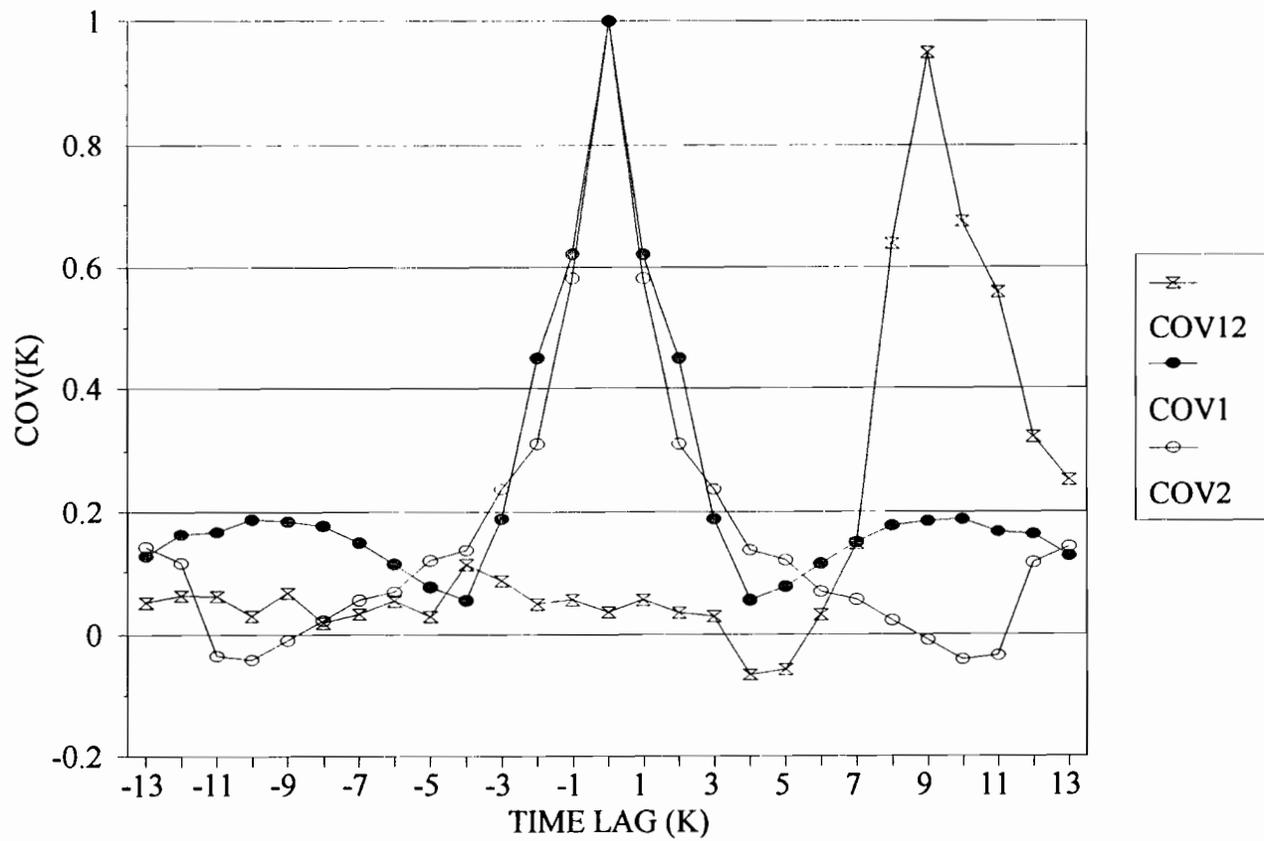


Figure 5.3: Illustration du décalage temporel sur la base de la fonction d'autocorrélation.

En d'autres termes, la hauteur de pluie $P(x_i, y_i, k)$ pour un k donné est plus dépendante des hauteurs de pluie $P(x_j, y_j, k + \tau_{\max}(ij))$ ($j=1, \dots, N-1$), $\tau_{\max}(ij)$ étant le décalage temporel pour lequel $C_{ij}(\tau)$ est maximale, que des hauteurs de pluie $P(x_j, y_j, k)$ ($j=1, \dots, N-1$) observées pour le même temps k dans le repère fixe. Ainsi, au lieu d'estimer la pluie en une station de coordonnée (x_o, y_o) pour le temps k_o comme étant

$$\hat{p}(x_o, y_o, k_o) = \sum_{i=1}^N \lambda_i P(x_i, y_i, k_o) \quad (5.3)$$

dans le cas d'une estimation 2D, ou par

$$\hat{p}(x_o, y_o, k_o) = \sum_{i=1}^N \sum_{k=1}^K \lambda_{ik} P(x_i, y_i, k) \quad (5.4)$$

dans le cas d'une estimation spatio-temporelle classique (Annexe A2), dans l'approche du krigeage dynamique cette estimation sera donnée pour le cas 2D par

$$\hat{p}(x_o, y_o, k_o) = \sum_{i=1}^N \lambda_i P(x_i, y_i, k_o + \tau_{oi}) \quad (5.5)$$

où τ_{oi} est le décalage temporel maximisant la corrélation entre les séries P_o et P_i . La valeur de τ_{oi} peut être estimée connaissant la vitesse U du système pluvieux par $\tau_{oi} = d_{oi}/U$ où d_{oi} est la distance entre les deux points i et o . La figure 5.4 présente dans le cas unidimensionnel une illustration du

krigeage dynamique. Au lieu d'estimer le champs k par les observations appartenant à la même horizontale, ce qui se fait dans l'approche traditionnelle, il sera estimé par les observations appartenant à la ligne oblique ab dont la pente est fonction de la vitesse U du système pluvieux. A cause de la nature du réseau et la vitesse U , d_{oi}/U peut ne pas être un multiple du pas de temps Δt des données; dans ce cas on considère le temps multiple le plus proche. Les difficultés relatives à une modélisation correcte des covariances spatio-temporelles ne facilitent en aucun cas la mise en oeuvre des techniques du krigeage spatio-temporel telles qu'elles ont été présentées à l'annexe A2. Egalement la structure particulière des SCM (convectif+stratiforme) pose un problème pour ces covariances lorsqu'on est à cheval sur ces deux parties. Pour toutes ces raisons, cette nouvelle technique d'interprétation des corrélations temporelles ne peut qu'être séduisante.

5.3.2 Du krigeage dynamique au krigeage dynamique lagrangien

La technique du krigeage dynamique telle qu'elle est formulée dans l'équation 5.5 présente une difficulté liée à l'estimation de τ_{oi} , donc du vecteur vitesse U mais surtout du fait que le champ des données à kriger varie selon le point d'interpolation considéré.

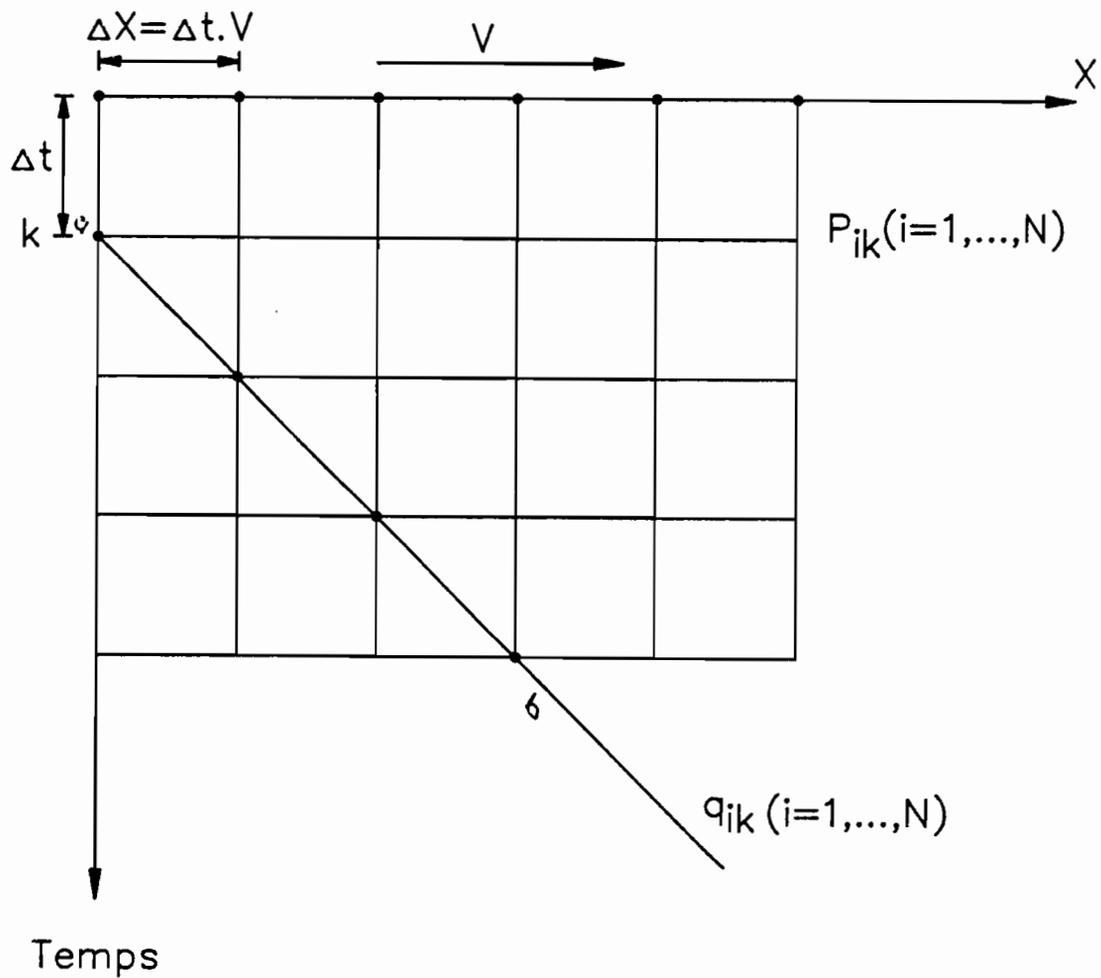


Figure 5.4: Illustration de l'estimation dynamique dans un cas unidimensionnel. V vitesse du système, P_{ik} et q_{ik} pluies dans les repères fixe et mobile.

Ainsi pour un même temps donné k_o , il faudra interpoler le champ de valeurs $P(x_j, y_j, k_o + \tau_{oj})$ pour reconstituer $P(x_o, y_o)$, alors que pour un point (x_a, y_a) , éloigné il faudra interpoler le champ des valeurs $P(x_j, y_j, k_o + \tau_{aj})$, τ_{aj} étant différent de τ_{oj} . C'est pourquoi il apparait préférable d'aller jusqu'au bout de la démarche qui sous-tend le krigeage dynamique en transformant les champs eulériens en champs lagrangiens et en appliquant une interpolation par krigeage 2D classique à ces champs lagrangiens (voir section suivante).

Rapportons nous maintenant au repère mobile (l'observateur se déplace avec l'orage). Dans ce repère d'observation, l'origine temporelle est propre à chaque station et celle-ci correspond au début de pluie au niveau de la station. Ainsi si $q(x_j, y_j, h)$ ($h=1, \dots, H$) est la série spatio-temporelle telle que observée dans le repère mobile, celle-ci est reliée à la série provenant du repère fixe par

$$q(x_j, y_j, h) = p(x_j, y_j, h + t_j / \Delta t) \quad (5.6)$$

avec t_j étant le temps mis par l'orage pour arriver à la station j (l'origine du temps étant le début de pluie sur le réseau). Connaissant la vitesse moyenne U de l'orage supposée constante, le temps t_j est calculé par $t_j = d_j / U$ avec d_j la distance séparant le point j au point de début de pluie sur le réseau. On verra par la suite qu'avec cette astuce, le krigeage dynamique revient à kriger les champs lagrangiens

tels qu'observés par un observateur se déplaçant avec l'orage. C'est ce que nous appellerons krigeage dynamique lagrangien.

Les champs de pluie dans le repère mobile sont obtenus à partir des champs dans le repère fixe en tenant compte de la dynamique du système pluvieux. Deux procédures sont possibles. La première qui est dite indirecte est basée sur la vitesse de déplacement du système pluvieux. La vitesse permet de passer du temps à l'espace et vice-versa. Par exemple si U est la vitesse moyenne de l'orage, la pluie $P(x_i, y_i, t)$ observée dans le repère fixe, la même hauteur de pluie sera observée dans le repère mobile au temps $t \pm d_i/U$ avec d_i étant la distance parcourue par l'orage du début de la pluie sur le réseau pour arriver au point i . Le signe $+$ ou $-$ sera fonction de la direction de déplacement de l'orage.

Cette procédure nécessite l'estimation de la vitesse moyenne de l'orage. L'utilisation d'une vitesse moyenne introduit des erreurs à cause du fait que l'orage se déplace réellement avec une vitesse variable d'un point à un autre. En annexe B4, une nouvelle méthodologie pour l'estimation de la vitesse d'approche variable de l'orage est proposée (vitesse variable d'un point à un autre). Celle-ci est basée sur la technique du krigeage dual appliquée sur les temps

d'arrivée de l'orage (début, centre ou fin de l'orage). Ainsi la pluie $P(x_i, y_i, t)$ dans le repère fixe sera observée dans le repère mobile au temps $t -/+d_i/v_i$ avec cette fois-ci, v_i étant la vitesse d'approche de l'orage au point i . Cette procédure de détermination des champs dans le repère mobile à partir des données observées par rapport au repère fixe, à cause de l'étape intermédiaire d'estimation de vitesse de l'orage, peut introduire des erreurs d'ordre numérique.

La deuxième procédure qui fait l'objet d'application ici et qui n'introduit apparemment pas de problème numérique, consiste à considérer pour chaque série temporelle sa propre origine. C'est-à-dire que pour passer des champs observés par rapport au repère fixe à ceux observables dans le repère mobile, il suffit de considérer que chaque station a sa propre origine de temps correspondant au début de la pluie pour la station en question. Ainsi avec cette approche, la mise en oeuvre du krigeage dynamique lagrangien ne requière que la connaissance des temps de début de pluie aux différents points du réseau. La figure 5.5 illustre à titre d'exemple, les isochrones de début de pluie pour la ligne de grain du 20/08/91.

5.3.3 Interpolation par krigeage dynamique lagrangien

. Estimation ponctuelle

L'interpolation de $P(x_o, y_o)$ au temps $h_o + d_o / (U \Delta t)$ par le krigeage dynamique lagrangien est donné, en remplaçant k_o par $h_o + d_o / (U \Delta t)$ dans l'équation (5.5), par

$$\hat{p}(x_o, y_o, h_o + d_o / U) = \sum_{i=1}^N \lambda_i p(x_i, y_i, h_o + d_o / U + \tau_{oi}) \quad (5.7)$$

or $d_o / U + \tau_{oi} = d_i / U$ ce qui fait que

$$p(x_i, y_i, h_o + d_o / U + \tau_{oi}) = p(x_i, y_i, h_o + d_i / U) = q(x_i, y_i, h_o) \quad (5.8)$$

Ainsi l'équation (5.7) devient

$$\hat{p}(x_o, y_o, h_o + d_o / U) = \sum_{i=1}^N \lambda_i q(x_i, y_i, h_o) \quad (5.9)$$

D'où le krigeage dans le repère mobile du champ $q(x_i, y_i, h_o)$ ($i=1, \dots, N$) au point (x_o, y_o) n'est rien d'autre que le krigeage dynamique au même point, (x_o, y_o) , mais pour un temps égal à $h_o + d_o / U$ dans le repère fixe (d_o / U est le temps que met l'orage pour arriver au point o). Ainsi le champ h à kriger dans le repère mobile pour estimer la pluie au point (x_o, y_o) au temps k_o dans le repère fixe est donné par l'entier le plus proche de $k_o - d_o / (U \Delta t)$.

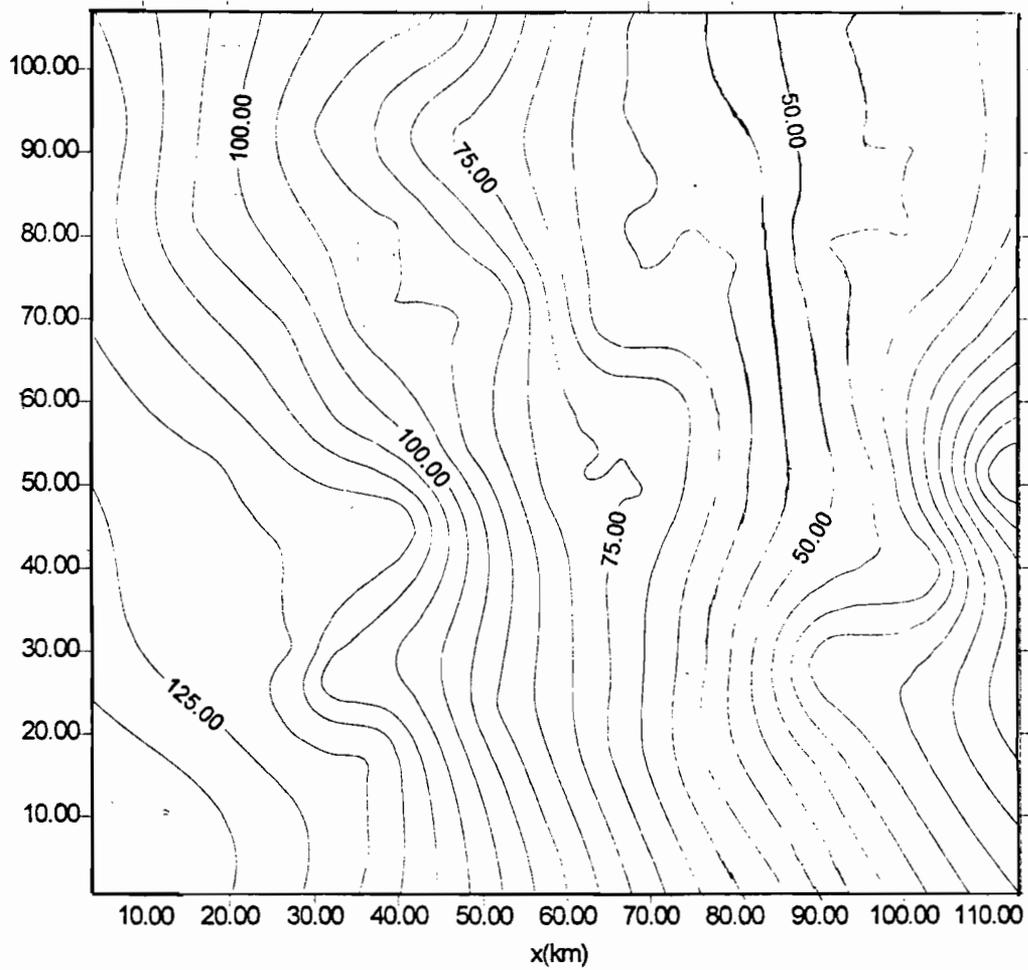


Figure 5.5: Isochrones de début de pluie pour la ligne de grains du 20/08/91.

Comme on peut le vérifier avec des données réelles, le processus spatial $q(x,y,h)$ possède des propriétés qui concordent mieux aux techniques du krigeage que le processus $p(x,y,k)$. En effet le processus $q(x,y,h)$ est continu et isotrope et peut être considéré stationnaire. Le processus $q(x,y,h)$, à cause de la faible variation de la structure spatiale d'un champ à un autre, se prête mieux que les champs $p(x,y,k)$ aux techniques de krigeage climatologique et vectoriel (voir partie sur l'analyse structurale), conduisant ainsi à la modélisation d'un seul variogramme (variogrammes climatologique et/ou multivariable). Ce qui réduit énormément l'aspect numérique du krigeage dynamique lagrangien.

A cause de la nature du réseau et de la nature discrète des données sur l'axe du temps (champs k , $k=1,..K$), le krigeage dynamique lagrangien à champs discrets introduit des erreurs systématiques dues aux arrondis provenant du calcul de $h + d/(U*\Delta t)$ et $h + \tau/\Delta t$ (entier le plus proche). Il est possible de les réduire, en procédant aux krigeages dynamiques spatio-temporels (3D) dans le repère mobile. Le processus $q(x,y,h)$ est considéré comme un processus 3D continu sur lequel on peut appliquer les techniques du krigeage.

L'estimation en (x_o, y_o) au temps t_o dans le repère mobile sera donnée par

$$\hat{q}(x_o, y_o, t_o) = \sum_{i=1}^M \lambda_i q(x_i, y_i, t_i) \quad (5.10)$$

avec $t_i = (h-1) \cdot \Delta t$ ($h=1, \dots, H$), H le nombre de champs considéré. Cette estimation correspond dans le repère fixe au temps $t_o + d_o/U$. Pour estimer $p(x_i, y_i, t_o)$, il suffit de kriger en 3D dans le système mobile les données spatio-temporelles au point de coordonnées $(x_o, y_o, t_o - d_o/U)$. Le krigeage sous sa forme duale est plus approprié à cette approche de krigeage dynamique du moment où dans la forme duale, le système de krigeage est résolu une seule fois et permet d'exprimer la fonction krigée directement en fonction des coordonnées du point (x, y, t) . En supposant que le processus 3D $q(x, y, t)$ est stationnaire, alors la fonction krigée est donnée par

$$\hat{q}(x_o, y_o, t_o - d_o/U) = \hat{p}(x_o, y_o, t_o) = a + \sum_{i=1}^M b_i K_{oi} \quad (5.11)$$

où a et b_i ($i=1, \dots, M$) s'obtiennent par la résolution du système d'équations suivant

$$\sum_{i=1}^M b_i K_{ij} + a = q_j \quad j=1, \dots, M$$

(5.12)

$$\sum_{i=1}^M b_i = 0$$

avec K_{ij} la valeur de la fonction de covariance entre les points d'observations (x_i, y_i, t_i) et (x_j, y_j, t_j) dans le repère mobile et K_{oi} est par contre la valeur de la fonction de covariance entre les points (x_i, y_i, t_i) et $(x_o, y_o, t_o - d_o/U)$ toujours au niveau du repère mobile. On suppose que cette fonction de covariance n'est fonction que de la position relative des observations.

. Etimation de la pluie moyenne

Soit A l'aire sur laquelle on aimerait estimer la pluie moyenne au champ k_o dans le repère fixe. La pluie moyenne $P_A(k_o)$ donnée par

$$P_A(k_o) = \frac{1}{A} \int_A P(x, y, k_o) dA \quad (5.13)$$

est d'habitude estimée dans le cas 2D par

$$\hat{P}_A(k_o) = \sum_{i=1}^n \lambda_i P(x_i, y_i, k_o) \quad (5.14)$$

où n est le nombre de stations appartenant à l'aire A .
Avec le krigeage dynamique, il est possible d'utiliser une plus grande quantité de données spatio-temporelles pour estimer $P_A(k_o)$. En raisonnant comme au niveau de l'estimation ponctuelle, à chaque valeur de $p(x_i, y_i, k_o)$ il est possible, à cause de la dynamique du système pluvieux, de lui associer un champ $q(x_j, y_j, h_o)$ avec $j=1, \dots, N$ et $h_o=k_o-d_j/U$ dans le repère mobile. Les valeurs caractérisant ce champ sont fortement liées à $p(x_i, y_i, k_o)$ et donc peuvent être prises en ligne de compte dans toute estimation faisant intervenir $p(x_i, y_i, k_o)$. Comme l'aire A contient plusieurs observations (repère fixe), alors les données spatiales de plusieurs champs (repère mobile) doivent être considérées pour l'estimation de $P_A(k_o)$.

Les champs à considérer et leur nombre sont fonction de la position de l'aire A par rapport au réseau de mesure et de son étendue. Si T est le temps que met l'orage pour traverser entièrement l'aire A , le nombre Z de champs (repère mobile) à considérer est égal à $T/\Delta t$ (entier le plus proche). Maintenant, si T' est le temps que met l'orage pour arriver au niveau de l'aire A (début de pluie sur A), alors les champs (repère mobile) à considérer seront k_o-Z_1 à k_o-Z_1+Z avec Z_1 étant l'entier le plus proche de $T'/\Delta t$.

Ainsi l'estimation de $P_A(k_o)$ par le krigeage dynamique sera donnée par

$$\hat{P}_A(k_o) = \sum_{h=k_o-Z_1}^{k_o-Z_1+Z} \sum_{i=1}^N \lambda_{ih} q(x_i, y_i, h) \quad (5.15)$$

où λ_{ih} est le poids associé à l'observation $q(x_i, y_i, h)$ dans le repère mobile. L'application des conditions de non biais et de variances minimales conduit, dans le cas où le processus est considéré stationnaire au système d'équations suivant

$$\sum_{h=k_o-Z_1}^{k_o-Z_1+Z} \sum_{i=1}^N \lambda_{ih} G(i, h, j, h') + \mu = \frac{1}{A} \int_A G(x, k-d/(U*\Delta t), j, h') dA \quad j=1, \dots \quad (5.16)$$

$$\sum_{h=k_o-Z_1}^{k_o-Z_1+Z} \sum_{i=1}^N \lambda_{ih} = 1$$

avec $G(i, h, j, h')$ étant la fonction de covariance entre les observations (i, h) et (j, h') et $G(x, k_o-d/(\Delta t*U), j, h')$ est la covariance entre un point quelconque (x, y) de l'aire A et les différentes observations (j, h') . On voit que l'estimation de la pluie moyenne par l'approche du krigeage dynamique nécessite la modélisation d'une covariance spatio-temporelle comme dans le cas de l'estimation ponctuelle avec krigeage dynamique 3D.

5.3.4 Analyse structurale pour le cas sahélien

Tout comme les autres techniques du krigeage, la mise en oeuvre du krigeage dynamique lagrangien nécessite l'estimation et la modélisation de la fonction de structure. En dehors du fait que cette étape détermine pour une bonne partie la qualité de l'estimation du krigeage, une attention particulière doit être portée quant à la nature des structures spatiales caractérisant les champs à kriger. Dans cette application, on se limite au KDL 2D. Les champs de pluie sont considérés dans le repère mobile. L'analyse structurale des données spatio-temporelles des lignes de grains du 27/07/90, 20/08/91 et 30/08/92, par rapport aux repères fixe et mobile conduit aux différentes observations mentionnées ci-après.

Pour chacun des événements pluvieux, on distingue deux structures spatiales, une caractérisant les champs intenses (partie convective de l'orage) et l'autre caractérisant la queue de l'orage (partie stratiforme). La structure des champs appartenant à la partie stratiforme est caractérisée par un variogramme à effet de pépite pur (Figures 5.6a et ainsi que les figures 1a et 2a de l'annexe B3).

La comparaison des champs pluvieux de l'événement pluvieux du 20/08/91 dans les deux repères (fixe et mobile), montrent que les champs dans le repère mobile n'ont aucune direction privilégiée et peuvent être considérés comme étant isotropes. Par contre les champs dans le repère fixe présentent manifestement une certaine anisotropie (Figures 5.2a et 5.2b).

Pour les trois lignes, les variogrammes climatologiques possèdent un palier voisin de l'unité, ce qui est en conformité avec la théorie du krigeage climatologique. Par contre dans le cas du repère fixe, les variogrammes correspondants sont quasiment linéaires (Figures 5.6b ainsi que les figures 1b et 2b de l'annexe B3). La séparation de structures entre les parties convective et stratiforme est d'une grande importance pour l'estimation aux faibles pas de temps. Ceci confère au krigeage dynamique lagrangien un avantage certain par rapport à l'approche classique dans le repère fixe où cette séparation est plus délicate et laborieuse à effectuer.

Avec le krigeage dynamique lagrangien 2D, l'estimation d'une partie convective s'effectuera avec les données de la partie convective; de même, l'estimation d'une partie stratiforme se fera avec les données provenant de la partie

stratiforme. Le krigeage dynamique lagrangien en dehors de ses avantages par rapport aux approches traditionnelles de krigeage se rapproche donc de la physique du phénomène étudié (dynamique, séparation entre les parties convective et stratiforme).

5.4 MISE EN OEUVRE DU KRIGEAGE DYNAMIQUE LAGRANGIEN 2D SUR LES LIGNES DE GRAINS SAHELIENNES

5.4.1 Procédure pour l'interpolation ponctuelle

Soit un point de coordonnées (x_0, y_0) , on désire estimer la pluie en ce point au temps t (l'origine du temps est ici par rapport au repère fixe). La procédure suivante est utilisée:

1- à partir de l'interpolation des temps d'arrivée de la pluie, estimer le temps d'arrivée de la pluie, t_0 au point o (exemple de la figure 5.5).

2- si $t - t_0$ est inférieur à zéro, alors le système pluvieux n'est pas encore arrivé en o , d'où $P(x_0, y_0, t) = 0$;

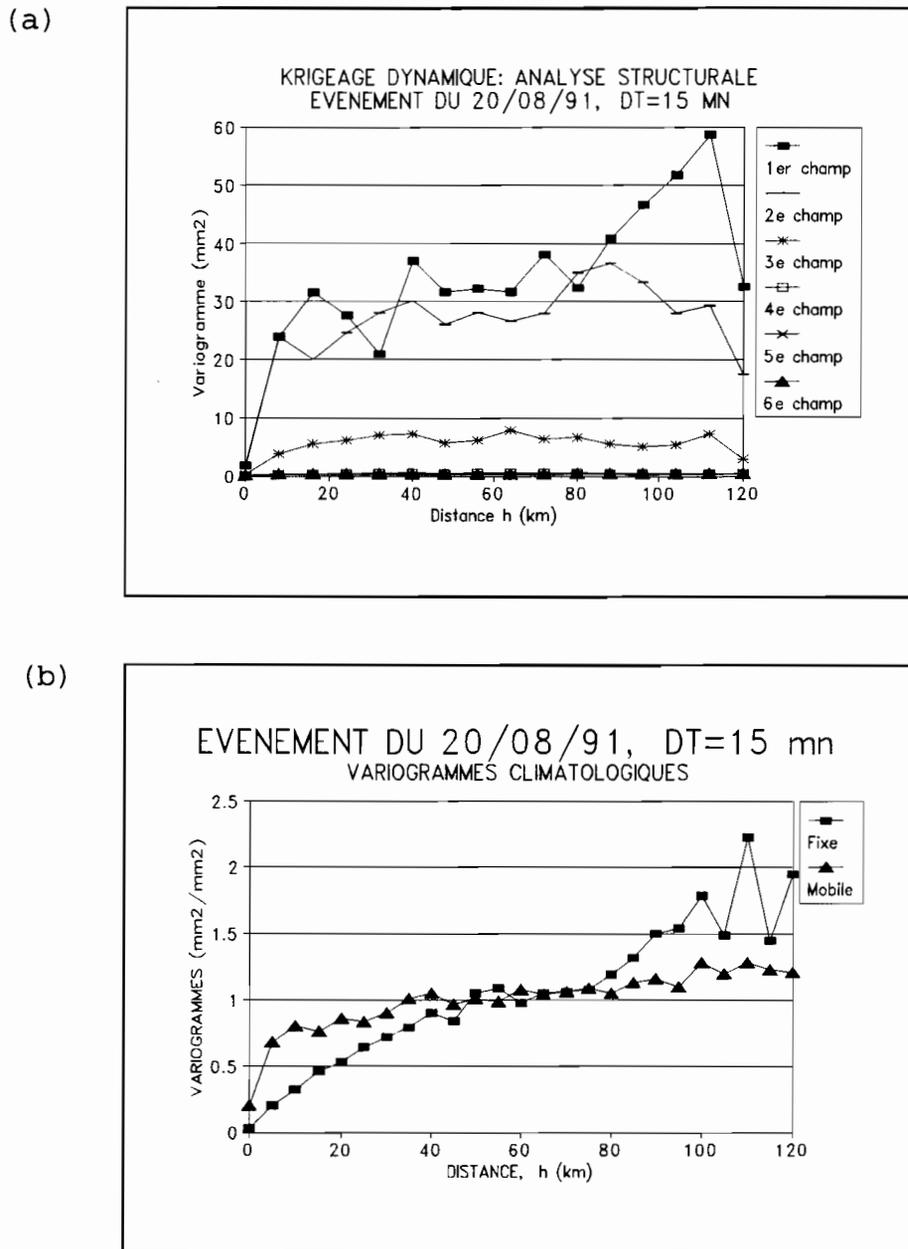


Figure 5.6: Variogrammes des champs pour l'événement du 20/08/91, a) repère mobile; b) variogrammes climatologiques

3- si $t-t_0 > 0$, le champ de pluie à considérer dans le repère mobile est celui du rang $(t-t_0)/\Delta t$ dans le cas où cette expression donne un entier (Δt est le pas de temps de cumul). Dans le cas où cette expression n'est pas entière, il doit exister un entier k tel que $(k+1) \Delta t > t-t_0 > k \Delta t$. Deux approches sont dans ce cas possibles:

a) considérer le champ de rang k tel que l'entier k soit le plus proche du réel $t-t_0$. Ainsi le champ à kriger en (x_0, y_0) dans le repère mobile est celui de rang k .

b) kriger individuellement les deux champs k et $k+1$ au point (x_0, y_0) donnant respectivement les estimations $P_0(k)$ et $P_0(k+1)$. On obtient l'estimation en o au temps t par une simple interpolation linéaire:

$$P_0(t) = P_0(k) + (P_0(k+1) - P_0(k)) ((t-t_0)/\Delta t - k) \quad (5.17)$$

Cette approche a deux désavantages: celui de lisser trop les données et celui de prendre numériquement deux fois plus de temps que la première. Elle a été rejetée au cours du traitement en faveur de la première approche.

4- Pour chacun des champs k , l'estimation au point de coordonnées (x_0, y_0) sera donnée par une combinaison linéaire des observations du champ en question, c'est-à-dire que

$$\hat{P}_k(x_0, y_0) = \sum_{i=1}^N \lambda_i P_k(x_i, y_i) \quad (5.18)$$

avec λ_i la pondération associée à la mesure $P_k(x_i, y_i)$ et N , le nombre total des stations de mesure. La détermination de ces pondérations diffère en fonction de la l'appartenance du champ k à l'une des deux partie de l'orage (parties convective et stratiforme). Dans le cas où le champ est dans la partie convective, les valeurs des λ_i ($i=1, \dots, N$) sont données par le système classique de krigeage climatologique. Par contre dans le cas de la partie stratiforme, l'estimation est donnée par une simple moyenne arithmétique en voisinage glissant (structure pépitique):

$$\hat{P}_k(x_o, y_o) = \frac{1}{M} \sum_{i=1}^M P_k(x_i, y_i) \quad (5.19)$$

avec M étant l'ensemble de stations de mesure contenues dans un cercle centré en o et de rayon R ($N > M$).

5.4.2 Estimation de la pluie moyenne (méthode simplifiée)

L'estimation de la pluie moyenne sur une surface A donnée appartenant au réseau de base par l'intermédiaire du krigeage dynamique pour un temps donné (champ donné dans le repère fixe) se fait en deux étapes. La première consiste à déterminer en fonction de la position de cette surface par rapport au réseau de base, les champs du repère mobile et leur nombre à considérer dans l'estimation. Si T' est le temps que met l'orage pour intéresser la zone A et si T est

le temps que met l'orage pour traverser l'aire A alors le nombre de champs à considérer est donné par $T/\Delta t$ et ils ont pour rang $(t-T')/\Delta t$ à $(t-T'+T)/\Delta t$.

Telle que présentée au paragraphe 5.2.3, la mise en oeuvre du krigeage dynamique pour l'estimation de la pluie moyenne nécessite à première vue la modélisation des structures spatiales croisées des champs à utiliser pour cette estimation. La modélisation de ces structures croisées est souvent délicate et n'est pas sans créer des problèmes numériques au niveau de la résolution du système de cokrigeage. Le nombre de champs qu'il est possible d'inclure est de ce fait limité, et l'ensemble représente de toutes façons un calcul assez lourd.

Pour toutes ces raisons, une approche simplifiée est considérée ici. Elle consiste à diviser la surface initiale A en surfaces élémentaires par l'intermédiaire d'un maillage. La pluie au centre de gravité de chacune des surfaces élémentaires est estimée en utilisant la procédure présentée au précédent paragraphe pour l'estimation ponctuelle. La pluie moyenne de la surface A est donnée par la moyenne arithmétique des pluies estimées aux différents centre de gravité des surfaces élémentaires. Cette estimation sera d'autant plus précise que le maillage est plus fin.

5.4.3 Procédure de validation

Afin de comparer l'efficacité du KDL à celle du KF (krigeage fixe), quatre interpolateurs sont considérés: le krigeage fixe (KF), la moyenne glissante par rapport au repère fixe (MF), le krigeage dynamique (KDL ou KD simplement) et la moyenne glissante par rapport au repère dynamique (MD). La procédure classique suivante de validation est considérée:

1- pour un pas de temps donné, les champs de pluie sont extraits à partir de deux réseaux sensiblement uniformes issus de la séparation du réseau initial en deux. L'un des deux est considéré pour la modélisation et l'autre pour la validation. Leur rôle est ensuite inversé. On suppose que K est le nombre total de champs extraits et N_k le nombre de stations ayant fonctionné pour le champ k ($k=1, \dots, K$).

2- pour un champ k donné (repère fixe), les pluies $P_{i,k}$ observées par les N_k stations sont reconstituées par chacun des quatre interpolateurs en fonction des données de l'autre réseau. On note par $P_{i,k,e}$ ($i=1, \dots, N_k$) les estimations obtenues pour un interpolateur e donné ($e=KF, MF, KD$ et MD). Les erreurs d'estimations sont définies par trois statistiques à savoir le RMSE (Root Mean Square Error, Erreur quadratique

moyenne), le EM (Erreur Moyenne) et Emax (l'erreur maximale). Elles sont données par

$$RMSE(k) = \sqrt{\frac{\sum_{i=1}^{N_k} (P_{i,k,e} - P_{i,k})^2}{N_k}} \quad (5.20)$$

$$EM(k) = \frac{\sum_{i=1}^{N_k} (P_{i,k,e} - P_{i,k})}{N_k} \quad (5.21)$$

$$Emax(k) = \max_{i=1}^{N_k} |P_{i,k,e} - P_{i,k}| \quad (5.22)$$

Dans le cas où la comparaison des erreurs entre les champs d'un même événement pluvieux ou entre les champs des événements pluvieux différents est nécessaire, il est préférable d'utiliser les statistiques standardisées. L'erreur quadratique standardisée est obtenue en divisant le RMSE par l'écart type S de l'échantillon ($PRMSE = RMSE/S$); cette erreur en pourcentage est notée PRMSE. L'erreur moyenne est quant à elle standardisée par la moyenne, m de l'échantillon ($PME = EM/m$). Elle est notée par PME en pourcentage aussi. Plus l'estimation est précise, plus ces erreurs sont faibles; à la limite, elles sont égales à zéro, ce qui équivaut à une estimation parfaite. Les mêmes calculs

sont effectués en inversant le rôle des deux réseaux et une moyenne des erreurs sur les deux réseaux est considérée.

3- L'étape 2) est appliquée pour toutes les valeurs de k ($k=1, \dots, K$).

La même procédure est utilisée dans le cas de la validation pour l'estimation de la pluie moyenne à la différence qu'un maillage doit être considéré à la place des stations de mesure. On divise la surface en surfaces élémentaires. La pluie au centre de gravité de chaque surface élémentaire est supposée caractériser la pluie moyenne de la surface élémentaire. Ici la pluie de référence à utiliser au niveau des équations (17,18 et 19) est estimée pour chaque maille à partir du réseau initial global. Au niveau de ces équations N_k sera remplacée par le nombre total de mailles.

5.4.4 Création d'un étalon à l'aide d'un hyétogramme théorique de ligne de grains

Un modèle théorique de ligne de grains est construit afin de valider la procédure d'estimation par KDL. Deux modèles de ligne de grains sont construits. Le premier est une ligne de grains dite parfaite (reproduction d'un hyétogramme de pluie sans déformation d'un point à un autre)

et le second est une ligne de grains bruitée (le hyétogramme subit une déformation aléatoire d'un point à un autre).

La simulation de la ligne de grains est faite sur la base du modèle d'hyétogramme donné à la figure 5.7 dont les caractéristiques sont les suivantes: durées des parties convective et stratiforme 30min et 3h, l'intensité maximale de 100mm/h et celle de la partie stratiforme supposée constante est supposée égale à 3mm/h. Le choix de cet hyétogramme est bien en accord avec les caractéristiques temporelles de lignes de grains sahéliennes. La vitesse de déplacement de la ligne est considérée égale à 50 km/h et une direction est-ouest est supposée. Ces valeurs sont bien en accord avec les caractéristiques dynamiques (vitesse moyenne et direction) des lignes de grain sahéliennes, comme le montre la figure 5.9 donnant les histogrammes de vitesses moyennes et direction pour les événements pluvieux de 1990, 1991 et 1992 observés sur le réseau EPSAT-NIGER, valeurs calculées à partir de la méthode de trois points (Niemczynowicz et Dahlblom, 1984; Johnson et Bras, 1979). Pour la simulation d'une ligne de grains bruitée, le bruit est supposé être proportionnel à l'intensité (ou à la hauteur pour un pas de temps de cumul donné) et est considéré comme gaussien avec un coefficient de variation constant. Deux niveaux de bruit sont considérés, $CV = 20\%$ et $CV = 50\%$.

La simulation d'une L.G est faite sur un réseau de 100km par 100km dont la maille élémentaire est de 12.5 km par 12.5 km maille qui, est identique à celle du réseau de base d'EPSAT-Niger. Le choix de ce maillage permet de comparer les ordres de grandeur des erreurs d'estimation sur les L.G simulées à celles obtenues sur les lignes de grain réelles observées sur le réseau EPSAT-NIGER.

Soit un point de grille (i,j) de coordonnées (x_i, y_j) et $P_{ij}(k)$ la hauteur de pluie simulée au k^{eme} champ et soit (x_o, y_o) les coordonnées d'un point repère du réseau (point de début de pluie) et V la vitesse de déplacement de la L.G dans la direction des X (Figure 5.8). On note par d_{ij} la distance entre le point de grille (i,j) et le point de référence o.

On note également par $q(k*\Delta t)$ ($k=1, \dots, K$) la série temporelle de cumul de pluie du hyétogramme modèle. L'algorithme de simulation de la L.G.P (Ligne de Grain Parfaite) est le suivant:

1) pour un point de grille (i,j) et un k donné, calculer les distances:

- $d_{ij} = |x_i - x_o|$, distance entre le point (i,j) et le point repère;

- $d_k = U*k*\Delta t$, distance parcourue par la L.G pour le champ k ;

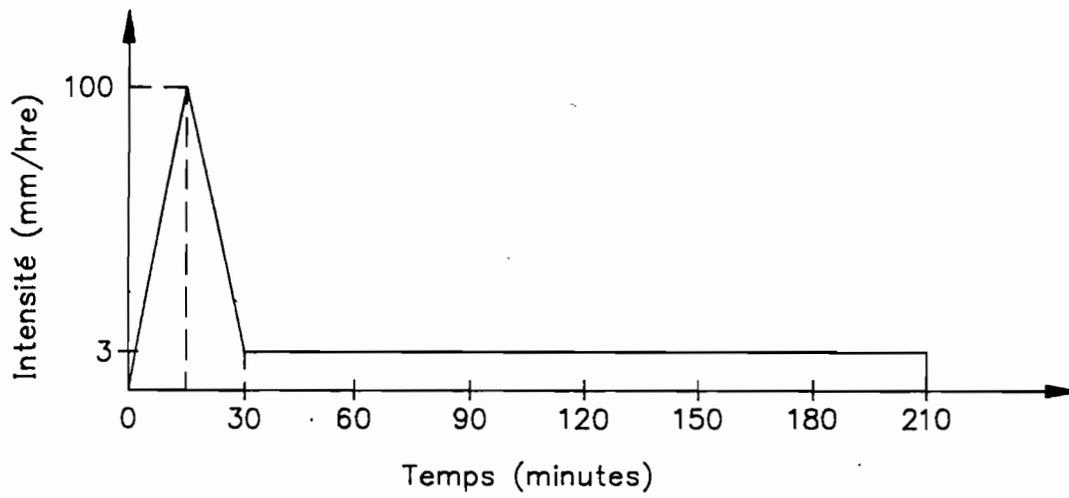


Figure 5.7: Modèle théorique d'hyéotgramme pour la simulation de la ligne de grains

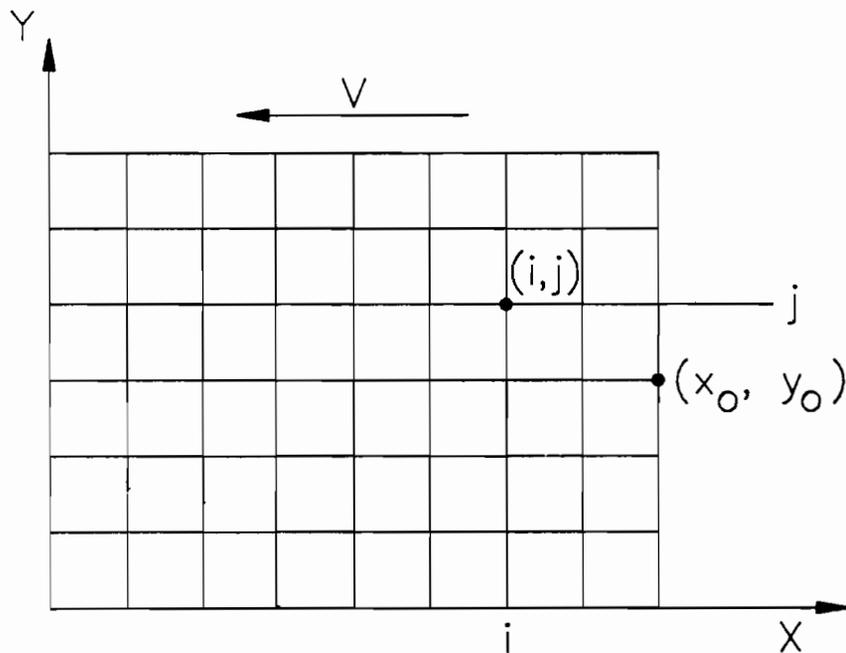
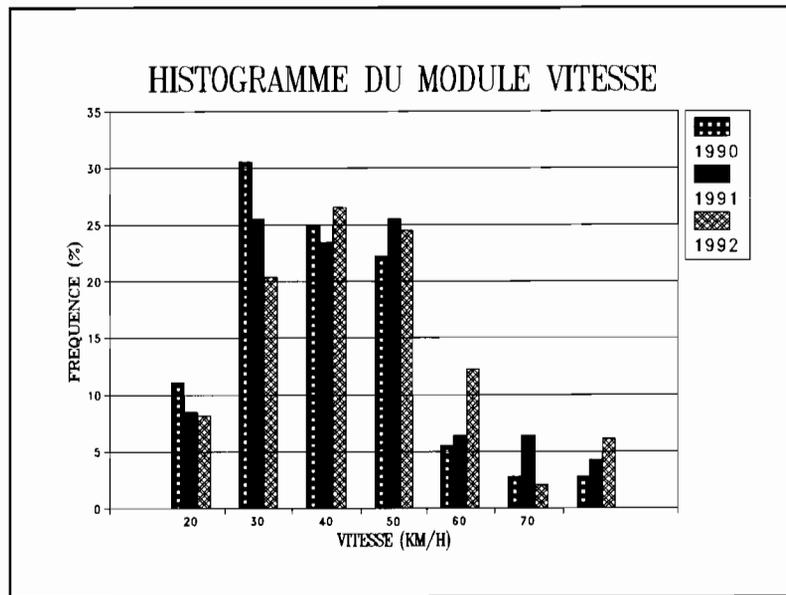


Figure 5.8: Maillage pour la simulation de la ligne de grains

(a)



(b)

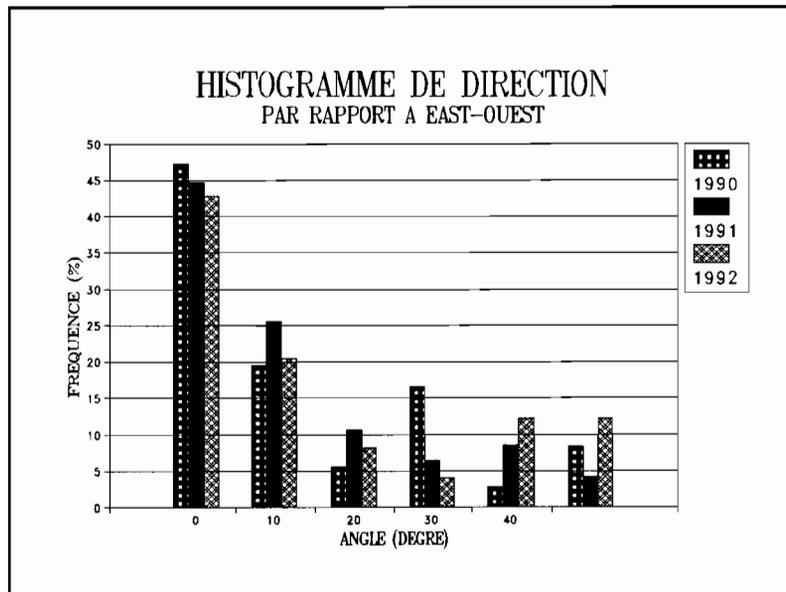


Figure 5.9: Histogrammes du vecteur vitesse des orages au Sahel, a) module vitesse b) direction

2) si $d_{ij}-d_k \geq 0$, alors la pluie simulée est $P_{ij}(k)=0$ du moment où la ligne n'a pas encore atteint le point de grille (i,j) .

3) si par contre $d_{ij}-d_k < 0$, calculer

$$- \tau = |d_{ij}-d_k| / (U \cdot \Delta t)$$

$$- H = \text{Ex}(\tau) + 1 \text{ si } \tau - \text{EX}(\tau) \geq 0.5$$

$$= \text{Ex}(\tau) \text{ sinon}$$

avec $\text{Ex}()$ étant la fonction partie entière. Ainsi la pluie simulée au point (i,j) est donnée par $P_{ij}(k) = q(H \cdot \Delta t)$. Pour une ligne de grains parfaite, les hauteurs de pluie simulées dans la direction perpendiculaire à la ligne de grains sont les mêmes d'un point de grille à un autre (i constant).

4) Répéter les opérations 1), 2) et 3) pour les différentes valeurs de k , ce qui donne un jeu de données spatio-temporelles.

Pour ce qui est de la simulation d'une ligne de grains bruitée), l'algorithme est le même que celui d'une ligne de grain parfaite à la seule différence qu'une étape intermédiaire de simulation de nombre aléatoire est nécessaire en début d'algorithme. Soit CV le niveau de bruit supposé constant ($CV = S/m$, S est l'écart type et m étant la moyenne). A toute valeur de pluie $P_{ij}(k)$ on associe un bruit aléatoire e gaussien de telle sorte que la valeur bruitée est

donnée par

$$P_{b_{ij}}(k) = P_{ij}(k) + l \cdot S = P_{ij}(k) + l \cdot CV \cdot m \quad (5.23)$$

En considérant $m = P_{ij}(k)$, la valeur simulée sera donnée par

$$P_{b_{ij}}(k) = P_{ij}(k) \cdot (1 + l \cdot CV) \quad (5.24)$$

avec $P_{ij}(k)$ valeur simulée sans bruit au niveau de la ligne de grains parfaite. Le paramètre l est le facteur de fréquence de la loi normale, facteur de fréquence à tirer aléatoirement.

La simulation de l se fait en deux étapes. La première étape consiste à simuler un réel compris entre 0 et 1 (niveau de probabilité de dépassement) auquel on associe dans la deuxième étape un facteur de fréquence a par l'intermédiaire de la fonction de distribution inverse de probabilité de la loi normale réduite. Pour chaque point de grille et chaque k , une valeur de l est simulée, ce qui fait que l'hystogramme initial du modèle théorique de L.G.P (ligne de grains parfaite) subit une déformation aléatoire gaussienne. La simulation est effectuée pour $CV=20\%$ et 50% .

5.5 RÉSULTATS ET DISCUSSIONS

L'étude comparative des quatre interpolateurs (KF, MF, KD et MD) sera faite d'une part sur les données spatio-temporelles simulées par le modèle théorique de ligne de grains et d'autre part sur les données spatio-temporelles de

lignes de grain réelles observées sur le réseau EPSAT-NIGER. L'application sur les données simulées nous permettra de vérifier la pertinence de la nouvelle approche proposée.

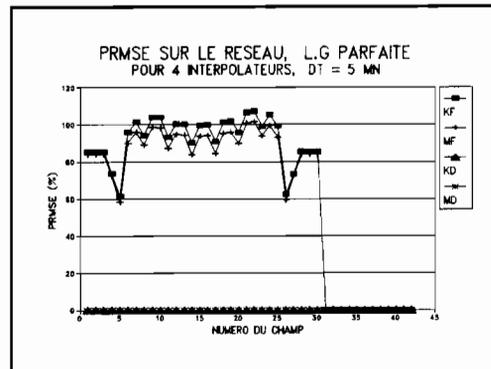
5.5.1 Données simulées

La procédure de validation présentée en section 5.3.3 a été appliquée aux trois jeux de données spatio-temporelles (LGP; LGB20; LGB50) du modèle théorique de ligne de grains. Les résultats sur la L.G.P sont en parfait accord avec la théorie de l'estimation dynamique. En effet les erreurs de reconstitution (PRMSE, PME et EMAX) des deux interpolateurs KDL et MD issus de l'approche dynamique sont nulles pour quasiment tous les champs, tandis que, celles des interpolateurs KF et MF sont de loin différentes de zéro (figure 5.10). Le fait que les derniers champs ne possèdent pas des erreurs de reconstitution identiquement égales à zéro s'explique par le fait que l'algorithme mis en place considère des hauteurs de pluie nulles au delà de la durée de la pluie sur le réseau de mesure. Cette situation explique aussi comme on le verra plus loin le fait que des erreurs très grandes pour les interpolateurs KD et MD sont associées aux derniers champs.

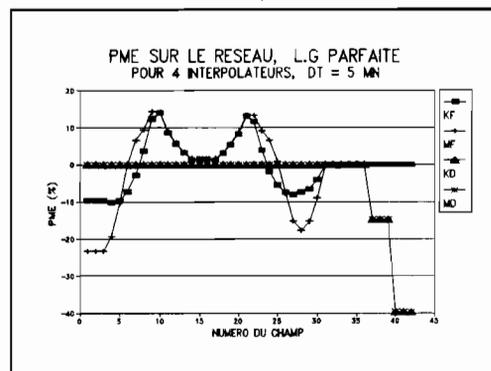
Les résultats concernant les lignes de grains "bruitées" (LGB20 et LGB50) sont représentés en figures 5.11 et 5.12. On observe que les erreurs associées aux interpolateurs KDL et MDL sont sans équivoque inférieures à celles associées aux deux autres méthodes. Cependant, la supériorité des modèles issus de l'approche dynamique par rapport aux modèles fixes est fonction du niveau de bruit. Plus la ligne est bruitée, plus l'apport de la dynamique pour l'estimation spatiale diminue. L'amélioration obtenue en effectuant l'estimation par un modèle dynamique par rapport à un modèle fixe est d'environ 75% et 50% en ce qui concerne le PRMSE pour les valeurs de CV de 20% et 50%. Pour un CV de zéro (ligne de grains parfaite) l'amélioration est de 100%.

L'analyse des erreurs des modèles KDL et MDL montre qu'il n'y a aucune différence significative entre les deux interpolateurs, ce qui signifie que l'advection est le facteur prédominant de structuration des champs pluvieux dans un tel cas. Dans la mesure où les paramètres choisis pour définir la ligne de grains théorique correspondent aux moyennes établies durant l'expérience EPSAT-NIGER, on peut conclure que les lignes de grains réelles se propageant avec peu de déformations, le KDL donnera des résultats meilleurs que le KF.

(a)



(b)



(c)

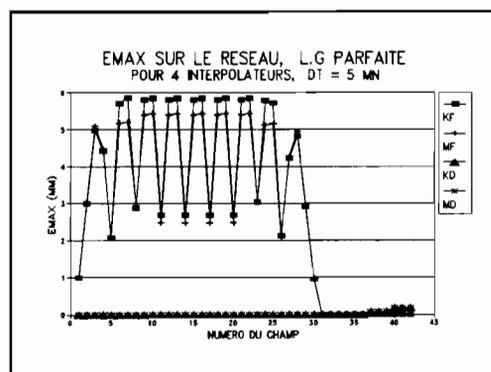
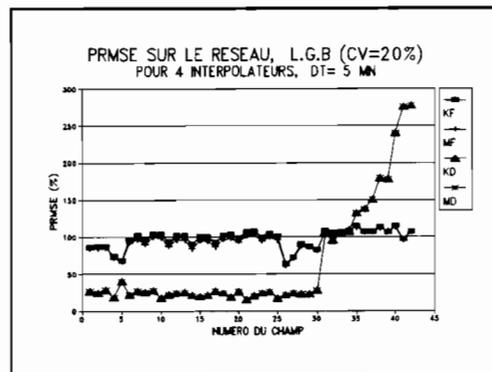


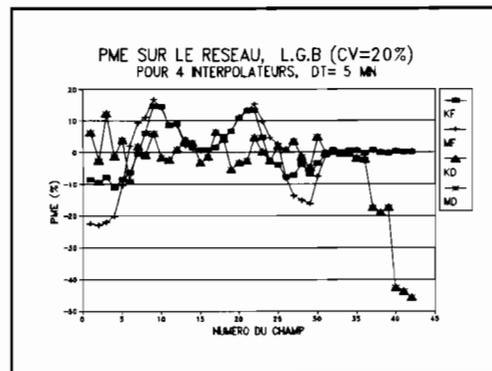
Figure 5.10: Erreurs de validation ponctuelle pour la LGP

a) PRMSE, b) PME et c) Emax

(a)



(b)



(c)

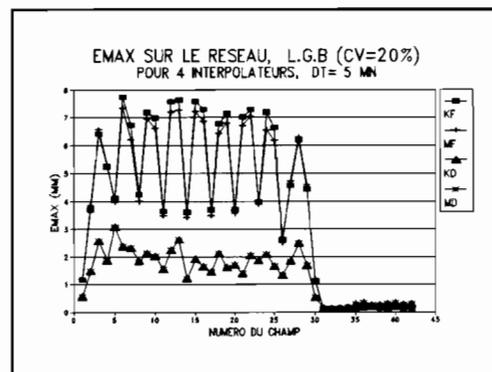
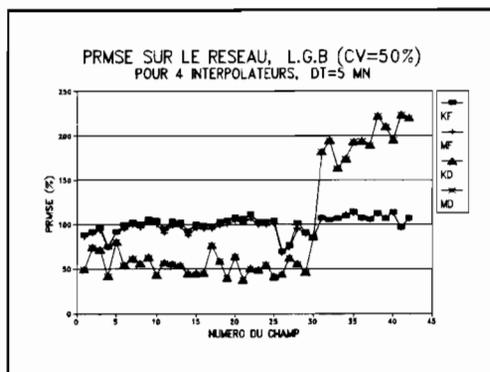
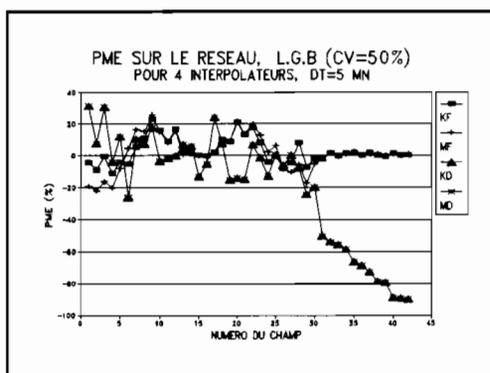


Figure 5.11: Erreurs de validation ponctuelle pour la LGB (CV=20%), a) PRMSE, b) PME et c) Emax

(a)



(b)



(c)

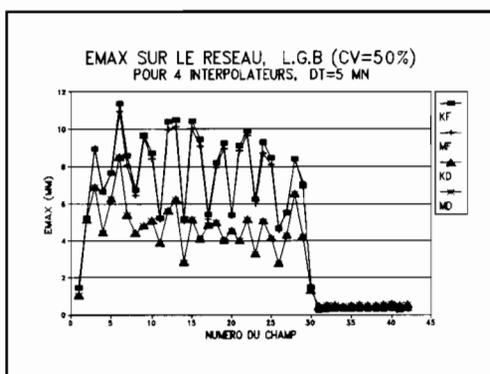


Figure 5.12: Erreurs de validation ponctuelle pour la LGB (CV=50%), a) PRMSE, b) PME et C) Emax

5.4.2 Données réelles (estimation ponctuelle)

On utilise ici les données spatio-temporelles de lignes de grains du 27/07/90, 20/08/91 et 30/08/92 observées par le réseau EPSAT-NIGER. L'application est faite non seulement sur le réseau de base, mais aussi sur la cible, ceci afin d'évaluer la sensibilité du KDL aux échelles spatiales considérées (distances entre stations de mesure et surfaces d'interpolation). Différents pas de temps de cumul ont sélectionnés: 10 et 15 minutes pour le réseau de base; 5 et 10 minutes pour la cible. Les résultats obtenus étant peu sensibles au pas de temps, seuls les résultats correspondants au pas de temps de 5 min pour la cible et au pas de temps de 10min pour le réseau de base sont présentés de façon détaillée.

La figure 5.13 et les figures 3 et 4 de l'annexe B3 présentent les erreurs de reconstitution PRMSE, PME et Emax sur le réseau de base respectivement pour les lignes de grains du 20/08/91, 27/07/90 et 30/08/92, pour le pas de temps de cumul de 10min. Pour ce qui est des erreurs de reconstitution sur la cible, elles sont présentées respectivement à la figure 5.14 et aux figures 5 et 6 de l'annexe B3 pour le pas de temps de 5min. Également, afin d'illustrer la similarité des résultats sur le réseau de base

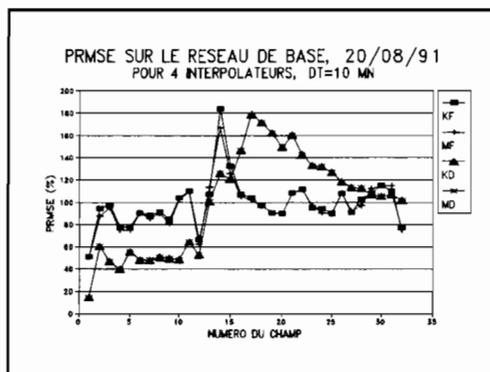
entre les pas de temps de 10 et 15min, la figure 7 de l'annexe B3 donne pour le pas de temps de 15 min les erreurs (PRMSE seulement) pour les trois lignes de grains, et la figure 8 présente les PRMSE sur la cible pour le pas de temps de 10min. L'analyse de ces différents résultats sur l'application des quatre interpolateurs KF, MF, KD et MD pour l'estimation des lignes de grain réelles conduit aux observations ci-après. A l'échelle du réseau de base, la performance des interpolateurs KDL et MDL par rapport à KF et MF, est moins bonne que celle obtenue sur les données des lignes de grains simulées que ça soit la L.G.P ou les deux L.G.B.

Néanmoins la performance des méthodes KD et MD comparée à celle des méthodes classiques KF et MF pour chacune des trois L.G.R est nettement meilleure. En effet, pour ce qui est des champs observés avant que la L.G ne commence à quitter le réseau de base (champs de numéro inférieur à 15 pour le pas de temps de 10min et inférieur 10 pour le pas de temps de 15 min), on observe une amélioration des méthodes dynamiques (KD et MD) par rapport aux méthodes classiques sur la base du PRMSE de l'ordre de 25%, 40% et 20% respectivement pour les L.G.R du 27/07/90, 20/08/91 et 30/08/92.

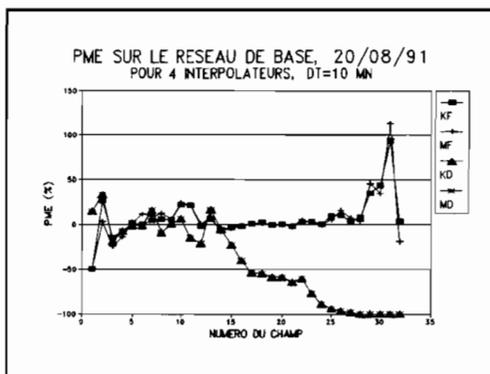
Les mêmes ordres de grandeur sont observés pour ce qui est du pas de temps de 15 min (Figure 7 de l'annexe B3). La supériorité des méthodes dynamiques est manifeste également pour ce qui est des critères PME et EMAX.

Pour ce qui est de l'échelle de la cible, les résultats sont moins intéressants que ceux obtenus à l'échelle du réseau de base. Cette situation s'explique très facilement par le fait qu'à l'échelle de la cible l'effet de la dynamique est moins perceptible. Ainsi, au pas de temps de 5 min, il existe une amélioration encore appréciable du modèle KD par rapport au modèle KF. Celle-ci est, pour les champs appartenant à la cible, de l'ordre de 20%, 25% et 25% pour respectivement les L.G.R du 27/07/90, 20/08/91 et 30/08/92. Ces valeurs sont nettement inférieures pour le pas de temps de 10 et 15 min. Ce qui confirme que l'apport de la méthode dynamique est fonction des échelles de temps et d'espace considérées. Bien que les résultats des méthodes dynamiques soient pour la plupart des champs, meilleurs que ceux des méthodes classiques, il existe des champs pour lesquels la situation totalement inverse est systématiquement observée. Ces champs sont ceux appartenant à la partie stratiforme de la L.G. Une fois que la partie convective de la ligne est passée, le comportement de la traîne est beaucoup moins structurée.

(a)



(b)



(c)

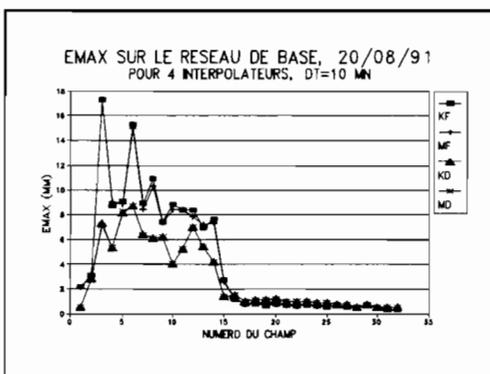


Figure 5.13: Erreurs de validation ponctuelle pour l'événement du 20/08/91 sur le réseau de base, a) PRMSE b) PME et c) Emax

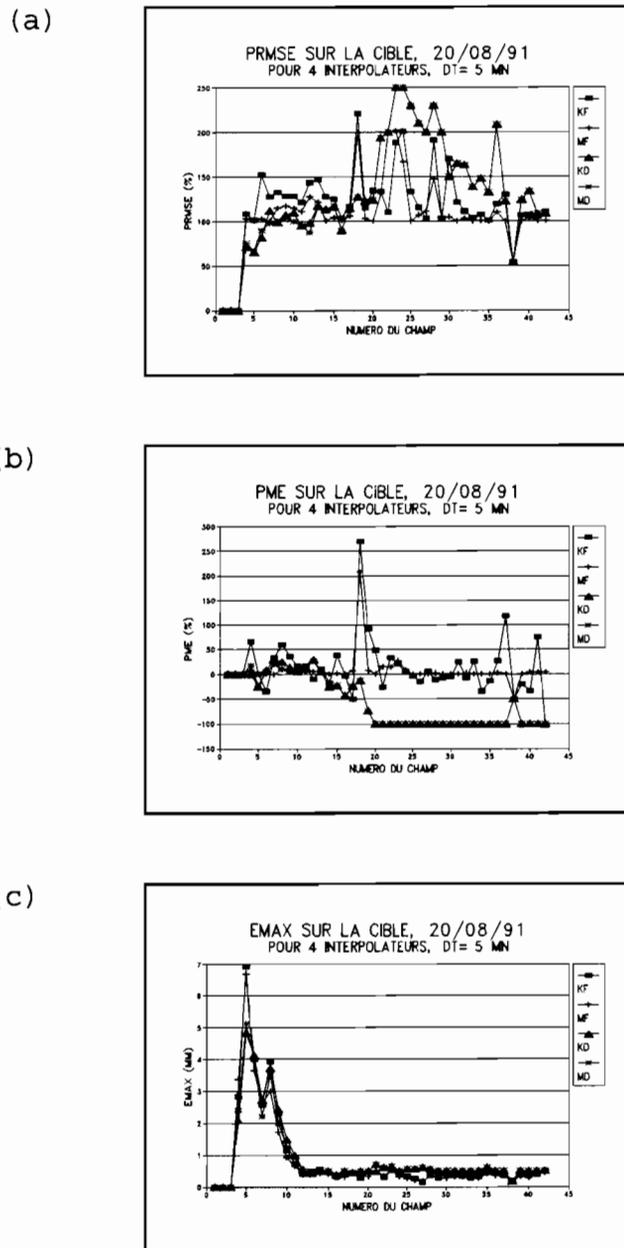


Figure 5.14: Erreurs de validation ponctuelle pour l'événement du 20/08/91 sur la Cible a) PRMSE, b) PME et c) Emax

La sous-estimation systématique (biais négatifs) de ces champs par les interpolateurs KDL et MDL provient de la dimension finie du réseau de mesure. En effet dans le repère mobile, les champs appartenants à la partie stratiforme (derniers champs), contiennent progressivement de plus en plus de stations avec des valeurs nulles.

Il découle de cette analyse que l'estimation dynamique doit se limiter à l'interpolation des champs de la partie convective de la L.G. Tant que la partie convective de l'orage n'a pas quitté le réseau de mesure, l'advection peut être mise à profit pour améliorer l'estimation par KDL. Pour les autres champs (partie stratiforme), la faible dynamique et la limite finie du réseau, conduisent à une sous-estimation systématique des valeurs reconstituées. Il est de ce fait préférable d'estimer ces champs par le biais d'une méthode classique, la moyenne arithmétique glissante par exemple. Ainsi pour une L.G.R observée sur un réseau de dimension DL dans la direction de déplacement de la L.G.R, en considérant que sa vitesse de déplacement est de U et que la durée de la partie convective est de T_c , alors les N_p premiers champs pouvant être bien estimés avec une approche dynamique sont donnés par $DL/(V*\Delta t)+T_c/\Delta t$ avec Δt étant le pas de temps de cumul ou par $T/\Delta t$ avec T étant le temps de passage de la partie convective de l'orage sur le réseau.

On vérifie très bien que pour les lignes de grains simulées que $N_p = (100/(5*50/60)+30/5= 24+6= 30$, ce qui correspond bien à la valeur observée sur les figures 5.10, 5.11 et 5.12. Pour les L.G.R, les valeurs de N_p observés à partir des différentes figures sont de 15, 15 et 13 pour le pas de temps de 10 min et de 10, 10 et 9 pour le pas de temps de 15min sur le réseau de base respectivement pour les L.G.R du 27/07/90, 20/08/91 et 30/08/92. Le biais systématique introduit par l'approche dynamique (interpolateurs KD et MD) pour les champs appartenant à la partie stratiforme, contribue de beaucoup à réduire l'efficacité de l'approche dynamique qui a été testée pour l'estimation des cumuls totaux par le cumul des champs (pour le pas de temps de cumul Δt).

5.5.3 Pluies moyennes

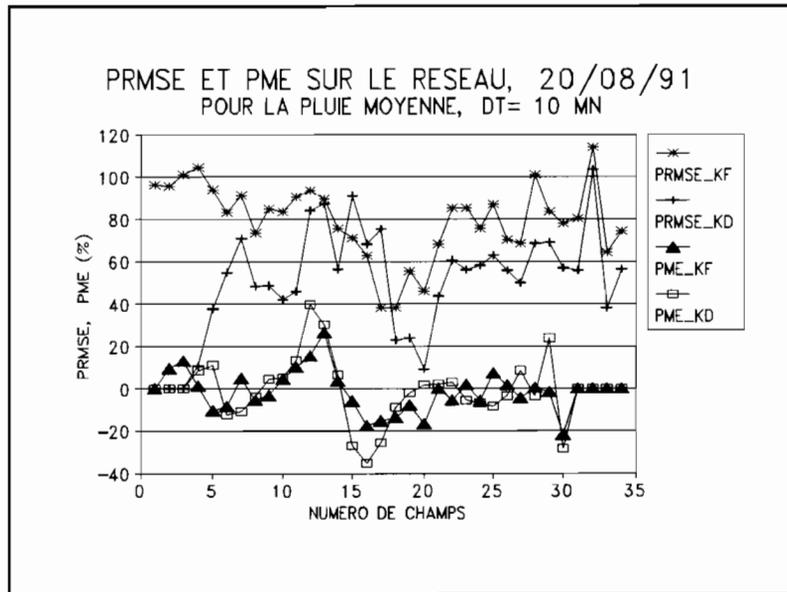
La figure 5.15 et les figures 9 et 10 de l'annexe B3 présentent les erreurs de constitution (PRMSE et PME) dans le cas de la validation sur la pluie moyenne pour les pas de temps de cumul de 10 et 15 min sur le réseau de base respectivement pour les lignes de grains du 20/08/91, 27/07/90 et 30/08/92. Dans le cas de la cible, la figure 11 de l'annexe B3 illustre les erreurs correspondantes pour le pas de temps de cumul de 10 min. Pour cette validation, le

maillage utilisé pour le réseau est de 10×10 , soit des mailles de $12\text{km} \times 11\text{km}$ de côté, alors que pour la cible, il est de 5×5 . Le choix de ces maillages est motivé uniquement par le temps d'exécution du programme. Ce temps est de 50 à 60 min pour le réseau global et de 15 min à 18 min pour la cible.

Les résultats de la validation pour l'estimation de la pluie moyenne sur le réseau de base et la cible, conduisent pratiquement aux mêmes observations que dans le cas de la validation ponctuelle à la seule différence que les erreurs sont légèrement inférieures.

A cause de l'effet de lissage (fonction de la densité du maillage), l'approche dynamique donne des meilleurs résultats même dans le cas des champs observés après passage de la partie convective de l'orage sur le réseau. L'amélioration obtenue est très significative, de l'ordre de 40%.

(a)



(b)

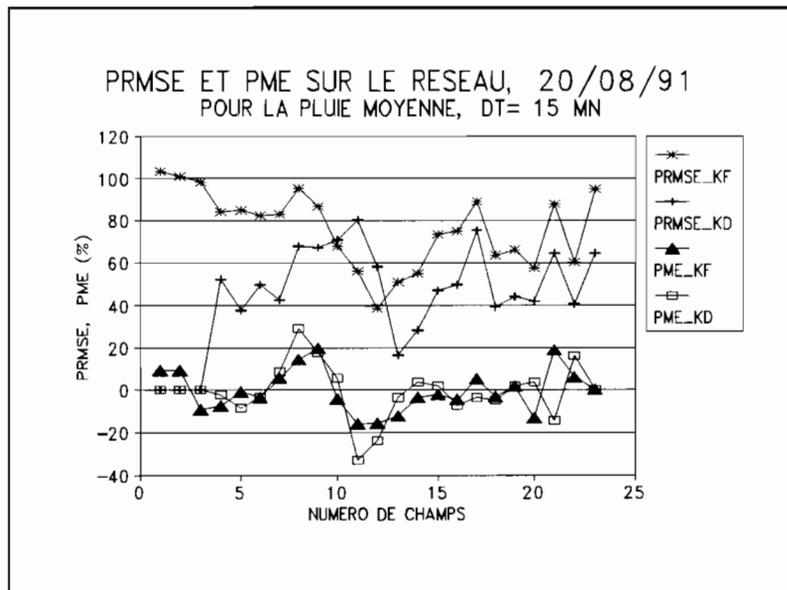


Figure 5.15: Erreurs de validation (PRMSE et PME) pour la pluie moyenne pour l'événement du 20/08/91, a) $\Delta t = 10$ min et b) $\Delta t = 15$ min.

5.6 CONCLUSION

Une nouvelle approche pour l'estimation spatio-temporelle de la pluie aux faibles pas de temps a été proposée et appliquée d'une part, à des données spatio-temporelles simulées par un modèle théorique de ligne de grains et d'autre part, aux données de trois lignes de grains (27/07/90, 20/08/91 et 30/08/92) observées au cours de l'expérience EPSAT-NIGER. Quatre interpolateurs ont été considérés. Il s'agit de deux modèles classiques d'estimation, krigeage fixe (KF) et moyenne glissante fixe (MF), et de deux modèles issus de l'approche dynamique, krigeage dynamique lagrangien (KDL) et moyenne arithmétique glissante lagrangienne (MDL). Cette approche nous permet de palier aux problèmes majeurs rencontrés dans les approches classiques du fait de l'intermittence et de l'anisotropie des champs pluvieux aux faibles pas de temps.

Les résultats de l'analyse structurale montrent que le KDL se conforme bien mieux que le KF aux hypothèses régissant la théorie du krigeage (continuité, isotropie, palier unitaire pour le cas du krigeage climatologique). Autre avantage, l'estimation dynamique distingue les deux parties de l'orage (convective et stratiforme) tant du point de vue de l'analyse structurale que de celui de l'estimation

proprement dite. Cette propriété des méthodes dynamiques est en droite ligne avec les idées développées dans la première partie de notre travail de thèse, notamment la nécessité d'algorithmes d'estimation différenciés par classe de champs pluvieux.

L'étude comparée des interpolateurs fixes (KF et MF) et dynamiques (KDL et MDL), sur la base d'une procédure de validations ponctuelle et spatiale a démontré la supériorité des méthodes dynamiques sur les méthodes fixes dans le cas du réseau de base, que ce soit sur les données simulées ou sur les données réelles. Par contre aux petites échelles d'espace (résultats de la cible), cette supériorité est moins nette. Ce qui s'explique par l'importance moindre de la dynamique à ces échelles. Les données simulées de L.G.P et L.G.B, nous ont permis de vérifier l'adéquation théorique de l'approche dynamique. La performance des méthodes dynamiques est fonction du niveau de bruit (variabilité) de la ligne de grains. Dans le cas des lignes de grains réelles, cette performance sera fonction non seulement du niveau de variabilité spatiale de la ligne, mais aussi du niveau de variabilité pour ce qui est de la vitesse de déplacement de la ligne. A cause de la nature finie du réseau de mesure, l'advection ne peut être mise à profit que pour les champs observés avant que la partie convective de l'orage ne quitte

le réseau. Afin d'améliorer les résultats sur l'utilisation de l'approche dynamique pour l'estimation des cumuls totaux, on recommande d'estimer les champs convectifs par les méthodes dynamiques et les champs stratiformes par les méthodes classiques.

Enfin, il est important de noter que la performance de l'approche dynamique est principalement due à l'effet d'advection et non à l'interpolateur, les interpolateurs KD et MD donnant des résultats identiques.

CHAPITRE VI: CONCLUSION GÉNÉRALE ET PERSPECTIVES DE RECHERCHE

L'estimation des pluies sahéliennes à l'échelle de l'événement ainsi qu'aux échelles de temps plus petite a été l'objet de ce travail de thèse. Le processus constitué par le cumul de pluie ayant des caractéristiques très différentes suivant ces deux échelles de temps, des méthodologies propres à chacun des pas de temps ont été considérées.

A l'échelle de l'événement pluvieux, nos travaux montrent sans aucun doute que la stratification des événements pluvieux se doit d'être considérée pour une bonne estimation et modélisation des précipitations pour des régions comme le Sahel où les pluies proviennent de différents systèmes météorologiques. En effet que ça soit au niveau des modèles globaux ou au niveau des méthodes de la géostatistique pour l'estimation des précipitations, on se doit de distinguer les champs pluvieux suivant leur structure spatiale. Pour le krigeage climatologique, une distinction doit être faite entre les champs continus pour lesquels la technique du krigeage climatologique peut être appliquée sans grande difficulté et les champs intermittents pour lesquels l'hypothèse de base du krigeage n'est pas vérifiée. Pour ces champs, une autre catégorie de modèle est à considérer pour

leur estimation.

Parmi les méthodes de classification analysées, la méthode de classification (DUPA) qui est basée sur le nouveau paramètre UPA s'est très illustrée. La définition de ce paramètre et les conséquences qui y en ont découlées constituent un des résultats importants de ce travail.

Il est important d'insister ici sur l'apport de la classification, plus particulièrement l'apport de la méthode DUPA sur la méthode des seuils. Nos excellents résultats au niveau de la méthode des seuils, montrent que nous avons intérêt à considérer séparément les groupes obtenus par les différentes méthodes de classification ayant fait l'objet d'application, plus particulièrement la méthode DUPA. Aussi la méthode de classification DUPA, avec sa stabilité et sa robustesse, peut être considérée pour la région Sahélienne, comme un excellent outil pour le diagnostic de la saison de pluie.

Aux faibles pas de temps, la plupart des techniques de la géostatistique disponibles pour l'estimation spatio-temporelle d'un processus espace-temps sont mal adaptées pour l'estimation de la pluie. A première vue, le krigeage spatio-temporel faciliterait la prise en compte de la dépendance

temporelle qui existe entre les champs d'un même événement pluvieux. Nos résultats sur l'application du krigeage spatio-temporel, même avec l'hypothèse de Taylor nous permettant dans le calcul de la distance euclidienne de convertir le temps en distance par le biais de la vitesse du système pluvieux, se sont avérés décevants. En effet, trois importantes caractéristiques de la pluie ne sont pas considérées dans ces techniques: la forte intermittence des champs, la non stationnarité et l'existence d'une certaine anisotropie des champs.

Afin de tenir compte de ces importantes caractéristiques pour l'estimation des pluies aux faibles pas de temps et afin d'intégrer de façon efficace la dynamique du système pluvieux, une technique originale appelée krigeage dynamique a été développée. Cette technique repose sur une exploitation conjointe de deux repères: un repère fixe et un autre mobile avec l'orage. Par construction, les champs à kriger par cette nouvelle technique sont continus et peuvent être considérés comme isotropes. La mise en oeuvre de cette méthode nécessite soit la connaissance de la vitesse en chacun des points du réseau (l'hypothèse de vitesse constante n'étant pas très bien adaptée) ou bien la connaissance par interpolation des temps de début de pluie aux différents points du réseau. Une nouvelle technique d'estimation des

vitesses frontales de déplacement de système pluvieux a été mise au point, mais n'a pas encore fait l'objet de validation par manque des données appropriées.

La supériorité de cette nouvelle approche de krigeage par rapport à l'approche classique est sans équivoque. En effet sur la base de notre application, une amélioration de l'ordre de 40% a été observée en appliquant le krigeage dynamique à la place du krigeage fixe. Il est important d'insister ici sur le fait que l'advection est prise en compte de façon plus globale qu'au niveau des méthodes récentes de prévision radar. L'advection apparait d'ailleurs comme la principale responsable de cette bonne performance, la nature de l'interpolateur n'étant pas déterminante.

Les perspectives de recherche suite à ce travail existent. Pour ce qui est des méthodes de classification, il serait intéressant d'appliquer la méthode de classification basée sur le paramètre UPA pour des données satellitaires afin d'améliorer si possible les algorithmes d'estimation de pluie par satellite surtout que le lien entre la surface pluvieuse et le seuil joue un grand rôle au niveau de ces algorithmes. Pour ce qui est du Sahel, l'analyse des événements des différents groupes obtenus par la méthode de classification DUPA est à approfondir. Aussi de nouvelles

approches d'estimation des événements intermittents une fois séparés des événements continus, sont à proposer. Par exemple pour un champ intermittent, on peut faire la distinction entre la partie pluvieuse où les techniques de la géostatistique peuvent être utilisées et la partie non continue où d'autres approches sont nécessaires. L'autre voie de recherche, qui est cette fois d'ordre théorique, concerne la caractérisation statistique du paramètre UPA. Peut-il être vulgarisé comme étant un paramètre statistique au même titre que le coefficient de variation ? Si oui, quel élément d'information apporte-t-il pour un échantillon donné.

Aux faibles pas de temps, la nouvelle méthode d'estimation des vitesses frontales du système pluvieux devrait être validée à l'aide par exemple des données radar. Pour une utilisation efficace des données disponibles, le krigeage dynamique spatio-temporel 3D est à mettre en oeuvre surtout pour l'estimation de la pluie moyenne. Notons enfin que cette approche peut être utilisée pour tous les processus dynamiques.

REFERENCES BIBLIOGRAPHIQUES

- AHMED, S., G.DE MARSILY, 1987:** Comparison of geostatistical methods for estimation transmissivity using data on transmissivity and specific capacity. *Water Resour. Res.*, 23(9): 1717-1737.
- AMANI, A., 1992:** Modélisation des précipitations: revue de littérature. Examen de Synthèse, Ecole Polytechnique de Montréal, génie civil, 110p.
- AMANI, A., J. ROUSSELLE, T. LEBEL et J.D. TAUPIN, 1993a:** Hyetogram analysis of Sahelian mesoscale convective systems, in : Sevruk B. and M. Lapin (Editors), *Precipitation Variability and Climate Change : Proceedings of the International Symposium on Precipitation and Evaporation*, Bratislava, Slovakia, September 1993, Vol. 2, pp 23-28. Printed by Slovak Hydrometeorological Institute, 1993a.
- AMANI, A., J. ROUSSELLE, T. LEBEL et J.D. TAUPIN, 1993b:** EPSAT-NIGER: Typology and spatial-temporal variability Sahelian precipitation at small time steps. *Proceedings of the Sixth International Conference on Urban Storm Drainage*, Niagara Falls, Canada, 12-17 september, J.Marsalek and H.C. Torno (Editors), vol. 1, 6-12.

- AMANI, A., T. LEBEL, J. ROUSSELLE et J.D. TAUPIN, 1995a:** Typology of Rainfall Fields to Improve Rainfall Estimation in the Sahel by the Area Threshold Method (ATM). Resoumis à Water Resources Research.
- AMANI, A., J. ROUSSELLE et F. TROCHU, 1995b:** Squall Line Rainfall Estimation at Small Time Steps in the Sahel by Dual Spatio-temporal Kriging. Soumis au Hydrological Sciences Journal.
- ANDERBERG, M.R., 1973:** *Cluster Analysis for Applications* Academic Press, New York, 1973.
- ANDRIEU, H., 1986:** Interprétation de mesures du radar Rodin de trappes pour la connaissance en temps réel des précipitations en Seine-Denis et Val-de-Marne, Thèse de Docteur ingénieur, ENPC Paris, 190p.
- ATLAS, D., D.ROSENFELD et D.A.SHORTT, 1990:** The estimation of convective rainfall by area integrals: 1. The theoretical and empirical basis, J. Geophys. Res., 95(D3): 2153-2160.
- BA, M.B et P.Y.DESCHAMPS, 1990:** Détermination des précipitations en région Soudano-Sahélienne à l'aide de l'image satellitaire: Tests de divers algorithmes. Annales Geophysiques, 8(9): 609-626.
- BACANCOURT, C., 1990:** Etude de l'intermittence et de la variabilité des champs de précipitation par une approche stochastique. Thèse de Doctorat, IMG. Grenoble. p.215.

- BACANCOURT, C., J.D. CREUTIN et M. RIVOIRARD, 1992:** A method for delineating and estimating rainfall fields. *Water Resour. Res.*, 28(4): 1133-1144.
- BARDOSSY, A., I. BOGARDI et W.E. KELLY, 1990a:** Kriging with Imprecise (Fuzzy) Variograms. I: Theory, *Math. Geology*, 22(1): 63-79.
- BARDOSSY, A., I. BOGARDI et W.E. KELLY, 1990b:** Kriging with Imprecise (Fuzzy) Variograms. II: Applications, *Math. Geology*, 22(1): 81-95.
- BARDOSSY, A., et E.J. PLATE, 1991:** Modelling Daily Rainfall Using Semi-Markov Representation of Circulation Pattern Occurrence, *J. Hydrol.*, 122(1-4): 33-47.
- BARRET, E, et D.M. MARTIN, 1981:** The use of Satellite Data in Rainfall Monitoring, Academic, New York, 340p.
- BASTIN, G., B.LORENT, C.DUQUÉ, et M.GERVERS, 1984:** Optimal estimation of the average areal rainfall and optimal selection of rain gauge locations. *Water Resour. Res.*, 20(4): 463-470.
- BEER, T., 1990:** Modelling rainfall as a fractal process. *Math.simul* 32: 119-124.
- BELLON, A., et I. ZAWADZKI, 1993:** Forecasting of Hourly Accumulations of Precipitation by Optimal Extrapolation of Radar Maps, *J. Hydrol.*, 157, 211-233.

- BOGARRDI, I., A.BARDOSSY et L.DUCKSTEIN., 1985:**
Multicriterion network design using geostatistics. *Water Resour. Res.*, 21(2): 199-208.
- BOURGAULT, G. et D.MARCOTTE, 1991:** Multivariate variogram and its application to the linear model of coregionalisation. *Math geology* 23(7): 899-928.
- BRAS, L.R. et I.RODRIGUEZ-ITURBE, 1985,** *Random Functions and Hydrology*. Addison-Wesley, 559p.
- BRAUD, I, P. CROCHET et J.D. CREUTIN, 1992:** A Method for estimating mean areal rainfall using lagrangian trend functions of intensities. Submitted to *J. Appl. Meteor.*
- BRAUD I., J.D. CREUTIN, et C. BARANCOURT, 1993:** The relation between the mean areal rainfall and the fractional area where it rains above a given threshold, *J. Appl. Meteor.*, 32, 193-202.
- BREMAUD, P.J., and Y.B. POINTIN, 1993:** Forecasting Heavy Rainfall from Rain Cell Motion using Radar Data, *J. Hydrol.*, 142: 373-389.
- CADET, M, 1988:** EPSAT- Estimation des Précipitations par Satéllite. ORSTOM, 63p.
- CERON, J.P, 1993:** L'influence des conditions météorologiques dans l'hémisphère nord sur la pluie au Sahel. *La Météorologie* 8(2): 54-61.

- CHUA, S.H. et R.L.BRAS, 1982:** Optimal estimation of mean areal precipitation in region of orographic influence. *J.Hydrol.*, 57(1/2): 23-48.
- COX, D.R, F.R.S. et V.ISHAM, 1988:** A simple spatial-temporal model of rainfall, *Proc.R.Soc.Lond.* A415. 317-328.
- CREUTIN, J.D et C.OBLED, 1982:** Objective analyses and mapping techniques for rainfall fields: An objective comparaison. *Water Resour. Res.*, 18(2): 413-431.
- CREUTIN, J.D., H.ANDRIEU et G.DELRIEU., 1989:** Une simplification du cokrigeage appliquée à l'étalonnage d'images de télédetection en hydrométéorologie. *M.Armstong(ed), Geostatistics, 2*, 1761-772.
- DELAY, F., et G. MARSILY, 1994:** The Integral of Semi-variogram: A Powerful Method for Adjusting the Semi-variogram in Geostatistics. *Math. Geology*, 26(3): 301-321.
- DELHOMME, J.P., 1978:** Kriging in the hydrosiences. *Adv. Water Resour.*, 1(5): 251-256.
- DELRIEU, G., A.BELLON et J.D.CREUTIN, 1988:** Estimation de lames d'eau spatiales à l'aide de données de pluviomètres et de radar météorologique: Application au pas de temps journalier dans la région de Montréal. *J. Hydrol.*, 98: 315-344.

- DESBOIS, M., T. KAYIRANGA, B. GNAMIEN, S. GUESSON et L. PICON, 1988:** Characterization of some elements of the Sahelian climate and their inter-annual variations for July 83, 84 and 85 from the analysis of METEOSAT ISCCP data. *J. Clim.*, 1(9): 867-904.
- DESBORDES, M., 1987:** Contribution à l'analyse et à la modélisation des mécanismes hydrologiques en milieu urbain. thèse d'Etat, Université des sciences et techniques du Languedoc (USTL), Montpellier, 242p.
- DHONNEUR, G., 1985:** Traité de météorologie tropicale. Météo-France, Paris, 150p.
- DISKIN, M.H., 1987:** On the Determination of the Speed of Moving Rainfall Patterns, *Hydrol. Scie. J.*, 32(1): 1-13.
- DISKIN, M.H., 1990:** The Speed of Two Moving Rainfall Events in Lund, *Nordic Hydrol.*, 21, 153-164.
- DONEAUD, A. A., P. L. SMITH, A.S. DENNIS et S. SENGUPTA, 1981:** A simple method for estimating convective rainvolume over area. *Water Resour. Res.*, 17(6), 167.
- DONEAUD, A. A., S.I. NISCOV, D.L. PRIEGNITZ, et P. L. SMITH, 1984:** The area- time integral as an indicator for estimating convective rain volume. *J. Appl. Meteor.*, 23, 555-561.
- DOWD, P.A., 1992:** A Review of Recent Developments in Geostatistics. *Compt. Scie*, 17(10): 1481-1500.

- EAGLESON, P.S, N.M.FENNESSEY, W.QINLIANG et I.RODRIGUEZ-ITURBE, 1987: Application of spatial poisson models to air mass thunderstorm rainfall. J. Geophys. Res., 92(D8): 9661-9678.
- FAURE, H., et J.Y. GAC, 1981: Will the Sahelian drought end in 1985. Nature, 291, 475-478.
- GAMBOLATI, G et G.VOLPI, 1979: Ground contour mapping in venice by stochastic interpolators: 1. Theory. Water Resour. Res., 15(2): 281-290.
- GANDIN, M., 1965: Objective analysis of meteorological fields. Isreal program for scientific translation, Jerusalem, Isreal, 242p.
- GEORGAKAKOS, K.P et M.L.KAVVAS, 1987: Precipitation analysis, modelling, and prediction in hydrology. Rev. Geophys., 25(2): 163-178.
- GOUTORBE, J.P., T. LEBEL, A. TANGA, P. BESSEMOULIN, J. BROUWER, H. DOLMAN, E.T. ENGMAN, J.H.C. GASH, M. HOEPPFNER, P. KABAT, Y.H. KERR, B. MONTENY, S. PRINCE, F. SAID, P. SELLERS et J. WALLACE, 1994: HAPEX-SAHEL: a large-scale study of land-atmosphere interactions in the semi-arid tropics. Annales Geophysicae, 12, 53-64.
- GUPTA, V.K et ED.WAYMIRE, 1987: On taylor's hypothesis and dissipation in rainfall. J. Geophys. Res., 92(D8): 9657-9660.

- GUPTA, V.K et ED.WAYMIRE, 1990:** Multiscaling proprieties of spatial rainfall and river flows distributions. *J. Geophys. Res.*, 95(D3): 1999-2009.
- HANDCOCK, M.S., et M.L. STEIN, 1993:** A Bayesian Analysis of Kriging. *Technometrics*, 35(4): 403-410.
- HANDCOCK, M.S., et J.R. WALLIS, 1994:** An Approach to Statistical Spatial-Temporal Modelling of Meteorological Fields. With Discussions. *J.Am.Stat.Ass.*, 89(426): 368-390.
- HAY, L.E, G.J. McCABE, JR, D.M. WOLOCK et M.A. AYERS, 1991:** Simulation of Precipitation by Weather Type Analysis, *Water Resour. Res.*, 27(4): 493-501.
- HAY, L.E, G.J. McCABE, JR, D.M. WOLOCK et M.A. AYERS, 1992:** Use of Weather Types to Disaggregate General Circulation Model Predictions, *J. Geophys. Res.*, 97(D3): 2781-2790.
- HUBERT, P et J.CARBONNELL, 1988a** Approche statistique de l'aridification de l'Afrique de l'Ouest. *J. Hydrol.*, 110: 349-367.
- HUBERT, P et J.P.CARBONNELL, 1988b:** Caractérisation fractale de la variabilité et de l'anisotropie des précipitations intertropicales. *C.R. Acad.Scie. Paris.*, t307, série II, 909-914.
- HUBERT, P et J.CARBONNELL, 1989:** Dimensions fractales de l'occurrence de pluie en climat Soudano-Sahélien *Hydrol.Continent.*, 4(1): 3-10.

- JANICOT, S, 1992a:** Spatiotemporal Variability of West African Rainfall. Part I: Regionalization and Typings. *J. Climate*, 5(5): 489-497.
- JANICOT, S, 1992b:** Spatiotemporal Variability of West African Rainfall. Part II: Associated Surface and Air Mass Characteristics, *J. Climate*, 5(5): 499-511.
- JANICOT, S et F. FONTAINE, 1993:** L'évolution des idées sur la variabilité interannuelle recente des précipitations en Afrique de l'Ouest. *La Météorologie*, 8(1): 28-53.
- JOHNSON, E.R et R.L.BRAS, 1979:** Real-time estimation of velocity and covariance structure of rainfall events using telemetered raingage data. A comparison of methods. *J.hydrol.*, 44: 97-123.
- JOURNEL, A.G et C.I.HUIJBREGTS, 1978:** *Mining geostatistics*. Academic press, New york, 597p.
- JOURNEL, A.G et R.FROIDEVEAUX., 1982:** Anisotropic Hole-Effet Modeling. *Math.geology*, 14(3): 217-239.
- KASIM, A.H.M et N.T.KOTTEGODA., 1991:** Rainfall network design trough comparative kriging methods. *Hydrol. Scie. J.*, 36, 223-240.
- KEDEM, B., D.A. SHORT et Z. KARNI, 1990a:** An analysis of the threshold method for measuring area-average rainfall, *J. Appl. Meteor.*, 29, 3-20.

- KEDEM, B., L.S. CHIU et G.R. NORTH, 1990b:** Estimation of Mean Rain Rate: Application to Satellite Observations. J. Geophys. Res., 95(D2): 1965-1972.
- KEDEM, B. et H.PAVLOPOULOS., 1991:** On the threshold method for rainfall estimation: Choosing the optimal threshold level. J.AM.Statis.A., 86(415): 626-633.
- KITANIDIS, M., 1988:** Statistical estimation of polynomial generalized covariance functions and hydrologic applications. Water Resour. Res., 19(4): 909-921.
- KOTTEGODA, N.T. et A.H.M.KASSIM, 1991:** Classification of storm profiles using crossing properties. J. Hydrol., 127: 37-53.
- KRAJEWSKI, M et J.D. CREUTIN, 1990:** Modélisation mathématique des champs de pluies: quelques types d'approches et leur application. La Météorologie 7, 4-16.
- LABORDE, J.P et R.LEMPEREUR, 1986:** Les pluies sur des petits bassins versants: une fonction aléatoire dont on peut estimer le variogramme. Hydrol. Cont., II(1): 3-13.
- LEBEL, T., 1984:** Moyenne spatiale de la pluie sur un bassin versant: estimation optimale, génération stochastique et GRADEX des valeurs extrêmes. Thèse de Docteur-Ingénieur, Institut National Polytechnique de Grenoble, Grenoble, France.

- LEBEL, T et G.BASTIN, 1985:** Variogram identification by the mean-squared interpolation error method with application to hydrologic fields. *J. Hydrol.*, 77: 31-56.
- LEBEL, T, G.BASTIN, C.OBLED et J.D.CREUTIN, 1987:** On the accuracy of areal rainfall estimation: A case study. *Water Resour. Res.*, 23(11): 2123-2134.
- LEBEL, T. et J.P.LABORDE, 1988:** A geostatistical approach for areal rainfall statistics assesement. *Stoch. Hydrol. Hydraul.*, 2: 245-261.
- LEBEL, T., H.SAUVAGEOT, M.HOEPFFNER, M.DESBOIS, B.GUILLOT et P.HUBERT, 1992:** Rainfall Estimation in the Sahel: The EPSAT-NIGER Experiment, *J. Hydrol. Scie.*, 37(2): 201-215.
- LEBEL, T. et LEBARBE, L, 1995:** Rainfall monitoring during HAPEX-Sahel: 2. Point and areal estimation at the event and seasonal scales. *Accepté pour publication dans le J. Hydrol.*
- LEBEL, T., A.AMANI et J.D.TAUPIN, 1995a:** La pluie au Sahel, une variable rebelle à la régionalisation. In Hoepffner, M., T. Lebel, B. Monteney (Editors), *Interactions Surface Continentale/ Atmosphère: l'expérience HAPEX-Sahel.* (Reprints of the 10eme journées hydrologiques de l'ORSTOM, Montpellier, septembre 1994). ORSTOM, Paris, sous presse.

- LEBEL, T., J.D. TAUPIN, et D'AMATO, 1995b:** Rainfall monitoring during HAPEX-Sahel: 1. General rainfall conditions and climatology. Accepté pour publication dans le J. Hydrol.
- LEROUX, M., 1988:** La variabilité des précipitations en Afrique Occidentale: les composantes Aérologiques du problème. Veille Climatique Sttélitaire, ORSTOM/LANION, 22(1): 26-45.
- MARAND, C. et J.F.ZUMSTEIN., 1990:** La modélisation des précipitations moyennes annuelles appliquée au massif Vosgien. Hydrol.Continent., 5(1): 29-39.
- MARSHALL, R.J, 1980:** The estimation and distribution of storm mouvement and storm structure, using a correlation analysistechnique and rain-gauge data. J.hydrol., 48: 19-39.
- MATHERON, G., 1970:** La théorie des variables régionalisées et ses applications. Chahier du centre de morphologie mathématique de Fontainebleau. F.5. p 212.
- MEJIA, J.M. et J.ROUSSELLE, 1976:** Disaggregation model in hydrology revisited. Water Resour. Res., 12(2): 185-186.
- MEYERS, D.E., 1981:** Matrix formulation of co-kriging, J. Math. Geol., 14(3): 249-257.
- MEYLAN, P., 1986:** Régionalisation des données entachées d'erreurs de mesure par le krigeage: application à la pluviométrie. Hydrol.Cont., I(1): 25-34.

- MORRISEY, M.L., 1994:** The Effect of Data Resolution on the Area Threshold Method. *J.Appl. Meteor.*, 33, 1263-1270.
- MORRISEY, M.L., W.F. KRAJEWSKI et M.J.MCPHADEN, 1994:** Estimating Rainfall in the Tropics using the Fractional Time Raining. *J.Appl. Meteor.*, 33, 387-393.
- MOUSSA, R., 1991:** Variabilité spatiotemporelle et modélisation hydrologique, application au bassin du Gardon D'Anduze. Thèse de Doctorat, Université de Montpellier II sciences techniques de Languedoc, 314p.
- MYERS, D.E. et A.JOURNEL., 1990:** Variograms with zonal anisotropies and noninvertible kriging systems. *Math.geology*, 22(7): 607-617.
- NICHOLSON, S.E, 1980:** The Nature of Rainfall Fluctuations in Subtropical West Africa, *Mon. Wea. Rev.*, 108, 473-487.
- NICHOLSON, S.E, 1981:** Rainfall and Atmospheric Circulation during Drought Periods and Wetter Years in West Africa, *Mon. Wea. Rev.*, 109, 2191-2208.
- NICHOLSON, S.E, 1986:** The Spatial Coherence of African Rainfall Anomalies: Interhemispheric Teleconnections. *Jour. Clim. Appl. Met.*, 25: 1365-1381.
- NICHOLSON, S.E, k. JEEYOUNG et J. HOOPINGARNER, 1988:** Atlas of African Rainfall and its Interannual Variability. Dept. of Met., The Florida State University of Tallahassee, Florida. 240p.

- NIEMCZYNOWICZ, J. et P.DAHLBLOM, 1984:** Dynamic proprieties of rainfall in lund, Nordic Hydrol., 15: 9-24.
- OMRE, H., 1987:** Bayesian Kriging. Merging Observations and Quantified Guess in kriging. Math. Geology, 19(1): 25-39.
- PAWLOWKY, V., R.A. OLEA et J.C. DAVIS, 1993:** Boundary Assessment Under Uncertainty: A Case Study. Math. Geology, 25(2): 125-144.
- RENDU, J.M., 1980:** Disjunctive Kriging: Comparison of theory with actual results. Math. Geol., 12(4): 305-320.
- RIBEIRO, L.T., F.H.MUGE et M. ARMSTONG, 1989:** A geostatistical approach to the modelling of a piezometric field. M.Armstong(ed), Geostatistics., 2: 651-660.
- RODRIGUEZ-ITURBE, I. et J.M.MEJIA, 1974a:** On the transformation of point rainfall to areal rainfall. Water Resour. Res., 10(4): 729-735.
- RODRIGUEZ-ITURBE, I. et J.M.MEJIA, 1974b:** The design of rainfall networks in time and space. Water Resour.Res., 10(4): 713-728.
- RODRIGUEZ-ITURBE, I., D.R.COX, F.R.S et P.S.EAGLESON, 1986:** Spatial modelling of total rainfall. Proc.R.Soc.Lond A 403: 37-50.

- RODRIGUEZ-ITURBE, I., B.F.DE POWER et J.B.VALDES, 1987a:**
Rectangular pulses point process models for rainfall:
Analysis of empirical data. J. Geophys. Res., 92(D8):
9645-9656.
- RODRIGUEZ-ITURBE, I., D.R.COX, F.R.S. et V.ISHAM, 1987b:** Some
models for rainfall based on stochastic point processes.
Proc.R.Soc.Lond. A410: 269-288.
- RODRIGUEZ-ITURBE, I., D.R.COX, F.R.S. et V.ISHAM, 1988:** A
point process model for rainfall: futher developments.
Proc.R.Soc.Lond. A417: 283-298.
- ROSENFELD, D., D. ATLAS et D.A. SHORT, 1990:** The estimation
of Convective Rainfall by Area Integrals. 2- The Height
Area Rainfall Threshold (HART) method. J. Geophys. Res.,
95(D3): 2161-2176.
- ROUHANI, S. et T.J.HALL, 1989:** Space-time kriging of
groundwater data. M.Armstong (ed), Geostatistics, 2:
639-650.
- ROUHANI, S. et D.E.MEYERS., 1990:** Problems in space-time
kriging of geohydrological data. Math.Geology, 22(5):
611-623.
- ROUHANI, S. et H.WACKERNAGEL, 1990:** Multivariate
geostatistical approach to space-time data analysis.
Water Resour. Res., 26(4): 585-591.

- SAUVAGEOT, M., 1994:** The probability density function of rain rate and the estimation of rainfall by area integrals. *J. Appl. Meteor.*, 33, 1255-1262.
- SEN, Z., 1989:** Cumulative Semivariogram Models of Regionalized Variables, *Math. Geology.*, 21(8): 891-903.
- SEN, Z., 1992:** Standard Cumulative Semivariograms of Stationary Stochastic Processes and Regional Correlation. *Math. Geology.*, 24(4): 417-435.
- SEO, D.J, W.F.FRAJEWSKI et D.S.BOWLES, 1990a:** Stochastic interpolation of rainfall data from rain gages and radar using Cokriging: 1- design of experiments. *Water Resour. Res.*, 26(3): 46-477.
- SEO, D.J, W.F.FRAJEWSKI et D.S.BOWLES, 1990b:** Sochastic interpolation of rainfall data from rain gages and radar using Cokriging: 2-Results. *Water Resour. Res.*, 26(5): 915-924.
- SHAMSI, U.M, R.G.QUIMPO et G.N.YOGANARASIMHAN, 1988:** An application of kriging to rainfall network design. *Nordic. Hydrol.*, 19: 137-152.
- SHORT, D.A., D.B. WOLFF, D.R. ROSENFELD et D. ATLAS, 1993:** A study of the threshold method utilizing raingage data, *J. Appl. Meteor.*, 32: 1379-1387.

- SIRCOULON, J., 1976:** Les données hydropluviométriques de la sécheresse récente en Afrique intertropicale. Comparaison avec les sécheresses 1913 et 1940. Cah. ORSTOM. Hydrol., XIII(2).
- SMITH, J.A et A.F.KARR, 1985a:** Statistical inference for point process models of rainfall. Water Resour. Res., 21(1): 73-79.
- SMITH, J.A et A.F.FARR, 1985b:** Parameter estimation for a model of space-time rainfall. Water Resour. Res., 21(8): 1251-1257.
- SOLOW, A.R., 1993:** On the Efficiency of the Indicator Approach in Geostatistics. Math. Geology, 25(1): 53-57.
- TABIOS III, G.Q et J.D.SALAS, 1985:** A comparative analysis of techniques for spatial interpolation of precipitation. Water Resour. bull., 21(3): 365-380.
- TAUPIN, J.D., T.LEBEL, F.CAZENAVE, M.GREARD, J.KONG, J.LECOCQ, M.ADAMSON, N.d'AMATO et A.B.MOHAMED, 1992:** EPSAT-NIGER, Campagne 1992. ORSTOM-DMN, 63p.
- TAUPIN, J.D, A. AMANI et T. LEBEL, 1993:** Small scale spatial variability of the annual rainfall in the Sahel, in : Bolle H.-J., Feddes, R.A. and Kalma, J. (Editors), *Exchange Processes at the Land Surface for a Range of Space and Time Scales* (Proceedings of the Yokohama Symposium, July 1993), IAHS Publ. N° 212, pp593-602.

- THAUVIN, V. et T. LEBEL, 1991: EPSAT-NIGER: study of rainfall over the Sahel at small time steps using a dense network of recording raingauges. Hydrol. Processes, 5, 251-260.**
- THAUVIN, V., 1992: Etude de la répartition spatiale des précipitations en milieu sahélien à l'aide du réseaudense de pluviographes de l'expérience EPSAT-NIGER: application à la détermination de la précision des moyennes surfaciques au pas de temps de l'événement pluvieux. Thèse de Doctorat, USTL, Montpellier.**
- THIAO, M.W., 1988: Contribution à l'estimation des précipitations au Sahel par l'imagerie satellitaire. Thèse de Docteur-Ingénieur, Université Blaise Pascal. 300p.**
- VALDES, J.B, I.RODRIGUEZ-ITURBE et V.K.GUPTA, 1985: Approximations of temporal rainfall from a multidimensional model. Water Resour. Res., 21(8): 1259-1270.**
- VALENCIA, D. et J.C.SHAAKE, 1972: Disaggregation processes in stochastic hydrology. Water Resour. Res., 9(3): 580-585.**
- VILLENEUVE, J.P, G.MORIN, B.BOBEE, D.LEBLANC et J.P.DELHOMME, 1979: Kriging in the design of streamflow sampling networks. Water. Resour. Res.,15(6): 1833-1840.**

- WAYMIRE, ED et V.K.GUPTA, 1981a:** The mathematical structure of rainfall representations: 1- A review of the stochastic rainfall models. *Water. Resour. Res.*, 17(5): 1261-1272.
- WAYMIRE, ED et V.K.GUPTA, 1981b:** The mathematical structure of rainfall representations: 2- A review of the theory of pointprocess. *Water. Resour. Res.*, 17(5): 1273-1285.
- WAYMIRE, ED et I.RODRIGUEZ-ITURBE, 1984:** A spectral theory of rainfall intensity at the meso-B-scale. *Water. Resour. Res.*, 20(10): 1453-1465.
- WILSON, L.L, et D.P. LETTENMAIER, 1992:** A Hierarchical Stochastic Model of Large-Scale Atmospheric Circulation Patterns and Multiple Station Daily precipitation, *J. Geophys., Res.*, 97(D3): 2791-2809.
- YATES, S.R, A.W.WARRICK et D.E.MEYERS, 1986a:** Disjunctive kriging: 1 Overview of estimation and conditional probability. *Water. Resour. Res.*, 22(5): 615-621.
- YATES, S.R, A.W.WARRICK et D.E.MEYERS, 1986b:** Disjunctive kriging: 2 examples. *Water. Resour. Res.*, 22(5): 623-630.
- ZAWADZKI, I., 1973a:** Errors and fluctuations of rain gauge estimates of areal rainfall. *J. Hydrol.*, 18: 243-255.
- ZAWADZKI, I., 1973b:** Statistical proprieties of precipitation patterns, *J. Appl. Meteorol.*, 12: 469-472.

- ZAWADZKI, I., 1987: Fractal structure and exponential decorrelation in rain. *Water. Resour. Res.*, 23(8): 9586-9591.
- ZAWADZKI, I., J. MORNEAU et R. LAPRISE, 1994: Predictability of Precipitation Patterns: An Operational Approach, *J. Appl. Meteor*, 33(12): 1562-1571.
- ZIMMERMAN, D.L. et M.B. ZIMMERMAN, 1991: A comparaison of spatial semivariogram estimators and corresponding ordinary kriging predictors. *Technometrics*, 33(1): 77-91.
- ZIMMERMAN, D.L, 1993: Another Look at Anisotrpy in Geostatistics, *Math Geology*, 25(4): 453-470.

ANNEXE A :

**ARTICLES SOUMIS POUR PUBLICATION DANS DES REVUES
SCIENTIFIQUES**

ANNEXE A1:

Typology of Rainfall Fields to Improve Rainfall Estimation in the Sahel by the Area Threshold Method (ATM), by A. Amani, T. Lebel, J. Rousselle and J.D. Taupin. Soumis après corrections des réviseurs au Water Resources Research.

Manuscript resubmitted to Water Resources Research

**TYPOLOGY OF RAINFALL FIELDS TO IMPROVE RAINFALL ESTIMATION IN
THE SAHEL BY THE AREA THRESHOLD METHOD (ATM)**

Abou AMANI

Département de Génie Civil, Ecole Polytechnique, C.P.6079,
Succ. centre- ville, Montréal, Québec, H3C 3A7, Canada

Thierry LEBEL

Groupe PRAO, Laboratoire d'étude des Transferts en Hydrologie
et Environnement; CNRS-INPG-UJP, BP 53, 38041 Grenoble
cedex 9; France

Jean ROUSSELLE

Département de Génie Civil, Ecole Polytechnique, C.P.6079,
Succ. centre-ville, Montréal, Québec, H3C 3A7, Canada

And

Jean Denis TAUPIN

Groupe PRAO, ORSTOM Mission au NIGER, BP 11416, Niamey, Niger

Abstract The stratification of rainfall fields to improve specific rainfall models is a subject that has received relatively little attention in the literature. It is shown here that objective stratification techniques, based on the characteristics of the rainfall fields (intensities and spatial distribution) rather than on the associated meteorological patterns, can produce meaningful results in terms of the ATM model and climatology. Four approaches for rainfall classification, using raingauge data, are proposed in order to improve the average areal rainfall estimation in the Sahel by the ATM model. Two of them are based on the structural behavior of the rainy area function against a threshold. Based on this function, a new parameter, called the UPA (Under Profile Area) has been proposed for the classification of rainfall fields. The groups obtained from the method based on this parameter are characterized by different average spatial structures. A significant improvement on the ATM model is observed by considering classification based on the UPA parameter. An average reduction of 34% of the root mean squared error is observed in a validation term. This improvement is a direct consequence of the fact that the optimal thresholds are different from one group to another, which is an important point when considering the impact of classification on the ATM model.

1- WHY A TYPOLOGY OF RAINFALL FIELDS?

For a large majority of the rainfall models presented in the literature, all the available rainfall data are used without any stratification, or, at best, with a crude separation based on some elementary characteristics of the rainfall field (its space-averaged value, for instance). Yet, it is well known that rainfall data originate from different types of rain storms, and Houze (1981) has shown that a separation between stratiform and convective rainfall is possible for many different precipitation systems. On a synoptic scale, meteorologists have developed the concept of weather type, which has proved efficient in relating the rainfall features on a large scale with some relevant meteorological variables (e.g. Benichou et al., 1988).

However, in many cases it is very difficult, if not impossible, to identify the meteorological nature of a rainfall event in a limited area of study, the more so if only inappropriate meteorological data are available. Furthermore, on smaller scales, which are of interest to hydrologists, relationships between the atmospheric circulation and rainfall characteristics on the ground are much more difficult to find. This is partly due to the lack of proper meteorological data and partly due to the

complexity of rainfall distribution in time and space when viewed with a resolution of, say, a few kilometres and a few hours (or less). On such scales, two events triggered by the same meteorological situation may have a different ground signature.

Besides a broad set of applications for rainfall modeling, rainfall stratification is especially useful for satellite-based rainfall estimation algorithms. For instance, the lower accuracy in satellite estimates during phase 2 of the GARP Atmospheric Tropical Experiment (GATE) was attributed by *Woodley et al. (1980)* to the presence of a higher proportion of a different kind of precipitating convective system as compared to phases 1 and 3. In their review of the grid-cell approach for satellite rain estimation over Florida, *Negri and Adler (1987)* felt the necessity to group the data into classes. However, the results of the proposed stratification technique, based on the mean rain rate, were inconsistent, and the authors implied that other classification schemes should be designed. *Desbois et al. (1988)* point out that in the Sahel, a distinction has to be made between local convection, mesoscale convection and tropical squall lines.

Given the theoretical and practical difficulties of implementing stratification schemes based on a meteorological analysis, it appears appropriate to derive alternate schemes based on the rainfall data set only, since they are the most readily available to users. Since the possible uses of a rainfall typology are numerous, the choice of a proper classification procedure will have to be based on the aims of this classification. Stratification of rainfall events in their stratiform and convective parts is considered to be very important in rainfall estimation by satellite remote-sensing. Also, the separation of rainfall events of a different nature or having different statistics is necessary to improve rainfall estimation by global methods (satellite rainfall estimation methods) like the ATM (Area Threshold Method) model.

The ATM model (Short et al., 1993; Rosenfeld et al., 1990), is derived from the ATI (Area Time Integral) model, first proposed by *Doneaud et al. (1981)* for the estimation of convective rain volume for a given rainfall event. In this study, we are interested in classification schemes based on raingauge data, in order to: i) improve the areal rainfall depth estimation by the ATM model, and ii) improve our understanding of the Sahelian rainfall climatology.

The topics covered in this paper are presented in the following order: area and data set of the study, review of possible classification criteria, classification methods relevant to rainfall estimation by the threshold method, applications of the classification methods and their contribution for improving Sahelian precipitation estimation by the ATM model, and conclusion.

2- A SAHELIAN CASE STUDY

Estimating rainfall over the Sahel (the semi-arid region lying south to the Sahara), where the networks of measurement stations are not highly developed, is necessarily more difficult than in a temperate region. This is because of the intermittent nature, in time and space, of the data they provide. It is of vital importance in this setting, therefore, to combine direct rainfall measurements (ground-based rainfall data) and data obtained through remote-sensing. However, the performance of the algorithms designed to integrate these raw data is a function of the homogeneity of the data. In order to test and improve these algorithms, the EPSAT-NIGER experiment (*Lebel et al., 1992*) was conducted in the Niamey region in which data were obtained through very accurate sampling of rainfall systems by means of a dense network of rainfall gauges and by radar (Figure 1). At this

time, the area under study, measuring 120 km x 110 km, was equipped with about one hundred static-memory raingauges. This network consisted of a basic network of raingauges installed on a regular grid and spaced approximately 12.5 km apart, at the center of which was a denser network covering an area of 400 km².

The various rainfall events observed in the Sahel originate mainly from three types of meteorological systems: isolated convective systems, organized moving convective systems and squall lines, the two latter being grouped together under the heading of Mesoscale Convective Systems (MCS). The amounts of precipitation recorded during a rainfall event, whatever its origin, are extremely variable in space, however. The rainfall regime includes a dry season (October-April) and then a rainy season (May-September). The rainfall events observed during the EPSAT-NIGER experiment have been defined as follows: i) a minimum of 30% of operating raingauges have recorded more than 2.5 mm of rainfall, and ii) a minimum of 30 minutes must separate two consecutive events. Respectively 37, 47 and 49 such events were recorded in 1990, 1991 and 1992 (*Lebel et al., 1995b*), totalling more than 95% of the seasonal rainfall (the remaining rainfall, less than 5%, is produced by isolated showers at the beginning and end of the rainy season).

Another feature of rainfall in this region is that the average annual rainfall is strongly correlated to the number of events during which rainfall accumulation exceeds a given threshold (*Taupin et al., 1993*). Recently, *LeBarbé and Lebel (1995)* point out that seasonal rainfall depth (week, month or year) in the Sahel can be easily estimated by the product of climatological mean rainfall depth per event and the number of recorded events. This observation can be improved if we can distinguish between the different contributions in number and total rainfall accumulation of the different types of rainfall events. Also, a robust method stratifying the Sahelian rainfall events can contribute to the improvement of the climatological diagnostics for the rainy season.

A major field of interest for classification is to improve areal rainfall estimation models, such as the ATM, widely proposed for rainfall estimation from remote sensing data. Since the Sahelian rainfall is produced by a few different precipitation systems, it would be helpful to classify rainfall fields with a view to improving the ATM model in the Sahel. An objective meteorological stratification of rainfall events in the Sahel is difficult to implement because of the lack of pertinent meteorological data and the soft gradient of Sahelian pressure fields. Therefore, classification based on raingauge data is very

attractive. Since it is based on rainfall characteristics observed at the ground level, the classification can also be useful for other specialists interested in water resources in the Sahel.

3- STRATIFICATION BASED ON RAINGAUGE DATA: A REVIEW OF POSSIBLE CRITERIA.

Since no widely accepted stratification method is available, several approaches will have to be compared, based on parameters characterizing the spatial or spatio-temporal features of the rainfall events. These parameters may be directly connected to the theoretical assumptions of the estimation model(s) of interest, or we can define, a priori, certain parameters characterizing the basic structure of the rainfall field, but with no direct connection with the estimation model(s).

One can imagine a very large array of rainfall stratification techniques. For rainfall field estimation models, the classification criteria must be necessarily linked to the spatial or spatio-temporal behavior of the rainfall fields. *Taupin et al. (1993)*, *Lebel et al. (1995a)* and *Lebel and Lebarbé (1995)* insisted on the high spatial variability of the rainfall in the Sahel at the event and

annual scales. Thus, a natural classification criterion for rainfall fields can be the spatial structure of the total rainfall depths of the event. Various attempts carried out by Amani (1995) to stratify the Sahelian rainfields according to such statistical parameters as the variogram parameters or the coefficient of variation of the rainfields proved to be not very efficient. We were thus led to consider alternate approaches.

Considering the spatio-temporal behavior of rainfall fields, the procedure proposed by *Kottegoda and Kassim (1991)* for the classification of temporal rainfall structures (punctual hyetographs) was generalized in space-time (areal hyetographs) by *Amani et al. (1993)*. The classification based on these approaches has two main drawbacks: i) the computation of the rainfall hyetograph type depends on the time-step duration of the hyetograph used for the analysis (*Amani et al., 1993*); ii) it cannot account for the rainfall distribution in space. An extension of these methods to describe the spatial structure of the rainfall as measured by a raingauge network is therefore proposed in the next section as a classification scheme directly related to the ATM method.

Two important characteristics to be considered in the classification schemes are the intermittent nature and intensity of rainfall fields in the Sahel. Even though intermittent rainfall events are generally not intense events, we can observe from Figure 2 that non intense events can be non intermittent events. As we can see from that Figure, there is a significant difference between the conditional and the non-conditional mean areal rainfall depth for very intermittent events. A good classification method for Sahelian rainfall must consequently have a separation of events based on their degree of intermittence or their intensity or both.

4- SELECTION OF METHODS REVELANT TO RAINFALL ESTIMATION BY THE ATM ALGORITHMS.

Classification methods capable of improving rainfall field estimation by the ATM model must necessarily be based on parameters characterizing the spatial structure of rainfall fields in terms of rainfall accumulation and/or rainfall extension. Several classification methods are proposed and compared in *Amani (1995)*. These methods are based on criteria characterizing the spatial distribution of the event: rainfall depths, the intensity of the event and its spatio-temporal behavior. The comparison presented in

Amani (1995) led us to select three classification methods which are better and more robust than the others.

They are: i) Method I, based on the clustering analysis of parameters characterizing the rainfall event intensity; ii) method II, based on a crossing analysis of the rainy area function versus threshold (this is an extension in space of the method proposed by *Kottegoda and Kassim (1991)* for the classification of the hyetograph of rainfall events observed at a given raingauge); and iii) method III, based on the analysis of the cumulative distribution function of the newly introduced UPA parameter. The UPA (Under Profile Area) is defined as the area under the rainy area function obtained by method II. Also, the two criteria, the UPA parameter and the parameters used in method I are combined to produce a fourth method.

4.1 Method I

For a rainfall event, the maximum total $C_{M\theta}$ observed over the whole raingauge network for various periods of accumulation θ are computed:

$$C_{M\theta} = \text{Max}(R_{\theta}(i)), \quad i=1,n \quad (1)$$

Usually, not all the maxima are recorded on the same gauge. This set of maximum values is thus a global indicator of the rainfall pattern over the area of study. Here, five durations θ were chosen: 5, 10, 15, 30 and 60 minutes. Other meaningful parameters may also be considered, such as total event recorded depth, mean recorded depth and event duration. The five parameters retained here proved to be the most significant for application to Sahelian rainfalls.

4.2 Method II

With this method, we propose to use the profile of the rainy area function as a classification parameter. The rainy area is the area over which the rain rate exceeds a given threshold. It is scaled by the total area of the study. Figures 3 and 4 illustrate the rainy-area approach. Let C be a rainfall threshold; then define AC as the percentage of the rainy area where rainfall exceeds this threshold C . AC is estimated from the recorded spatial data using the indicator technique, where an indicator variable is computed as:

$$I(x,y,t) = \begin{cases} 1 & \text{if } R(x,y,t) \geq C \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

The value of AC is then given by:

$$AC = E[I(x, y, t)] \quad (3)$$

In practice, AC is calculated for a rainfall field accumulation over a time period θ (θ may be a fixed time step, say one hour, or the event duration as considered in this study). The corresponding AC function is denoted by AC_θ and estimated as:

$$AC_\theta = \frac{1}{n} \sum_{i=1}^n I_\theta(x_i, y_j) \quad (4)$$

where $I_\theta(x_i, y_j)$ is the indicator variable associated with gauge i for the accumulation period θ .

AC_θ is a decreasing function of C . The complementary function

$BC = (1-AC)$ is the area of rainfall below C . In order to compare the BC functions for different events, the threshold axis must be scaled by a scaling parameter C_M

$$C_M = \text{Max}(R_\theta(i)), \quad i=1, \quad n \quad (5)$$

The threshold variate then becomes:

$$C^* = C/C_M \quad (6)$$

BC is a function analogous in time to a mass curve.

Following the crossing procedure applied by *Kottegoda and Kassim (1991)* to the hyetograph analysis, it is possible to count the number of crossing between the bisector and the observed BC curve, which produces a direct classification (Figure 5). The basic types of the classification are numbered 1a, 1b, 2a, 2b, ..and so on. The order number is defined as follows: if x is the number of intersections between the mass curve and the bisector excluding the two ends points, then the order of the mass curve is given by $(x+1)$. If the part of the mass curve before the first intersection is under the bisector, then the class of the mass curve is a. Otherwise, it is b.

4.3 Method III

This method is based on the *BC* function defined above. The area under the *BC* function (*UPA*) is computed (Figure 6). Since both axes are scaled, *UPA* takes values between 0 and 1. The value 0 corresponds to a spatial uniform rainfall and the value 1 corresponds to a Dirac-like event. It is demonstrated in section 4.4 that the *UPA* parameter is related to the coefficient of variation under certain assumptions. It will be observed in section 5 that the cumulative distribution of the *UPA* values obtained from the 133 Sahelian rainfall events leads to a natural classification in three groups. The *UPA*

parameter has some interesting statistical meaning, as will be shown below.

4.4 Links between the UPA parameter and CV

First, let us consider, the relationship between the rainy area AC and the threshold C . Studying this relationship for each event, it appears that $\log (AC)$ is linearly correlated with C . In 1991, 9 events only, out of a total of 47, have an r^2 under 0.90, and 26 events have an r^2 above 0.95. The results for 1990 and 1992 are even better, with respectively 5 events (out of a total of 37) and 4 events (out of a total of 49) with an r^2 under 0.90, and 20 events (1990) and 37 events (1992) with an r^2 above 0.95. For a given event j , we may thus write:

$$AC_j = \exp(-C/a_j) \quad (7)$$

It should be noted that the empirical justification of the expression (7) is backed by the finding of *Lebel et al.* (1995a) that the cumulative distribution function (cdf) of the event rainfall is exponential. Since there is an obvious, even if not straightfoward, relationship between the exceedance probability of a given rainfall event value Z_j and the area a_j corresponding to this value, it is not surprising

that the cdf and area distribution function are of the same functional type. Using the scaled threshold C^* defined in expression (6), (7) may be rewritten:

$$AC_j = \exp(-C^* \cdot C_{Mj}/a_j) \quad (8)$$

where C_{Mj} is the recorded maximum of $R_T(x,y)$ for event j . The scaled spatial profile BC_j is then expressed as:

$$BC_j = 1 - \exp(-C^* \cdot C_{Mj}/a_j) \quad (9)$$

BC_j , C_{Mj} and a_j are all realizations of random variables associated with the random process $R_T(x,y)$. This process is defined by its probability density function, with mean μ_T , standard deviation σ_T and coefficient of variation CV . From the frequency analysis point of view, the maximum rainfall recorded by a network of n independent gauges is also a random variable corresponding to the quantile Q_n . The expected probability of non exceedance of Q_n , P_n , is given by $(n-a) / (n+b)$, with the a and b parameters taking values between 0 and 1 (the values of a and b required to obtain an unbiased estimate of P_n depend on the probability distribution of $R_\theta(x,y)$; see *Cunnane, 1978*, for a review). Q_n may be expressed as:

$$Q_n = \mu + \lambda \cdot \sigma \quad (10)$$

where λ is a statistic which depends on P_n and on the probability distribution of $R_T(x, y)$. For a given realization (rain event) j , this expression yields:

$$C_{Mj} = m_j + l_j \cdot s_j \quad (11)$$

where m_j and s_j are the computed mean and standard deviation of the rainfall recorded by the network for the event j . l_j is the frequency factor of the maximum cumulated rainfall and it is linked to the return period of C_M recorded during event j . Substituting (11) into (9) yields:

$$BC_j = 1 - \exp(-C^* \cdot (m_j + l_j \cdot s_j) / a_j) \quad (12)$$

For the area under the scaled spatial profile, UPA_j , being the integral of BC for C^* varying from 0 to 1, we get:

$$UPA_j = \int_0^1 1 - \exp(-C^* (m_j + l_j s_j) / a_j) dC^* \quad (13)$$

Defining,

$$K_j = (m_j + l_j \cdot s_j) / a_j \quad (14)$$

leads to:

$$UPA_j = \left[1 - \frac{1}{K_j} \exp(-K_j C^*) \right]_0^1 \quad (15)$$

or:

$$UPA_j = 1 - \frac{1}{K_j} + \frac{\exp(-K_j)}{K_j} \quad (16)$$

Computing the values a_j and m_j for all the available events j shows that m_j may be used as an estimate of a_j , which means that (14) may be written simply as:

$$K_j = 1 + l_j \cdot CV_j \quad (17)$$

Furthermore, the values of CV_j and l_j show that $l_j \cdot CV_j$ is almost always greater than 1 and most often greater than 2. The Figure 7 presents the histogram of the frequency factor values l .

Consequently, $\exp(-K_j)$ is small compared to 1 and the third term in the summation of (16) may be neglected, so that (16) becomes:

$$UPA_j = \frac{l_j \cdot CV_j}{1 + l_j \cdot CV_j} \quad (18)$$

Using a Taylor second-order development, (18) may be written as:

$$UPA_j \approx 1 - \frac{1}{l_j CV_j} + \frac{1}{2 (l_j CV_j)^2} \quad (19)$$

It is possible to verify the theoretical expression (19) by regressing UPA on $1/CV$ and $(1/CV)^2$. The results are given in Table 1 for all three years. It can be seen that, depending on the year considered, between 75 and 90% of the variance of UPA is explained by the variance of CV . Thus the UPA distribution is related primarily to that of CV , which is a measure of the global spatial variability of a given rainfield, and also to that of the variable l , which appears to be related to the overall strength of a rainfall event. This explains (at least partly), two findings of our comparison between classification methods: one, not presented here, is that differences were found between the UPA classification and a classification based on the CV only (Amani, 1995), the UPA classification performing better; the other (presented below) is that the combination between methods I and III did not improve the performances obtained with method III alone (the information provided by method I is essentially related to the strength of a rainfall event).

5- STRATIFICATION OF SAHELIAN EVENT RAINFIELDS.

In order to evaluate the impact of classification on the spatial estimation of rainfall by the ATM model, rainfall data collected in connection with the EPSAT-NIGER experiment during the years 1990, 1991 and 1992 are used. The four classification schemes presented in section 3, methods I, II, III and IV, are applied to the stratification of each of the three event rainfall samples of 1990 1991 and 1992, samples of 1990 and 1991 combined, and the three years combined. The histogram of the UPA values (133 events recorded during the EPSAT-NIGER experiment) is presented in Figure 8. This figure clearly leads to a natural classification of rainfall fields into three distinct groups: group 1 with UPA values less than 0.66; group 2 with UPA values between 0.66 and 0.81 and group 3 with UPA values greater than 0.81. This systematic classification is called DUPA (Distribution of the UPA parameter), or method III. Thus, two methods are systematic (methods II and III) and the other two (methods I and IV) are based on a clustering analysis. The commercial software Statistical Analysis System (SAS, 1982), version 5.18, was used to perform the clustering. A review of clustering techniques may be found in *MacQueen, 1967 and Anderberg, 1973.*

Table 2 presents the size of the significant groups for each method. Method I leads to two groups having similar size. Group 1 is made of the stronger events, while group 2 is made of the less intense events. For method II, spatial crossing, the group of the event is given by the order of its spatial profile function (rainy area against threshold) when the crossing technique is applied. The majority of the events belong to group 1b. The other events, belonging to the other orders (2a or 2b), are characterized by a large spatial extension. The three groups given by method III are characterized by different degrees of intermittence. The events of group 1 are large spatial events, while those of group 3 are highly intermittent.

A major point to consider when estimating areal rainfall is the continuity and stationarity of the field. If the field may be assumed to be continuous and stationary in space, several statistical methods, most notably those of the BLUE (Best Linear Unbiased Estimator) family, are available. On the other hand, when the intermittence becomes a significant aspect of the field under consideration, it has to be taken into account, as proposed by *Braud et al. (1993)* for instance. In such cases, two parameters play a key role in the areal estimation process: the mean Areal Rainfall (AR) over the rainy areas, which is a characteristic of the event

magnitude, and the Percentage of Gauges Affected by rainfall (PGA), which characterizes the spatial extension of the event. Methods II and III, based on the spatial behavior of the events, tends to classify the rainfall events according to their PGA and their AR, whereas method I does so principally according to their AR (*Amani, 1995*).

Three important statistics of the obtained groups can be used to analyze the potential of the classification methods to improve rainfall estimation by the ATM model. They are: the mean average areal rainfall, the mean standard deviation and the mean coefficient of variation. These statistics are presented for the three groups obtained by the method III (CUPAA) in Table 3 as an example. It appears that both the non-conditional and conditional statistics are significantly different from one group to another. The probability of non-zero rainfall may be estimated as the ratio between the non-conditional average (including zero rainfall) and the conditional non-zero average, that is 97% for group 1, 85% for group 2 and 63% for group 3. Since both CV and the probability of non-zero rainfall are very different from one group to the other, this necessarily has an effect on the ATM model where, usually, a unique conditional distribution function of rain rate is considered.

Also, to point out the possible positive impact of these classification methods on improving the ATM model, the variance of the conditional mean rainfall depth for each group is computed and compared with the variance of the conditional mean rainfall depth (conditional to the threshold considered) of the sample for each year. A reduction in the variance of the conditional average rainfall depth is observed for the majority of groups. Figure 9a illustrates the reduction in the variance of the conditional raindepth against the threshold for method III for the 1990 sample. The conditional mean raindepth is also given as a function of the threshold in Figure 9b. Table 3 and Figure 9 display significant differences between the statistics of each group of Method III. Similar results were obtained with the other methods, which leads to expect an improvement in the ATM model by considering the classification procedure.

6- COMPARISON OF PERFORMANCE OF THE ATM EVENT RAINDEPTH ESTIMATES WITH AND WITHOUT STRATIFICATION.

The four classification methods are considered now in order to evaluate their impact on the ATM estimates of areal rainfall depth. The ATM model is therefore applied before and after classification and the resulting errors are compared. These errors are evaluated by considering the optimal ATM

model for each group.

6.1 Review of the Area Threshold Method (ATM)

The ATM model, initially developed in the meteorological domain for estimating average rainfall based on satellite images and radar, is applied here to the EPSAT-NIGER data to estimate average event rainfall. The ATM model is based on the existence of a strong linear correlation between the average rainfall over an area and the fractional area where the rainfall amounts exceed a given threshold. This relation was first observed empirically by *Doneaud et al. (1981; 1984)*. Thus, for a threshold rainfall C , the mean intensity $m(t)$ at time t on the surface under study is given by:

$$m(t) = S(c) F(t, c) + b(c) \quad (20)$$

where $F(t, c)$ is the fractional area where the rainfall recorded at time t exceeds threshold C . $S(c)$ and $b(c)$ are coefficients obtained through linear regression between the average rainfall amounts for the various fields and the corresponding fractional rainfall area. These coefficients are a function of the threshold, while the average rainfall

is independent of it. The threshold method is characterized by the fact that, as the threshold value increases, so does the regression coefficient R^2 between $m(t)$ and $F(t,c)$ until a maximum value for a certain threshold (optimum threshold) is reached, and then it decreases.

The theoretical framework of the threshold method has been recently revisited by several authors (*Kedem et al., 1990; Atlas et al., 1990; Rosenfeld et al., 1990*). *Kedem and Pavlopoulos (1991)* have proposed a statistical method for determining the optimum threshold, and *Braud et al. (1993)* have shown that, under certain hypothesis, the theoretical correlation coefficient between $m(t)$ and $F(t,c)$ is a function of the spatial structure of the rainfall depth amounts.

The evaluation of the impact of classification is based on the calculation of a criterion measuring the error between the average rainfall estimated by the model over the whole area under study and the average rainfall recorded by the raingauge network. The criterion selected here is the Root Mean Square Error (*RMSE*). In addition, the improvement brought about by the model is tested in calibration and validation modes. Given a classification method, the *RMSE* for each group and a global *RMSE* are computed.

6.2 Relation between the rainfall surface and the average accumulated rainfall for the event

The method is applied here to estimate average rainfall accumulations during an event over the area under study

(13,200 km²). First, the method is applied without classification of rainfall fields. It is then applied separately for each of the significant groups obtained using the four classification methods. The following nine basic threshold values are considered: 0.5, 2, 5, 7.5, 10, 12.5, 15, 17.5 and 20 mm. For some groups, to reach the optimum threshold, additional threshold values are considered. For each of these values and for each rainfall field, the fractional rainfall area (an area where the rainfall amounts exceed the fixed threshold value) is calculated using the indicator method. Let C be the value of the threshold and $R(i)$ the total amount of rainfall recorded at station i .

The indicator variable $I(i)$ is defined as:

$$\begin{aligned} I(i) &= 1 \text{ if } R(i) \geq C & (21) \\ &= 0 \text{ otherwise} \end{aligned}$$

The fractional rainfall area for rainfall event k is estimated by:

$$F(c, k) = \frac{1}{N} \sum_{i=1}^{i=n} I(i) \quad (22)$$

The average rainfall $m(k)$ of field k over the surface under study is calculated by the arithmetic mean, taking into account null values. For each set of rainfall events, the estimation model is obtained by carrying out, for a given value of threshold C , a linear regression between the average rainfall values $m(k)$ and the values of the fractional rainfall area $F(c, k)$. Basic models are determined first, without classification, by calculating the correlations for each year separately (1990 - 37 values; 1991 - 47 values and 1992 - 49 values). More substantial samples are formed subsequently by adding the years 1990 and 1991 (84 values), and then by including all three years (133 values). The evolution of the coefficient of determination R^2 as a function of the threshold for each sample is given in Figure 10. From the values of R^2 , we see that the optimum threshold for each year is in the neighborhood of 17.5 mm.

6.3 The ATM model and classification

The area threshold method is applied to each of the groups obtained by the four classification methods, for each of the three years, the years 1990 and 1991 combined and the

three years combined. The most important point in the application of the ATM model for each group, is that the optimum threshold and the coefficient of determination R^2 are different from one group to another for all the samples and classification methods considered. As an example, Figure 11 illustrates the variation in the coefficient of determination (regression equation characterizing the ATM model) versus the threshold for the three classification methods (methods I, II and III) when all three years are combined.

This result can be expected from the analysis given in section 4 where it was pointed out that the statistics (mean, standard deviation and coefficient of variation) of the various groups are different. The parameters of the optimal ATM model for methods I, II and III are given in Table 4. Figure 12a illustrates the threshold model for the (1990-1991-1992) sample before classification at the optimal threshold (17.5 mm). For the method III, Figures 12b, 12c and 12d present also the optimal threshold models for the groups 1, 2 and 3 respectively.

Before stratification, the RMSE value of the optimal ATM model for the (1990-1991-1992) sample was 2.24. After classification, by considering the optimal model for each group, the RMSE values are 1.93, 1.68, 1.54 for methods I, II

and III respectively. In terms of percentages, this corresponds to RMSE value reductions of 13.8%, 25% and 31.3% respectively for the three methods. Thus, classification results in a significant improvement of the ATM estimates in the calibration phase, particularly for method III. However, the effect of the classification has to be evaluated in a validation mode, especially since the model before classification has a smaller number of parameters (2) than the models after classification ($2 \times$ number of groups).

Two cases are considered for validation purposes. In the first case, the mean areal rainfall of the 1992 events are reconstituted from the optimal ATM models calibrated by combining the 1990 and 1991 events. A fourth classification scheme combining the methods I and III is considered here in order to see if improvement can be observed by stratification based on parameters related to the strength and the intermittence of the event rainfall. The parameters of the optimal ATM model are given in Table 5 for the four methods. Based on the reduction of the RMSE criterion after classification, the four methods can be ranked as follows: method III (23.6%), method II (20.8%), method I (13.5%), and method IV (7.6%).

To check for possible bias linked to 1992 being a non representative year, a second validation procedure was designed. For a given classification method, events are considered in a chronological order inside each group produced by the method. Then, one out of two events is put into a validation sample. This leads to obtain a calibration and a validation samples of identical size (plus or minus one unit) for each group. The overall calibration and validation samples are the addition of the the calibration and validation samples of each group.

The results of this second validation procedure are similar to those obtained with the first one. That is, method III produces the largest improvement of the ATM model performances (Table 6). A reduction of the RMSE criterion after classification of 31.3% is obtained when sample I is estimated from sample II, and a reduction of 36.5% is obtained when sample II is estimated from sample I, that is an average reduction of 34%. Group 1, composed of major events, contributes greatly to the performance of the method III. For this group, a reduction in RMSE validation values of 52% is obtained.

The good results in validation case I reveal the robustness of the classification combined with the ATM model,

since the event statistics of the 1992 season were somewhat different from those of 1990 and 1991 (see statistics in Table 3). Classification method IV does not perform as well as method III. Thus, there is no advantage in combining methods I and III. The impact of the classification is very small for medium events and is sometimes negative for small or very intermittent rainfall events.

Analysis of the optimal ATM models and the statistics of the various samples considered (each year, the 1990-1991 combination, and the three years together, and their corresponding groups for method III) leads to the following observations:

i) For each sample before classification, the optimal threshold is around 17.5 mm. This means that the three years of EPSAT-NIGER data can be considered homogeneous. Thus, the threshold method can be used in the Sahel as a climatological approach for areal rainfall estimation over large domains (over 10.000 km² in this case). Of course, this is by no way the best model when direct measurement (raingauges) are available but may constitute a reasonable trade off when only indirect measurements such as satellite infrared temperatures are available.

ii) On the other hand, it is observed that for each sample considered, the three groups obtained from the DUPA classification (method III) have different optimal thresholds, and that these optimal thresholds are identical for all the samples. Their values are 30, 20 and 15 mm respectively for groups 1, 2 and 3.

iii) Further confirmation of the stability and robustness of the DUPA (method III) classification is given by the contribution of the different groups in terms of total rainfall accumulation for each sample. Analysing Table 3 (third row for each sample), reveals that the events in group 1 contribute greatly to the total rainfall relative to their number. It is also seen from this table that the ratio of the contribution of the events in each group relative to the total rainfall (as a percentage) to their number (as a percentage) is relatively constant from sample to sample. These ratios are approximately equal to 2, 1 and 0.5 for groups 1, 2 and 3 respectively.

iv) Even though some further work is needed on that point, it is of particular interest to link the three groups of method III with the types of precipitation systems listed in section 2, namely isolated convective systems, organized moving convective systems and fully developed squall lines. As far

as isolated convective systems could be identified to group 3 and fully developed squall lines to group 1, it can be concluded then 1992 was markedly different from 1990 and 1991 in that the isolated convective systems produced a larger proportion of the seasonal rainfall than usual. More generally, studying the contribution of each group of the UPA stratification to the seasonal rainfall could provide meaningful insights for hydrological and agronomic applications.

7 CONCLUSION

The impact of classification on the average estimation of rainfall over an area has been evaluated on the ATM model. Four methods of classification have been proposed and tested. These methods are: method I, based on the maximum rainfall accumulation over an array of time steps; method II, based on the crossing analysis of the rainy area function against a threshold; and method III, based on a new parameter, called UPA (Under profile Area). The fourth method is a combination of methods I and III. It should be noted that the UPA parameter is directly linked to the distribution of the area where it rains above a given threshold, and that it is a scaled parameter.

A consequence of the application of these methods is that the non conditional and conditional statistics (means of average rainfall, standard deviation and coefficient of variation) for each obtained group differ from one group to another. Also, a reduction in the variance of the average conditional mean rainfall depth is observed for the majority of groups.

These methods were applied to the estimation of mean areal rainfall on the EPSAT-NIGER study area by the ATM model. For every method, the optimal threshold of the ATM model is different from one group to another, which points up the importance of classification in improving the performances of the ATM model. At the calibration or validation level, classification methods II and III are the best in terms of RMSE reduction. The impact of classification is best for groups composed of large events (as in the case of group 1 for method III). For this group, the reduction in validation error are 52 % and 48.8% respectively for the two validation cases. Natural classification based on the three groups from the UPA distribution analysis (method III) is the best for the Sahelian rainfall events. A combination of method I, characterizing the intensity of the event, and method III characterizing the spatial organization of the event, does not challenge method III. This means that the

spatial organization of the rainfall event is more important than rainfall intensity for estimation with a global model such as the ATM model.

The excellent results of classification applied to the ATM model, combined with the fact that the optimal threshold is different from one group to another, have a major consequence. The hypothesis of one probability distribution function of the conditional rain rate for the sample assumed by the ATM model has to be reconsidered. More specifically, it appears that separating between different types of convective events may prove to be as rewarding as separating between stratiform and convective rainfall.

Finally, it was pointed out that the classification method based on the UPA parameter is climatologically stable and robust. Thus, this method can be used to stratify the various events observed during a rainy season in the Sahel based only on raingauge data, and can contribute to the improvement of our understanding of the interannual rainfall fluctuation in this region.

Acknowledgments. The suggestions and comments of three anonymous reviewers and also from Dr. Zawadzki from McGill university help us to improve the presentation and the

writing of this manuscript. They are gratefully acknowledged. Special thanks are addressed to the scholarship program of the Gouvernement of Canada (Bourse de la Francophonie) for their financial support of the graduate studies of the first author. The implementation of the EPSAT-NIGER experiment involved the TOA and DEC departments of ORSTOM, as well as the Direction de la Météorologie du Niger and the French Ministry of Cooperation. Their support is gratefully acknowledged. Finally, this research was partially supported by the Natural Sciences and Engineering Research Council of Canada on individual research grant (OGP0008415).

References

- Amani, A., Estimation des pluies au Sahel: quelques problèmes de modélisation analysés à partir des données de l'expérience EPSAT-NIGER. *Ph.D thesis*, Ecole Polytechnique de Montréal, 289p, 1995.
- Amani, A., J. Rousselle, T. Lebel and J.D. Taupin, Hyetogram analysis of Sahelian mesoscale convective systems, in: Sevruc B. and M. Lapin (Editors), *Precipitation Variability and Climate Change: Proceedings of the International Symposium on Precipitation and Evaporation*, Bratislava, Slovakia, September 1993, Vol. 2, pp 23-28. Printed by the Slovak Hydrometeorological Institute, 1993.
- Anderberg, M.R., *Cluster Analysis for Applications*, Academic Press, New York, 1973.
- Atlas, D., D.Rosenfeld and D.A.Short, The estimation of convective rainfall by area integrals: 1. The theoretical and empirical basis, *J. Geophys. Res.*, 95(D3): 2153-2160, 1990.
- Benichou P., M. Lavielle and P. Veyre, A study of the systematic behaviour of the French hemispheric NWP model as determined by a non-hierarchical clustering analysis. *OMM Workshop on Systematic Errors in Models of the Atmosphere*. Toronto, Canada, 19-23 Sept. 1988.

Braud I., J.D. Creutin and C. Barancourt, The relation between the mean areal rainfall and the fractional area where it rains above a given threshold, *J. Appl. Meteor.*, 32, 193-202, 1993.

Cunnane, C., Unbiased plotting positions- A review, *J. Hydrol.*, 37(3/4), 205-222, 1978.

Desbois M., T. Kayiranga, B. Gnamien, S. Guesous and 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. *J. Climatol.*, 1(9), 867-904, 1988.

Doneaud, A. A., P. L. Smith, A.S. Dennis and S. Sengupta, A simple method for estimating convective rain volume over area. *Water Resour. Res.*, 17(6), 1676-1682, 1981.

Doneaud, A. A., S.I. Niscov, D.L. Priegnitz and P. L. Smith, The area- time integral as an indicator for estimating convective rain volume. *J. Appl. Meteor.*, 23, 555-561, 1984.

Houze, R.A., Jr., Structure of atmospheric precipitation: A global survey, *Radio Science*, 16, 671-689, 1981.

Kedem, B., D.A. Chiu and Z. Karni, An analysis of the threshold method for measuring area-average rainfall, *J. Appl. Meteor.*, 29, 3-20, 1990.

Kedem, B. and H. Pavlopoulos, On the Threshold Method for Rainfall Estimation: Choosing the Optimal Threshold Level. *J. of American Statistical Association*, 86(415), 626-633, 1991.

Kottegoda, N.T and A.H.M. Kassim, Classification of storm profiles using crossing properties, *J. Hydrol.*, 127(1/4), 37-53, 1991.

LeBarbé, L and T. Lebel, Rainfall climatology of the HAPEX Sahel during the years 1950-1990. In print, *J.Hydrol.*, 1995.

Lebel, T and L. LeBarbé, Rainfall monitoring during HAPEX-Sahel: 2. Point and areal estimation at the event and seasonal scales. In print, *J.Hydrol.*, 1995.

Lebel, T., J.D. Taupin and N. d'Amato, Rainfall monitoring during HAPEX-Sahel: 1. General rainfall conditions and climatology. In print, *J.Hydrol.*, 1995a

Lebel, T., J.D. Taupin and M. Gréard, Rainfall monitoring: the EPSAT-NIGER set up and its use for HAPEX-Sahel, in Lebel T. (Editor), *Hydrologie et Météorologie de Mésos-échelle dans HAPEX-SAHEL : dispositif de mesures au sol et premiers résultats*, Editions ORSTOM, Paris, France, 1995b.

Lebel, T., H. Sauvageot, M. Hoepffner, M. Desbois, B. Guillot and P. Hubert, Rainfall estimation in the Sahel: The EPSAT-NIGER experiment, *Hydrological Sciences Journal*, 37(3), 201-215, 1992.

MacQueen, J., *Some methods for classification and analysis of multivariate observations*, Proceedings. 5th Berkeley Symp. on Math. Sta. and Prob., 1, 281-297, 1967.

Negri, A.J. and R.F. Adler, Infrared and visible satellite rain estimation. Part I: A grid cell approach, *J. Climate Appl. Meteor.*, 26, 1553-1564, 1987.

Rosenfeld, D., D. Atlas and D.A. Short, The estimation of Convective Rainfall by Area Integrals. 2- The Height-Area Rainfall Threshold (HART) method. *J. Geophys. Res.*, 95(D3), 2161-2176, 1990.

SAS (Statistical Analysis System), Statistics User's Guide, SAS Institute, North Carolina, 500p, 1982.

Short, D.A, D.B. Wolff, D.R. Rosenfeld and D. Atlas, A study of the threshold method utilizing raingage data. *J. Appl. Meteor.*, 32, 1379-1387, 1993.

Taupin, J.D, A. Amani and T. Lebel, Small-scale spatial variability of the annual rainfall in the Sahel, in: Bolle H.-J., Feddes, R.A. and Kalma, J. (Editors), *Exchange Processes at the Land Surface for a Range of Space and Time Scales*, IAHS Publ. N° 212, pp 593 - 602, 1993.

Woodley, W.L., C.G. Griffith, J.S. Griffin and S.C. Stromatt, The inference of GATE convective rainfall from SMS-1 imagery, *J. Appl. Meteor.*, 19, 388-408, 1980.

Figure captions

Fig. 1 The EPSAT-NIGER recording raingauge network, 1992.

Fig 2. Event rainfall versus proportion of zero-rainfall for the events of 1992.

Fig. 3 Illustration of the rainy area above a threshold (C_i , i^{th} threshold value and A_{ci} is the rainy area where it rains above C_i).

Fig. 4 Rainy area function of the threshold (C_m is the maximum rainfall depth and A_0 is the total area of the network).

Fig. 5 Classification of storm structure types by method II: a) type 1b and 1a; c) type 2b and 2a; c) type 3b and 3a. (After Kottegoda and Kassim, 1991).

Fig. 6 Illustration of the UPA parameter and spatial crossing approach.

Fig. 7 Histogram of the frequency factor values (three years combined).

Fig. 8 Histogram of the UPA values (the three years combined).

Fig. 9 Variance of the average conditional rainfall depth for the events of 1991 and the three groups of the DUPA method against the threshold considered. a) variance b) average conditional rain depth.

Fig. 10 Coefficient of determination of the ATM model for various thresholds for each year and combination of years.

Fig. 11 Coefficient of determination versus threshold for the ATM models after classification. a) method I, b) method II and c) method III.

Fig. 12 Optimal ATM models for the three years combined for the groups of method III. a) total sample, b) group 1, c) group 2 and d) group 3.

TABLE 1: Parameters of the linear regression of UPA against $1/CV$ and $(1/CV)^2$. ($UPA = a_0 + a_1/CV + a_2/(CV)^2$). l_1 and l_2 are the estimates of l computed respectively as $l = 1/a_1$ and $l = 1/(2a_2)^{1/2}$ (see expression 19).

	R^2	a_0	a_1	a_2	l_1 l_2	n
1990	0.753	1.10	-0.35	0.036	2.8 3.7	37
1991	0.873	1.02	-0.34	0.046	2.8 3.3	47
1992	0.813	1.01	-0.29	0.036	3.5 3.7	49

TABLE 2: Parameters and group size for each significant group obtained from each of the four methods for seasons 1990, 1991 and 1992.

Methods	Group number	Group size			Parameters
		1990	1991	1992	
Method I	1	15	15	24	Rainfall intensities
	2	19	15	22	
Method II	1b	32	39	35	Rainy area function
	2a,b	5	8	12	
Method III	1	6	10	5	UPA defined from rainy area function
	2	13	21	21	
	3	18	16	23	
Method IV	1	27	23	29	UPA and rainfall intensities
	2	10	20	12	

TABLE 3: Statistics and contributions of the different groups of method III to the total rainfall accumulation in the three seasons 1990, 1991 and 1992. For each group and for each year, we have, in the first row the non conditional statistics (average rainfall depth (mm), average standard of deviation (mm), and average coefficient of variation. In the second row, we have the corresponding conditional statistics, and the last row, we have the sample size, the percentage, represented by the group's events and the percentage of the group's events contribution for the total rainfall accumulation.

Sample	Group 1	Group 2	Group 3
Three years	22.8 12.8 0.60 23.6 12.5 0.56 21 15.8 34.7	11.1 10.2 1.15 13.1 10.4 0.84 55 41.4 43.7	5.3 8.4 1.79 8.4 9.6 1.16 57 42.8 21.6
1990	22.0 12.8 0.58 22.9 12.3 0.54 6 16.0 35.0	12.1 12.5 1.12 14.2 12.6 0.91 13 35.0 42.0	4.8 8.4 1.89 8.3 9.8 1.18 18 49.0 23.0
1991	21.4 12.7 0.65 21.9 12.5 0.61 10 21.3 41.0	10.9 10.2 1.01 12.3 10.1 0.83 21 44.6 43.9	4.9 8.3 1.83 8.2 9.6 1.17 16 34.0 15.1
1992	26.6 13.0 0.55 27.7 12.6 0.49 5 10.2 26.6	10.8 8.8 1.30 13.1 9.2 0.79 21 42.9 45.4	6.1 8.5 1.67 8.6 9.5 1.14 23 46.9 28.0

TABLE 4: Parameters of the optimal ATM model for the different groups obtained by Methods I, II and III for the combined rainfall events of 1990, 1991 and 1992 (coefficient of determination R^2 , sample size N , optimum threshold Opt , slope of the regression St and the ordinate at the origin B).

Method	R^2	sample N	opt (mm)	B (mm)	St (mm)
Method I	0.93	56	20	3.5	0.377
	0.94	50	12.5	1.5	0.253
Method II	0.88	26	30	10.4	0.436
	0.94	107	17.5	1.8	0.393
Method III	0.91	21	30	9.2	0.454
	0.96	55	20	2.6	0.436
	0.91	57	15	1.3	0.361
Total	0.94	133	17.5	2.1	0.361

TABLE 5: Calibration error of the 1992 mean areal rainfall events estimated by the ATM models before and after classification. (Optimal threshold Opt, coefficient of determination, slope of the regression St, the ordinate at the origin B, the error before classification RMSEb and the error after classification RMSEa). For method I, only the events belonging to the significant groups 1 and 2 are considered, this is why the RMSEb is 2.67 and not 2.50.

Method	Opt (mm)	R ²	St	B (mm)	RMSEb (mm)	RMSE a (mm)
Total	17.5	0.94	0.335	2.28	2.50	
Method III	30	0.94	0.430	10.43	2.50	1.91
	20	0.96	0.441	2.83	5.84	2.99
	15	0.88	0.338	1.21	2.28	2.20
Method II	30	0.95	0.432	12.94	2.50	1.98
	17.5	0.95	0.398	1.68	4.39	3.24
Method I	20	0.93	0.417	2.70	2.67	2.31
	15	0.94	0.304	1.90	3.48	2.92
Method IV	20	0.88	0.349	5.05	2.50	2.31
	15	0.92	0.308	1.70	4.00	3.61
					1.22	1.27

TABLE 6: Optimal ATM model parameters and the validation errors for validation case II for the method III. (Optimal threshold Opt, coefficient of determination, slope of the regression St, the ordinate at the origin B, the error before classification RMSEb and the error after classification RMSEa).

Sample	Opt (mm)	R ²	St	B (mm)	RMSEb (mm)	RMSE a (mm)
sample I	17.5	0.94	0.352	2.28	2.11	1.45
group 1	30	0.91	0.404	11.0	3.91	1.84
group 2	20	0.96	0.440	2.63	1.80	1.63
group 3	15	0.90	0.351	1.34	1.18	1.07
sample II	17.5	0.94	0.380	1.99	2.55	1.62
group 1	30	0.93	0.429	10.2	4.77	2.34
group 2	20	0.96	0.433	2.52	2.39	1.86
group 3	15	0.93	0.373	1.19	1.22	0.87

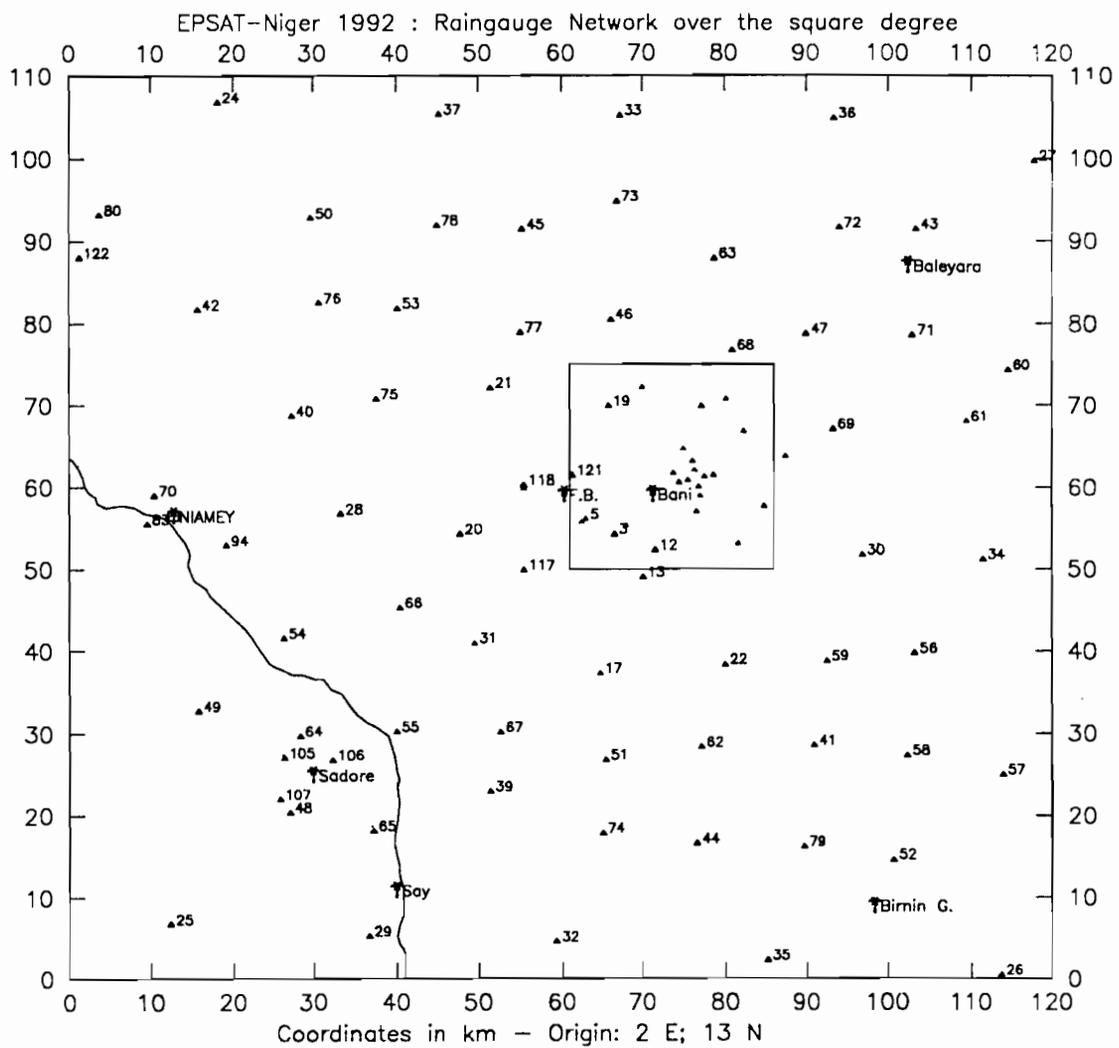


Figure 1

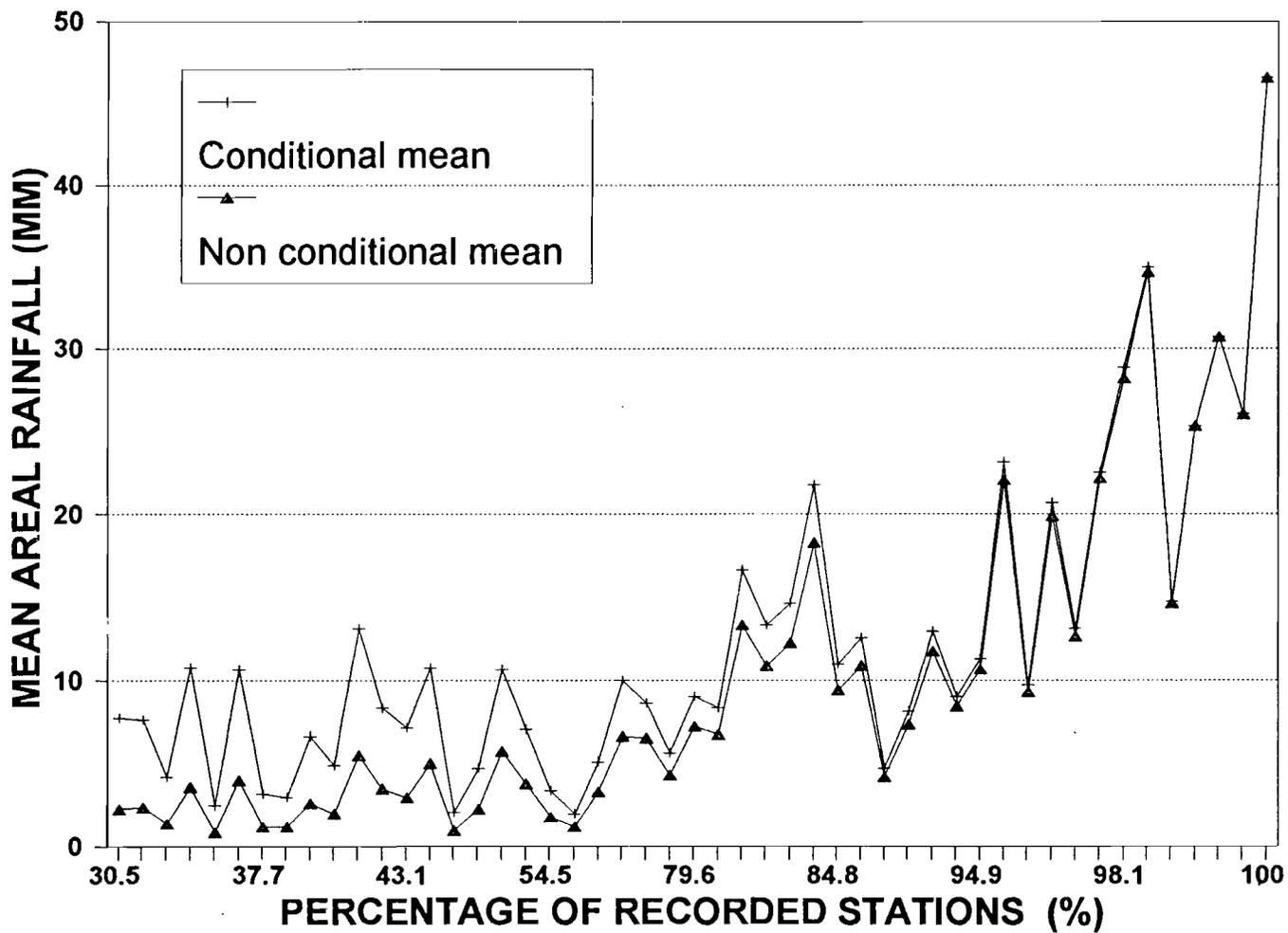


Figure 2

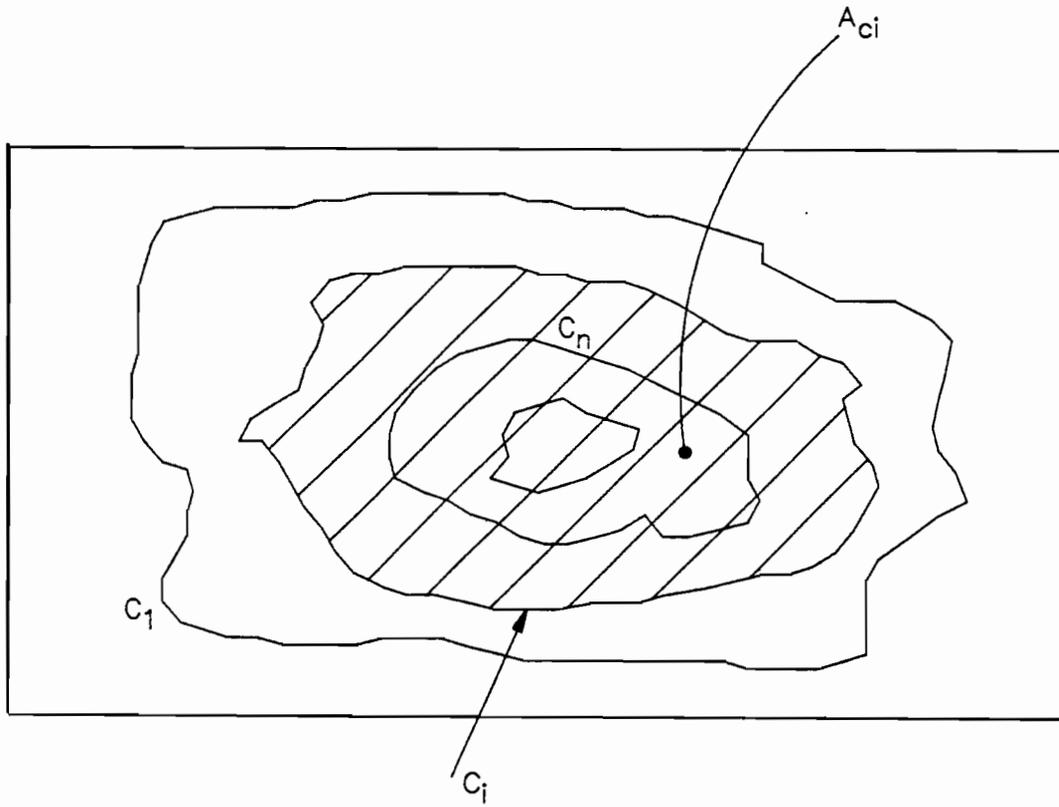


Figure 3

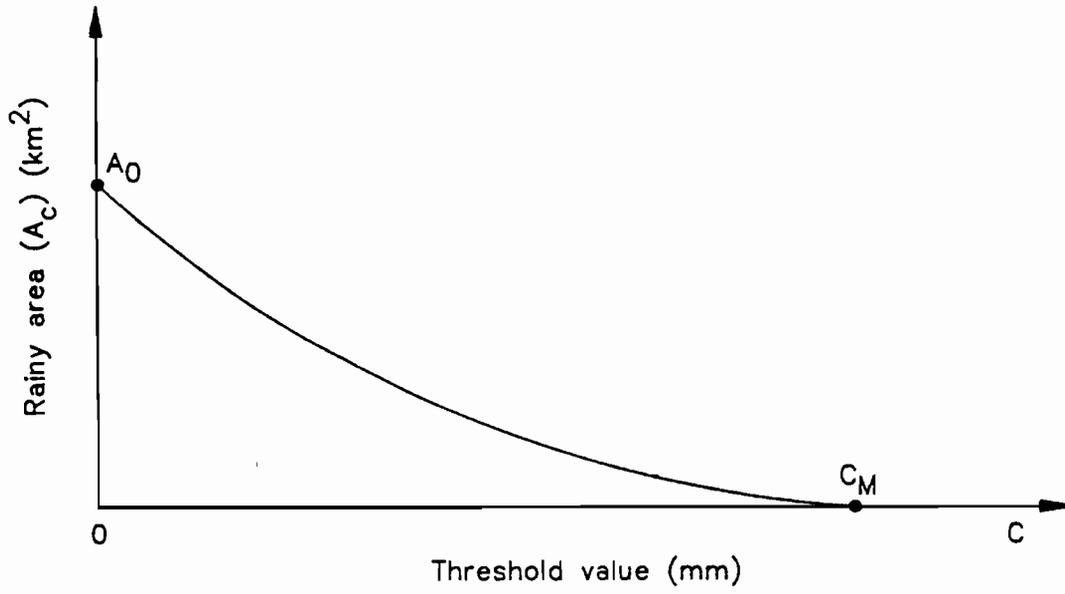


Figure 4

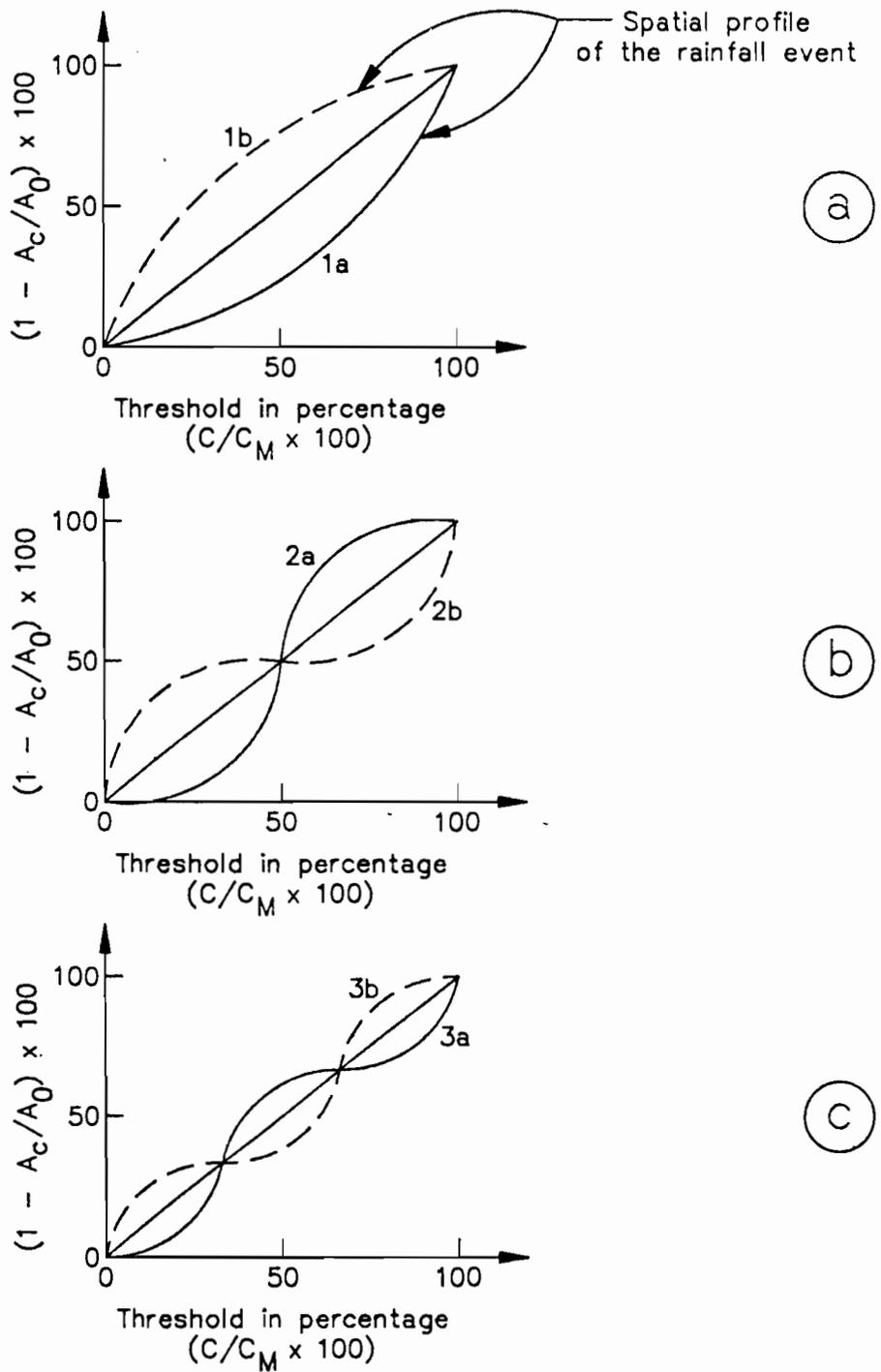


Figure 5

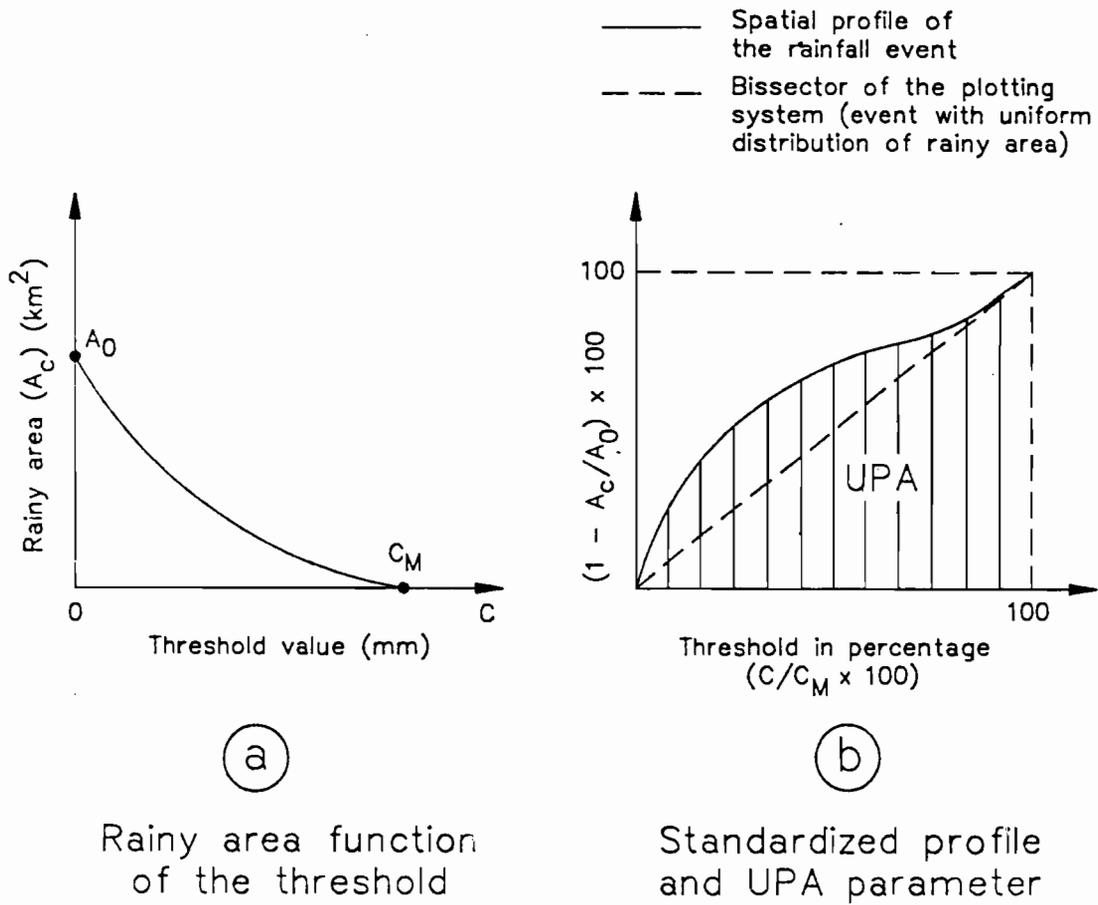


Figure 6

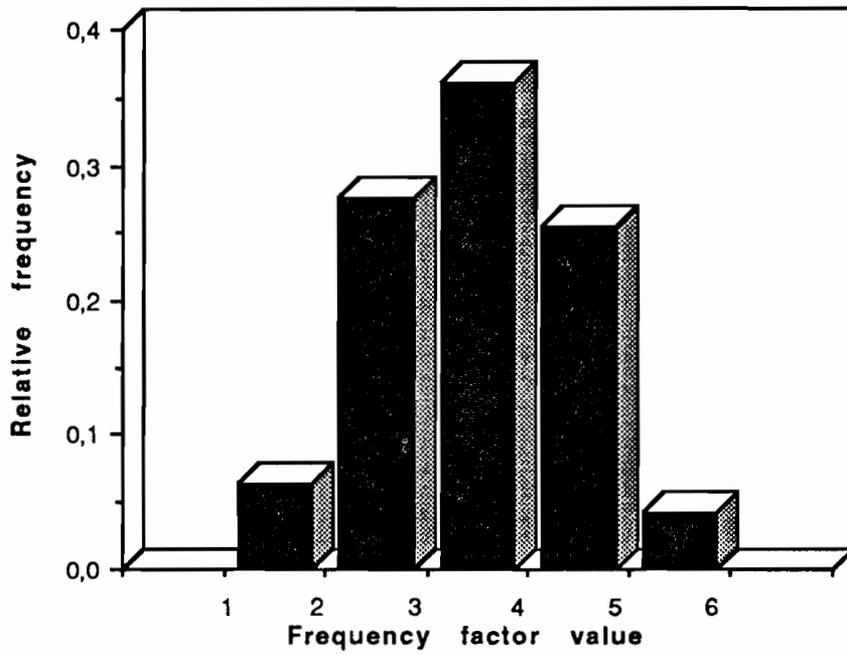


Figure 7

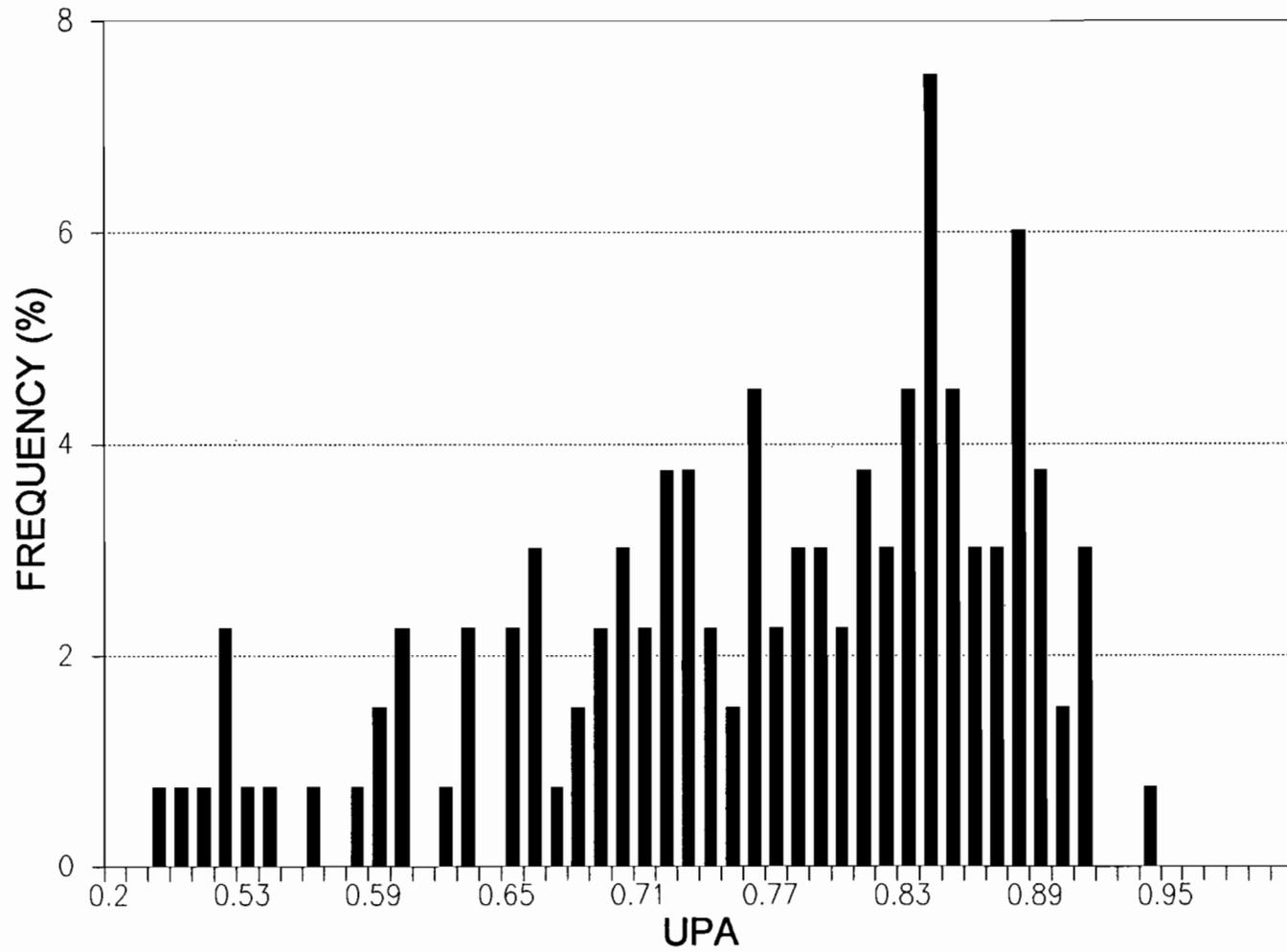


Figure 8

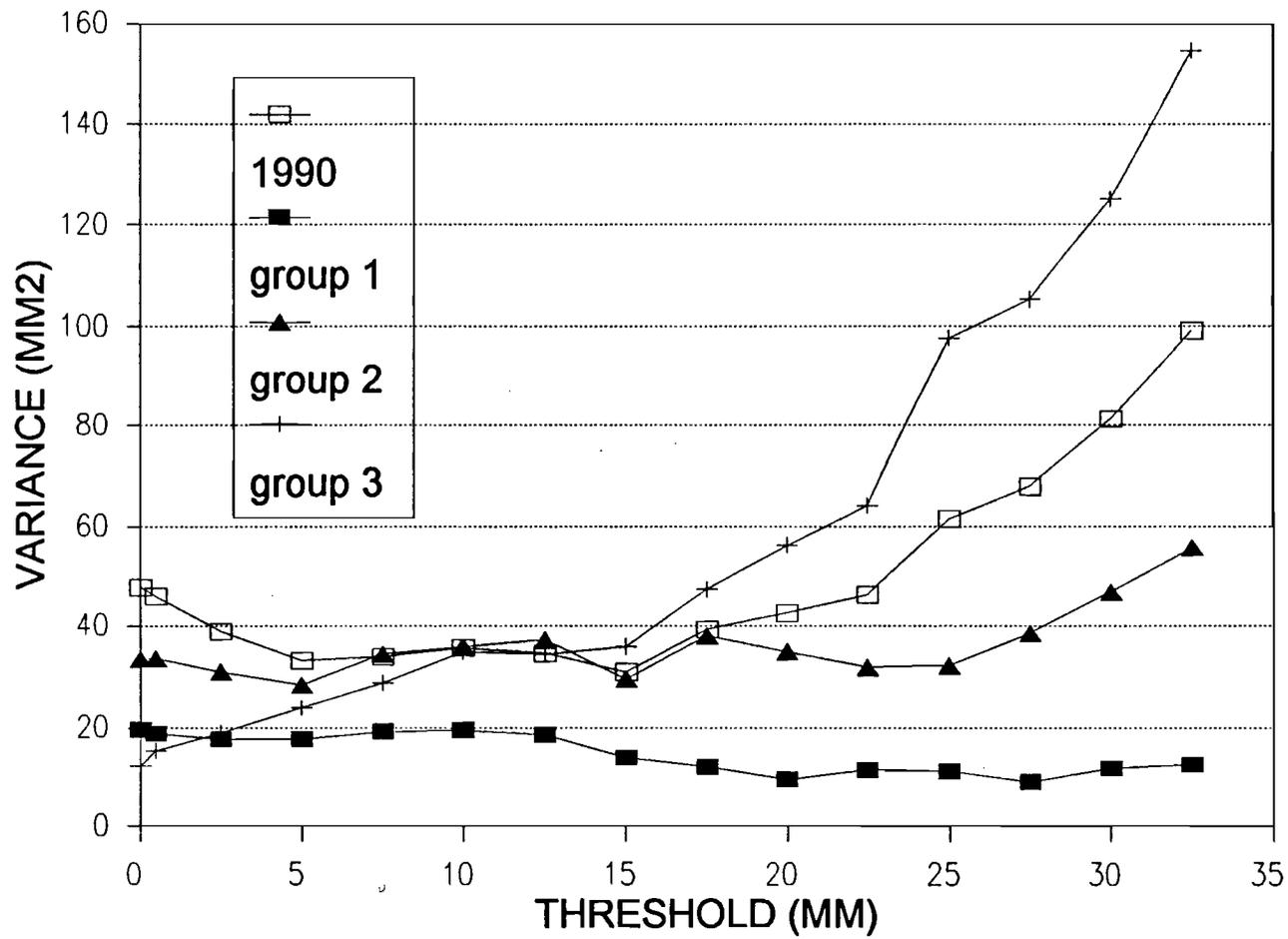


Figure 9a

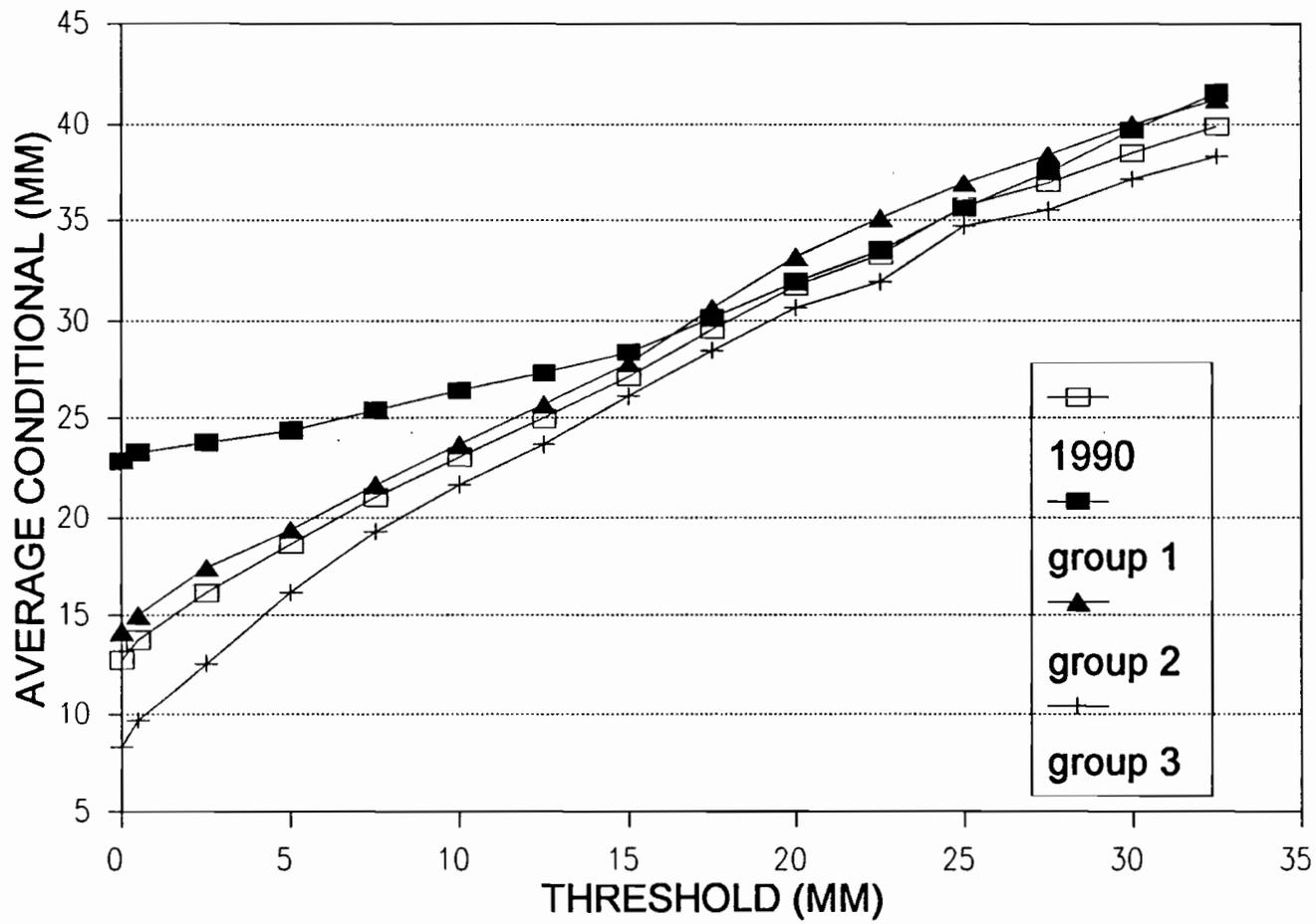


Figure 9b

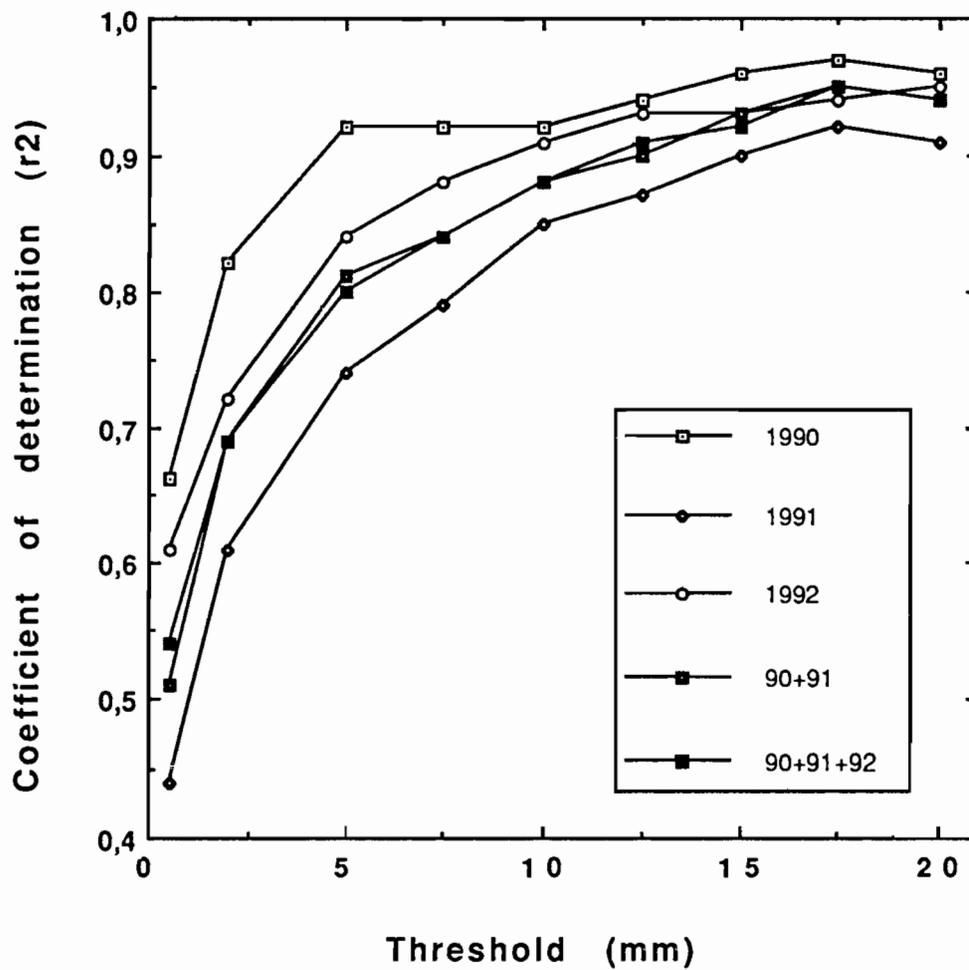


Figure 10

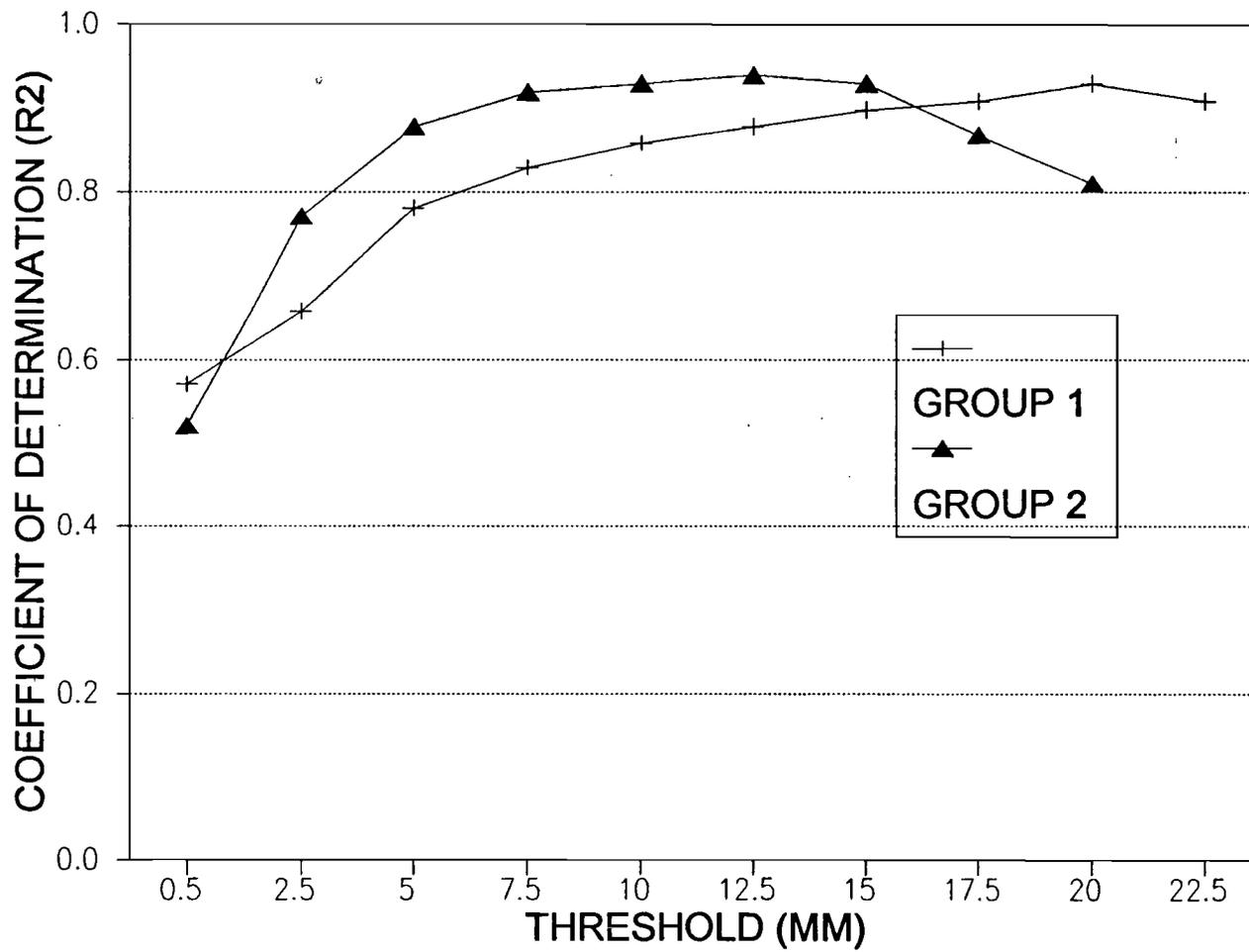


Figure 11a

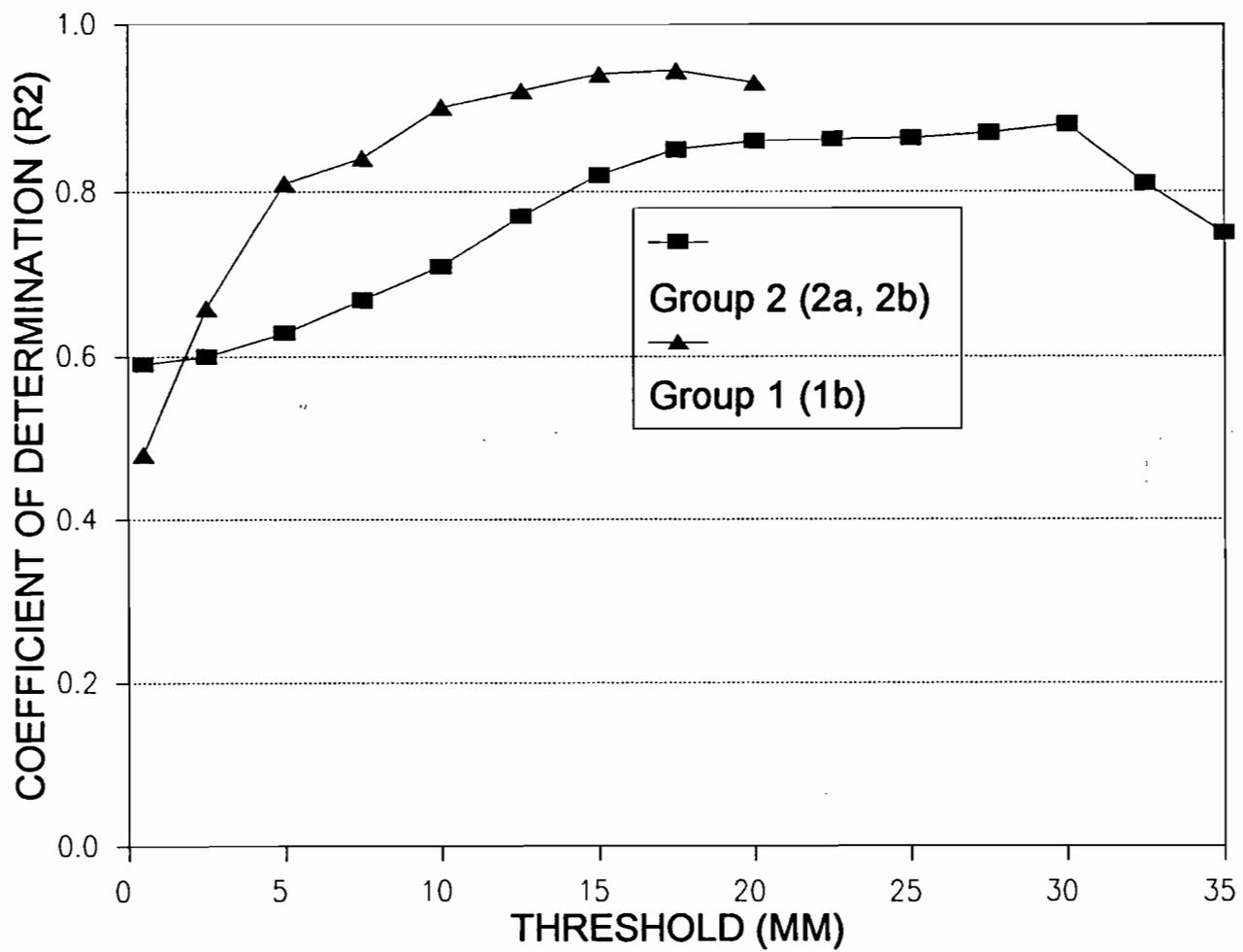


Figure 11b

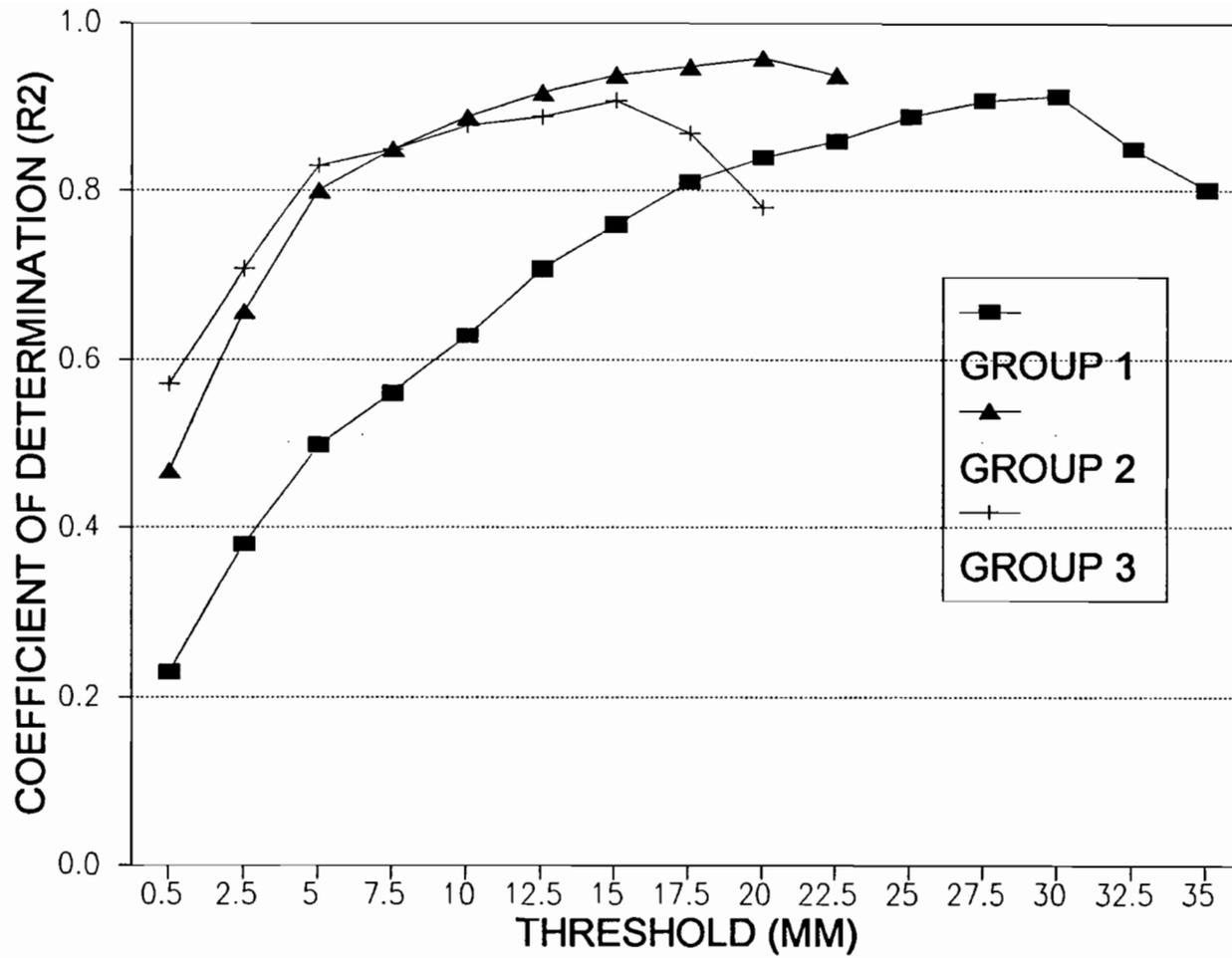


Figure 11c

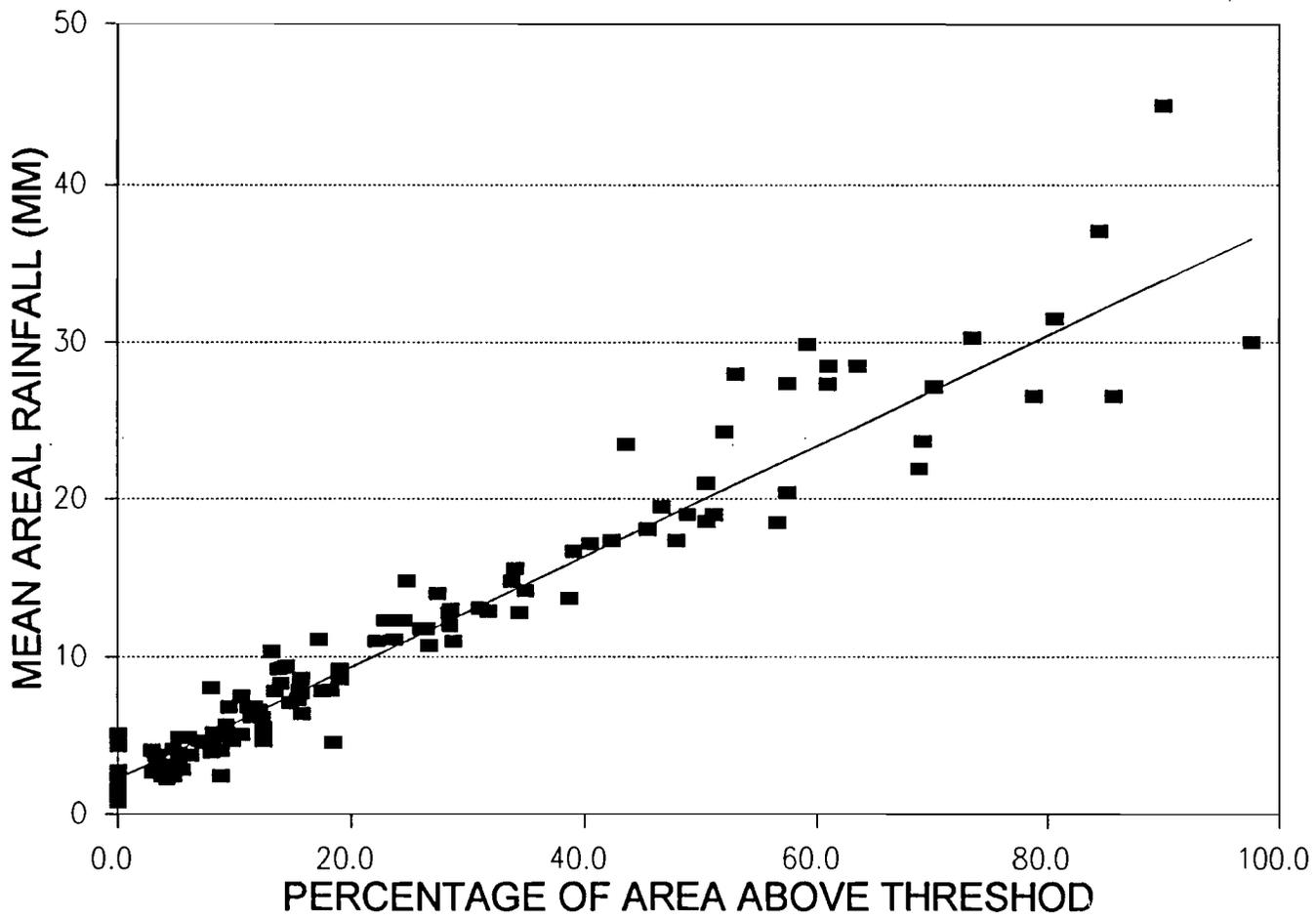


Figure 12a

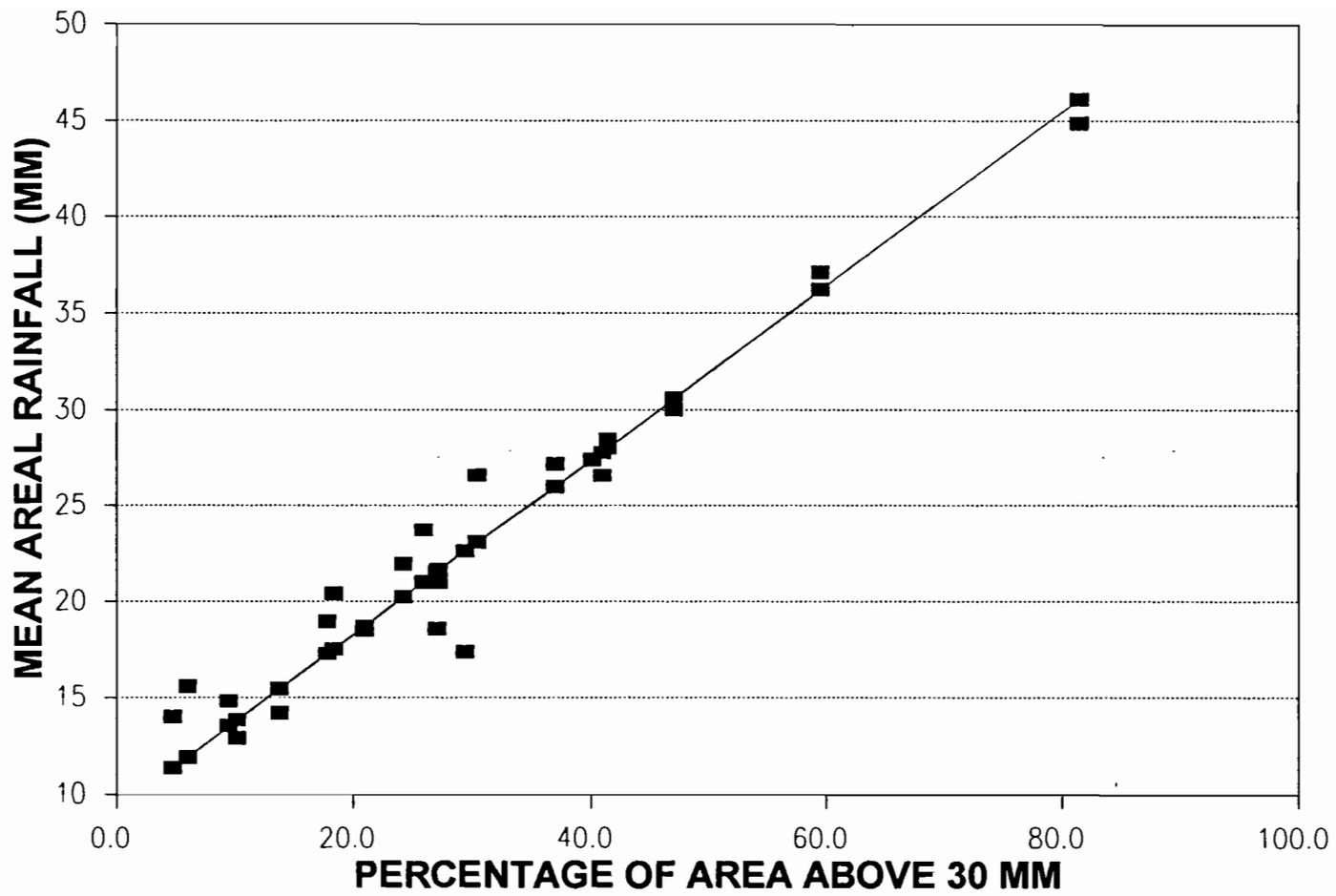


Figure 12b

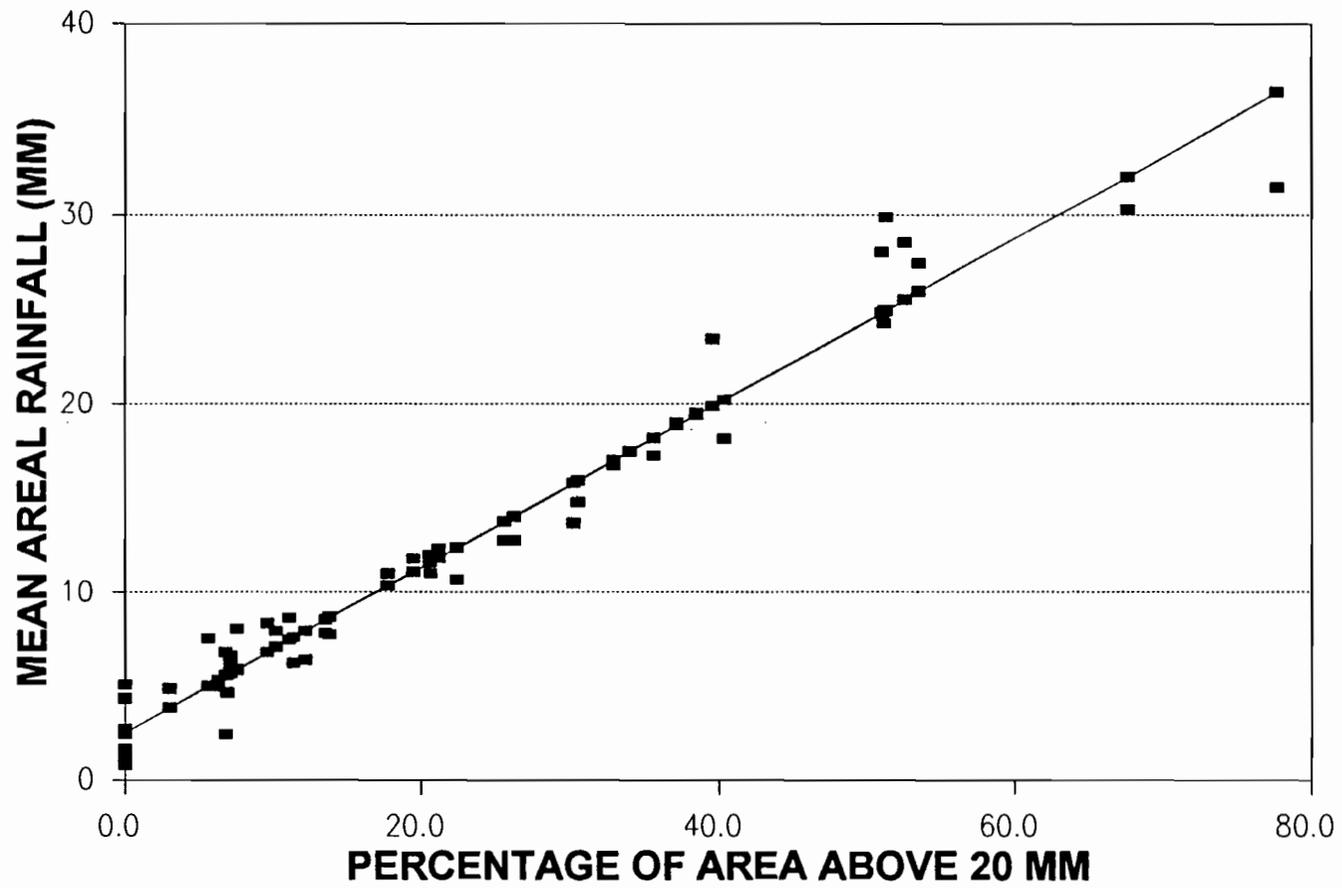


Figure 12c

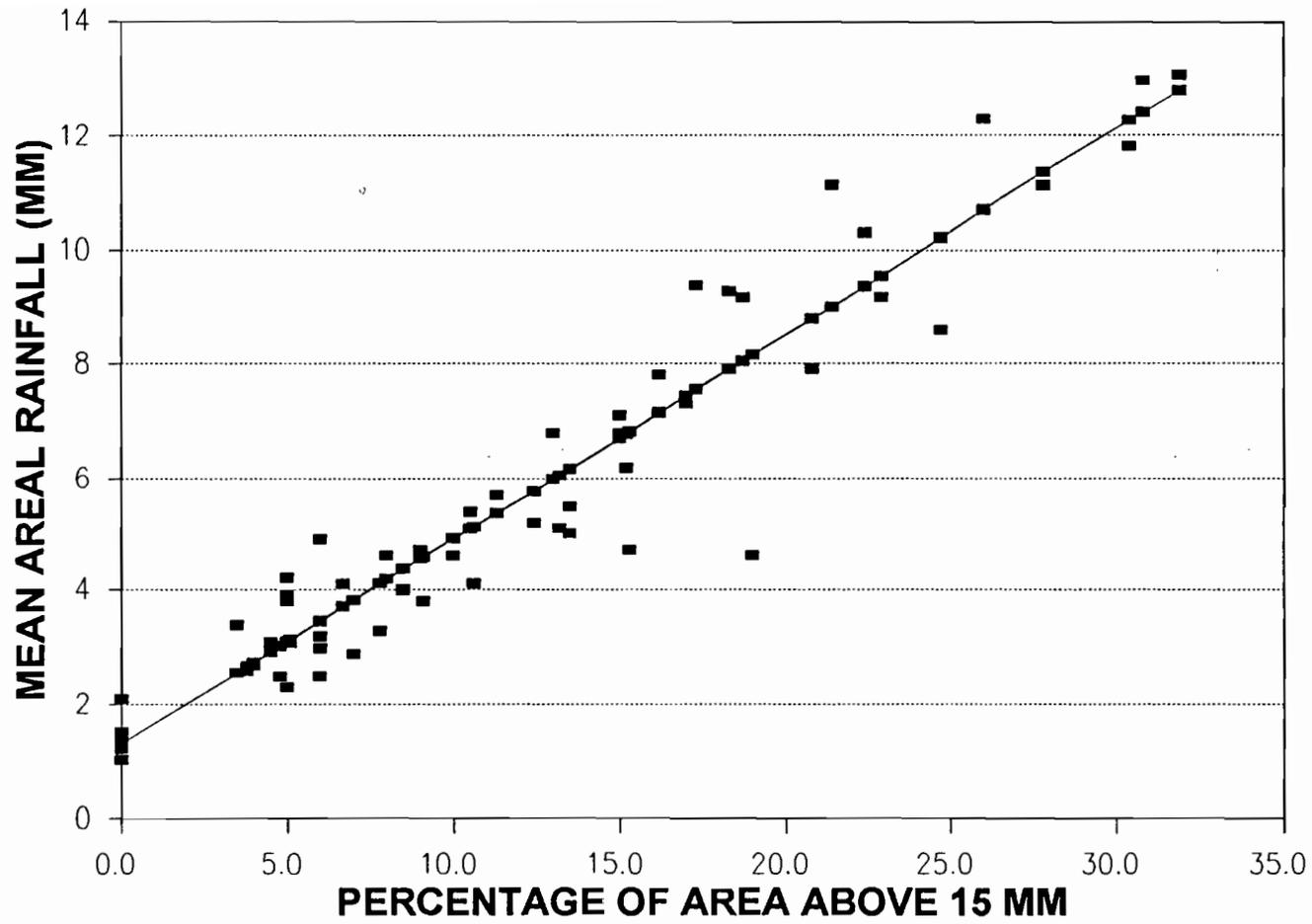


Figure 12d

ANNEXE A2:

Squall Line Rainfall Estimation at Small Time Steps in the Sahel by Dual Spatio-temporal Kriging, by A. Amani, J. Rouselle and F. Trochu. Soumis au Hydrological Sciences Journal.

**SQUALL LINE RAINFALL ESTIMATION AT SMALL TIME STEPS IN THE
SAHEL BY DUAL SPATIO-TEMPORAL KRIGING**

by

Abou AMANI and Jean ROUSSELLE

Ecole Polytechnique, Génie Civil,
C.P. 6079, Succursale Centre-ville,
Montréal, Québec H3C 3A7, Canada.

and

François TROCHU

Ecole Polytechnique, Génie Mécanique,
C.P. 6079, Succursale Centre-ville,
Montréal, Québec H3C 3A7, Canada.

Submitted to

Hydrological Sciences Journal

Abstract: At small time steps, it is well-known that there is a strong link between the spatial and the temporal behaviors of the rainfall process. In order to take into account this phenomenon, a spatio-temporal kriging technique was tested to improve, if possible, the estimation of Sahelian rainfall data (EPSAT-NIGER Data) at small time steps. The results, based on the EPSAT-NIGER basic network and the target area, lead to the fact that the improvement obtained by the spatio-temporal (3D) kriging is not significant compared to 2D kriging.

ESTIMATION AUX FAIBLES PAS DE TEMPS DES LIGNES DE GRAIN AU SAHEL PAR LE KRIGEAGE SPATIO-TEMPORAL DUAL

Résumé: Il est bien connu qu'il existe une très grande interaction entre les structures spatiale et temporelle de pluies aux petites échelles de temps. Cette dépendance est considérée par le biais du krigeage spatio-temporel (3D), afin d'améliorer si possible l'estimation des pluies sahéliennes (données EPSAT-NIGER) aux petites échelles de temps. Les résultats obtenus pour le réseau de base d'EPSAT-NIGER et pour la cible correspondante, montrent qu'il n'y a à toute fin pratique pas de gains significatifs obtenus avec le krigeage spatio-temporel au lieu du krigeage 2D pour l'estimation des pluies aux faibles pas de temps.

INTRODUCTION

As with many geophysical phenomena, rainfall exhibits a spatio-temporal behavior. The interaction between the space and time features of rainfall phenomena depends greatly on the time step duration of the observed data. In the majority of studies on rainfall estimation, temporal behavior is generally omitted and only a spatial estimation is performed for rainfall field data obtained for a given cumulative time step (Chua and Bras, 1982; Creutin and Obled, 1982; Lebel et al., 1987). For small cumulative time steps, the link between temporal and spatial rainfall features is so great that it might be interesting to model this interaction in order to improve, if possible, global rainfall estimations. In this study, a spatio-temporal kriging technique (3D kriging) is applied as a point rainfall interpolator at small time steps.

To our knowledge, this study is the first one in which the 3D spatio-temporal kriging has been applied to rainfall estimation. Spatio-temporal data sets of two rainfall events recorded during the EPSAT-NIGER experiment are used (Lebel et al., 1992). Two spatial scales are considered: the target area (400 km², 16 raingauges in 1990 and 18 raingauges in 1991) and the basic reference network area (12,000 km², 70 raingauges).

The paper is organized as follows: a review of the literature, formulation of the problem, presentation of the 3D kriging model, applications and conclusion.

REVIEW OF THE LITERATURE

Two main approaches are generally used in the literature to perform a spatio-temporal estimation. In the first approach, an indirect spatio-temporal data transfer is carried out via a global model which assumes indirectly a certain similarity between the rainfall fields. Climatological kriging and threshold method are being the principal examples of such indirect information transfer (Bastin et al., 1984; Lebel et al., 1987; Short et al., 1993). It was pointed out in Amani et al. (1995) that classification of spatio-temporal data sets can improve the accuracy of indirect transfer methods. In the second approach, also called the direct approach (Amani, 1995), the estimation at an ungauged site is considered to be a function of all available spatio-temporal data sets, and is given by equation (1) in the next section. The direct approach includes various methods: (i) cokriging techniques, (ii) multisite ARIMA models, and (iii) spatio-temporal kriging methods.

In cokriging, variograms and cross-variograms are needed (Delrieu et al., 1988; Seo et al., 1990). This leads to a practical numerical limitation related to the number of structural functions to be modeled. For a number of time steps K , cokriging requires $(K+1)K/2$ variograms and cross-variograms. Also the well-known ARIMA (AutoRegressive Integrated Moving Average) Box-Jenkins models present in practice a limitation; the time lag order of these models is generally less than or equal to 3.

Spatio-temporal kriging techniques allow us to overcome the main limitation related to cokriging and multisite ARIMA models. With spatio-temporal kriging, estimation at every ungauged site is possible at every time step. Both 3D kriging and multisite ARIMA models can be used for forecasting. Spatio-temporal kriging is a generalization of spatial kriging when time is considered as a third dimension (Bilonick, 1985; Rouhani and Hall, 1989). Depending on the structural function, two formulations of spatio-temporal kriging are possible (Amani, 1995): (i) 3D kriging in (x, y, t) with a structural function $K(h)$, where h is the euclidean distance in three dimensions; and (ii) 3D kriging with a structural function $K(h, \tau)$ depending on a spatial distance h and on a temporal distance τ . Meyers and Journel (1990) and Rouhani and Meyers (1990) have pointed out through some

examples that the decomposition of $K(h, \tau)$ into a sum or a product of temporal and spatial covariances, as proposed by some authors (Rodriguez-iturbe and Mejia, 1974; Rouhani and Hall, 1989), can generate noninvertible kriging systems.

FORMULATION OF THE PROBLEM

In this study, we are interested in evaluating the performances of spatio-temporal kriging (3D) and 2D kriging for rainfall estimation at a small cumulative time-step duration. Let $P(x_i, y_i, t_k)$ be the rainfall recorded at gauge i for the k^{th} time step for cumulative time-step duration θ . The raingauge network is composed of N gauges ($i=1, \dots, N$) and the total number of fields is K . For the process $P(x_i, y_i, t_k)$ with ($i=1, \dots, N$) and ($k=1, \dots, K$), we are interested in estimating the point rainfall $P(x_o, y_o, k_o)$ at an ungauged site located at (x_o, y_o) and at time k_o . For spatio-temporal estimation, $P(x_o, y_o, k_o)$ is estimated by

$$\hat{P}(x_o, y_o, k_o) = \sum_{k=1}^K \sum_{i=1}^N \lambda_{ik} P(x_i, y_i, k) \quad (1)$$

where λ_{ik} are the weighting coefficients associated with the rainfall $P(x_i, y_i, k)$ recorded at gauge i for the successive times $k=1, 2, \dots, K$. For 2D kriging, the estimator is given by

$$\hat{P}(x_o, y_o, k_o) = \sum_{i=1}^N \lambda_i P(x_i, y_i, k_o) \quad (2)$$

at each time step k_o .

The goal of this study is to evaluate the performance of the spatio-temporal kriging compared to one of the more classical spatial technique (2D kriging).

3D SPATIO-TEMPORAL KRIGING

3D kriging consists in expanding 2D kriging into time by considering time as a third dimension. Thus, only the calculation of the euclidean distance changes. The spatio-temporal data set $P(x_i, y_i, k)$ with $k=1, \dots, K$ and $i=1, \dots, N$ is assumed to be a realization of a three-dimensional random process $P(x, y, t)$, where (x, y, t) designates the location of the observed rainfall at time $t=(k-1)\theta$. The number of observed data M is the product of the number of gauges N and the number of fields K .

Kriging can be applied to the process $P(x, y, t)$, leading to the following estimation of rainfall at an ungauged site located at (x_o, y_o) at time $t_o= (k_o-1)\theta$

$$\hat{P}(x_o, y_o, t_o) = \sum_{i=1}^M \lambda_i P(x_i, y_i, t_i) \quad (3)$$

where λ_i are the weighting coefficients associated with the observed rainfall data $P(x_i, y_i, t_i)$ at a specific gauge location for a given time t_i . Suppose that $P(x, y, t)$ is characterized by a polynomial trend function $m(x, y, t)$:

$$m(x, y, t) = \sum_{p=1}^{l(q)} a_p f_p(x, y, t) \quad (4)$$

with

$$f_p(x, y, t) = x^{p_1} y^{p_2} t^{p_3} \quad p_1 + p_2 + p_3 \leq q \quad (5)$$

where q is the order of the polynomial trend, a_p are the parameters of the trend ($p=1, \dots, l(q)$) and $l(q)$ is the total number of parameters of the trend function. Applying the theory of the best linear unbiased estimator to the random process $P(x, y, t)$, the unbiasedness and minimum variance conditions give the estimation of $P(x_o, y_o, t_o)$ in the dual formulation of kriging by (Marche et al., 1990; Trochu, 1993):

$$\hat{P}(x_o, y_o, t_o) = \sum_{p=1}^{l(q)} a_p f_p(x_o, y_o, t_o) + \sum_{i=1}^M b_i K(h_{oi}) \quad (6)$$

where $K(h)$ is the generalized covariance function and a_p , for $p=1, \dots, l(q)$ and b_i , for $i=1, \dots, M$, are solution of the

following dual kriging system (Marche et al., 1990; Trochu, 1993):

$$\sum_{i=1}^M b_i K(h_{ij}) + \sum_{p=1}^{l(q)} a_p f_p(X_i) = P(x_j, y_j, t_j) \quad j=1, \dots, M$$

(7)

$$\sum_{i=1}^M b_i f_p(X_i) = 0 \quad p=1, \dots, l(q)$$

The advantage of the dual formulation of kriging is that the kriging system has to be solved once for a given sample, contrary to the primal formulation where a new kriging system has to be solved for each estimation.

When using 3D kriging to estimate rainfall phenomena, a particular problem arises, which is related to the physical meaning of the euclidean distance, whose expression is a combination of metric and time units. To overcome this problem, two approaches can be used: (i) normalization of the different coordinate axes; (ii) use of Taylor hypothesis to convert time into a metric distance by using the speed of the storm system (Zawadzki, 1973; Gupta and Waymire, 1987). For time-step durations less than 35 min, Zawadzki (1973) has observed that the behavior of rainfall phenomena is similar in space and time when time is converted into a metric distance by multiplication with the speed of the storm

system. Thus temporal units can be transformed into metric units in the following distance:

$$h_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + V^2 (t_i - t_j)^2} \quad (8)$$

where V is the average velocity of the storm system in the studied area. It is assumed here that the direction of movement is the x axis as it is generally observed for the squall lines in the Sahel.

The storm's average velocity can be estimated from remote-sensed data (radar images, for example) or from available spatio-temporal raingauge data (Johnson and Bras, 1979). In our case, an average velocity of 60 km/h was obtained for the two rainfall events and the wind direction is from east to west.

The following two approaches are considered in this study: (i) 3D kriging with standardized axes for the calculation of h , (ii) 3D kriging with velocity for the determination of h (equation (8)). 2D kriging is also performed to comparison purposes. All these approaches are applied for the target area and to the basic network for different cumulative time-step durations.

APPLICATIONS

The basic raingauge network of EPSAT-NIGER experiment is composed of 70 gauges spread out approximately evenly on a regular grid, with a 12.5 km * 12.5 km mesh covering the whole area of 12,000 km² (Lebel et al., 1992). A target area with a high density of available raingauges was considered for each year. In 1990, this target area was composed of sixteen (16) gauges installed in an area of about 400 km². In 1991, the target area was composed of 18 gauges over an area of 400 km² with a higher density of gauges. Figure 1 presents the EPSAT-NIGER raingauges network in 1991, and also the target area. The spatio-temporal data set of the best two squall-line rainfall events, 27/07/90 and 20/08/91, recorded during the EPSAT-NIGER experiments are considered and some of their characteristics are presented in Table 1.

Validation Methodologies

For each experiment, two uniform raingauge networks are formed with the same number of gauges: one for calibration and the other for validation. For the basic network, 35 raingauges are part of each subnetwork, but for the target area, each subnetwork is composed of 9 raingauges for the 20/08/91 rainfall event and 8 gauges for the 27/07/90 event

respectively. In the kriging estimation, constant trends and generalized covariance functions are considered. Two parameterized errors are considered: the Standardized Root Mean Square Error (SRMSE) and the ratio error (r). The SRMSE is defined as follows:

$$SRMSE = \sqrt{\frac{\sum_{i=1}^{i=N} (\hat{P}_i - P_i)^2}{N}} \frac{1}{\bar{p}} \quad (9)$$

where P_i and \hat{P}_i are the true and estimated rainfall values at gauge i respectively and \bar{p} represents the spatial average rainfall at the time considered. The ratio r_i , is proposed here since it is difficult from the SRMSE criterion to choose the best model for estimating rainfall at gauge i . This ratio is defined as follows:

$$r_i = \frac{|p_i - \hat{p}_i(3D)|}{|p_i - \hat{p}_i(2D)|} \quad (10)$$

where p_i , $\hat{p}_i(3D)$, $\hat{p}_i(2D)$ are the real, 3D and 2D estimated rainfalls at gauge i . The 3D model performs better than the 2D model at gauge i , if r_i is less than one; but if r_i is greater than one, the 2D model performs better than the 3D model. The two models are equivalent if $r_i=1$. The performance of 3D kriging will be evaluated by the total number of estimated values with a ratio r of less than one (percentage).

Example 1: target area

For the target area, we are interested in the performance of 3D kriging for small temporal and spatial scales. Cumulative time-step durations $\theta = 5, 10, 15, 20, 25$ and 30 min are considered. For each cumulative time duration, only the five most intense successive fields are considered for each event. For these five fields, four 3D kriging models are considered, each model being characterized by the number of different fields considered ($K = 2, 3, 4$ and 5, see equation (1)). For validation, only the rainfall values of the intense field are reconstituted from each of these four spatio-temporal kriging models. Then, the two subnetworks are inverted and the mean values of the resulting errors are considered.

Figures 2a, 2b and 3a, 3b show the variation of the SRMSE values against the number of fields considered K for 3D kriging and for different values of cumulative time duration for the two approaches and the two events. Figures 2a and 3a present results when a normalization of the three axes is performed and Figures 2b and 3b present results when the velocity is considered for the calculation of h . For comparison, 2D kriging results are indicated on these figures as $K=1$. It can be observed from these figures that the

velocity approach does not perform better than the axis standardization method, except for the cumulative time duration of 5 min where a decrease of SRMSE with K is observed.

The SRMSE of the velocity approach for the other time durations did not decrease with K. The SRMSE values can be assumed constant with K, which means that 2D and 3D kriging with the velocity approach yield the same results (Figures 2b and 3b). This result confirms indirectly the fact that, the Taylor hypothesis is also valid for the Sahelian squall lines rainfall events (Zawadzki, 1973). For the axes standardization approach, comparison of the 3D and 2D models leads to the following observations: 3D kriging generally performs better than the 2D model, but the improvement of the estimation is not significant. The performance of 3D kriging differs from event to event and depends on the cumulative time-step duration and on the number of rainfall fields considered.

Tables 2 and 3 show the percentage of values obtained by 3D and 2D krigings with a ratio r of less than one for different cumulative time-step durations and different 3D kriging models (here K is the number of fields considered in the 3D kriging) for the 27/07/90 and 20/08/91 rainfall

events. From Table 2, it can be seen that the 3D kriging models perform better than 2D kriging for cumulative time-step durations less than or equal to 20 min as it was observed for the parameter SRMSE, but the best performance is obtained for the 10 min duration. It can also be noted that the SRMSE parameters and the ratio r do not always lead to the same conclusions. For the 20/08/91 rainfall event, interpretation of the ratio r (percentage of r value less than one) is more difficult (see Table 3) since the 3D models perform better than 2D kriging for the cumulative times 5, 15 and 25 min, and less for times 10, 20 and 30 min.

Example 2: basic network

In this example, we are interested in the application of 3D kriging for large spatial scale. This leads us to consider the whole EPSAT-NIGER basic network and the cumulative time-step durations of 5, 15, 30, 45 and 60 minutes. Because of some numerical limitations due to the large number of samples, we have considered only the two most intense consecutive rainfall fields here. This is why only one 3D kriging model is used in this example and the validation is performed for each of these fields. The SRMSE values for the 3D and 2D krigings are presented together for each field as a function of the cumulative time-step

durations in Figures 4 and 5 for the 27/07/90 and 20/08/91 rainfall events only for the axes normalization approach. For the velocity approach, 3D and 2D krigings lead to the same estimation.

Table 4 presents the percentage of values of each of the two fields estimated by 3D and 2D krigings with a ratio of less than one for different cumulative time-step durations for the 27/07/90 and 20/08/91 events. As expected, the SRMSE values for 2D and 3D krigings models decrease with an increase in the cumulative time-step durations. This means that the estimation error decreases with the decrease of intermittency. Figures 4 and 5 and Table 4 lead to the following observations: the 3D kriging model generally performs better than the 2D kriging model. But again, the improvement is not really significant. For large cumulative times, there is practically no difference between the two models.

CONCLUSION

The dual formulation of kriging is applied to the spatio-temporal data sets of two EPSAT-NIGER rainfall events as a preliminary study of the rainfall estimation in the Sahel at small time steps. The large amount of interaction

between the spatial and temporal behavior of the rainfall phenomena at small time steps led to the design of a study to assess the performance of spatio-temporal estimation. Based on the results of this study, spatio-temporal kriging (3D kriging) is not really superior to the spatial estimation (2D kriging). The improvement obtained by 3D kriging is not very significant and depends principally on the cumulative time-step duration and on the number of fields considered. Thus, a simple direct extension of 2D kriging in time does not consider efficiently all the informations contained in the spatio-temporal structure of rainfall at small time steps.

Acknowledgments. Special thanks are addressed to the scholarship program of the Gouvernement of Canada (Bourse de la Francophonie) for their financial support of the graduate studies of the first author. We thank also the EPSAT-NIGER staff for providing us spatio-temporal rainfall data. The comments and suggestions of Dr. Zawadzki from McGill University and Dr. Lebel from ORSTOM, France are gratefully acknowledged. Finally, this research was partially supported by the Natural Sciences and Engineering Research Council of Canada on individual research grant (OGP0008415).

REFERENCES

- Amani, A. (1995) Estimation des pluies au Sahel: quelques problèmes de modélisation analysés à partir des données de l'expérience EPSAT-NIGER. Ph.D thesis, Ecole Polytechnique de Montréal, 289p.
- Amani, A., Lebel. T, Rousselle. J & Taupin, J.D. (1995) Typology of Rainfall Fields to Improve Rainfall Estimation in the Sahel by the Area Threshold Method (ATM). Resubmitted to Water Ressources Research.
- Bastin, G., Lorent. B, Duqué. C & Gervers, M. (1984) Optimal Estimation of Average Areal Rainfall and Optimal Selection of Rain Gauge Locations. Water Resour. Res., 20(4):463-470.
- Bilonick, R.A. (1985) The Space-Time Distribution of Sulfate Deposition in the Northeastern United States. Atmospheric Environment, 19(11): 1829-1845.
- Chua, S.H & Bras, R.L. (1982) Optimal Estimation of Mean Areal Precipitation in Region of Orographic Influence. J.Hydrol., 57(1/2): 23-48.
- Creutin, J.D & Obled, C. (1982) Objective Analysis and Mapping Techniques for Rainfall Fields: an Objective Comparaison. Water Resour. Res., 18(2): 413-431.

Delrieu, G., Bellon. A & Creutin, J.D. (1988) Estimation de lames d'eau spatiales à l'aide de données de pluviomètres et de radar météorologique: Application au pas de temps journalier dans la région de Montréal. *J.Hydrol.*, 98(2): 315-344.

Gupta, V.K & Waymire, E. (1987) On Taylor's Hypothesis and Dissipation in Rainfall. *J.Geophys.Res.*, 92(D8): 9657-9660.

Jonhson, E.R & Bras, R.L. (1979) Real-time Estimation of Velocity and Covariance Structure of Rainfall Events using Telemetered Raingauge Data. A comparison of Methods. *J. Hydrol.*, 44(1): 97-123.

Lebel, T., Bastin. G, Obled. C & Creutin, J.D. (1987) On the accuracy of areal rainfall estimation: A case study. *Water Resour.res.*, 23(11): p2123-2134.

Lebel, T., Sauvageot. H, Hoepffner. M, Desbois. M, Guillot. B & Hubert, P. (1992) Rainfall Estimation in the Sahel: The EPSAT-NIGER Experiment. *Hydrological Sciences Journal*, 37(3): 201-215.

Marche, C., Lessard, G & El Gharbi, B. (1990) Le krigeage dual, une technique pour la cartographie automatisée des zones d'inondation. *J. Hydraul. Res.*, 28(5): 629-643.

Meyers, D.E. & Journel, A. (1990) Variograms with Zonal Anisotropies and Noninvertible Kriging Systems. *Math.Geology*, 22(7): 779-785.

Rodriguez-Iturbe, I. & Mejia, J.M, 1974: The Design of Rainfall Networks in Time and Space. *Water Resour. Res.*, 10(4): 713-728.

Rouhani, S. & Hall, T.J. (1989) Space-Time Kriging of Groundwater Data. M.Armstrong (ed), *Geostatistics*, vol.2: 639-650.

Rouhani, S. & Meyers, D.E. (1990) Problems in Space-Time Kriging of Geohydrological Data. *Math.Geology*, 22(5): 611-623.

Seo, D.J., Frajewski, W.F & D.S.Bowles, D.S. (1990) Stochastic Interpolation of Rainfall Data from Rain Gauges and Radar using Cokriging: 1- Design of Experiments. *Water Resour. Res.*, 26(3): 466-477.

Trochu, F. (1993) A Contouring Program Based on Dual Kriging Interpolation. *Engineering with Computers*, 9: 160-177.

Zawadzki, I. (1973) Statistical Properties of Precipitation Patterns, *J. Appl. Meteorol.*, 12(2):469-472.

TABLE 1: Some characteristics of the 27/07/90 and 20/08/91 rainfall events.

	27/07/90 Event	20/08/91 Event
Duration (min)	335	320
Areal rainfall (mm)	23.9	30.9
Recorded gauges	75	85
5 min maximum accumulation (mm)	11.4	13
30 min maximum accumulation (mm)	46.0	36.4
60 min maximum accumulation (mm)	54.1	44.0

TABLE 2: Percentage of estimated rainfall values with a ratio r less than one for the different 3D kriging models (K) and cumulative time-step durations for the 27/07/90 event.

K: number of fields	5 min	10 min	15 min	20 min	25 min	30 min
2	50.0	62.5	62.5	50.0	37.5	12.5
3	62.5	75.0	62.5	62.5	50.0	37.5
4	62.5	87.5	62.5	62.5	37.5	37.5
5	37.5	75.0	50.0	62.5	37.5	37.5

TABLE 3: Percentage of reconstituted rainfall values with a ratio r less than one for the different 3D kriging models (K) and cumulative time-step durations for the 20/08/91 event.

K: number of fields	5 min	10 min	15 min	20 min	25 min	30 min
2	66.7	0.0	66.7	33.3	83.3	16.7
3	50.0	0.0	50.0	33.3	83.3	16.7
4	66.7	0.0	50.0	33.3	83.3	16.7
5	66.7	0.0	0.0	33.3	83.3	16.7

TABLE 4: Percentage of reconstituted rainfall values of the two fields considered using **3D and 2D krigings** for the cumulative time-step durations 5, 15, 30, 45 and 60 min for the 27/07/90 and 20/08/91 events.

Event	field	5 min	15 min	30 min	45 min	60 min
27/07/90	1	59.0	56.0	53.0	50.0	50.0
27/07/90	2	58.0	56.0	53.0	50.0	50.0
20/08/91	1	47.0	53.0	50.0	47.0	50.0
20/08/91	2	53.0	53.0	53.0	53.0	50.0

Figure Captions

Figure 1. EPSAT-NIGER raingauges basic network and target area in 1991.

Figure 2. Standardized Root Mean Square Error (SRMSE) versus the number of fields considered for the spatio-temporal kriging K of the target area and at cumulative time-step durations of 5, 10, 15, 20, 25 and 30 min for the 27/07/90 event; a) first approach, and b) second approach.

Figure 3. Standardized Root Mean Square Error (SRMSE) versus the number of fields considered for the spatio-temporal kriging K of the target area and at cumulative time-step durations of 5, 10, 15, 20, 25 and 30 min for the 20/08/91 event, a) first approach, and b) second approach.

Figure 4. Standardized Root Mean Square Error (SRMSE) versus the cumulative time-step duration for 2D and 3D kriging models for the basic network (27/07/90 event).

Figure 5. Standardized Root Mean Square Error (SRMSE) versus the cumulative time-step duration for 2D and 3D kriging models for the basic network (20/08/91 event).

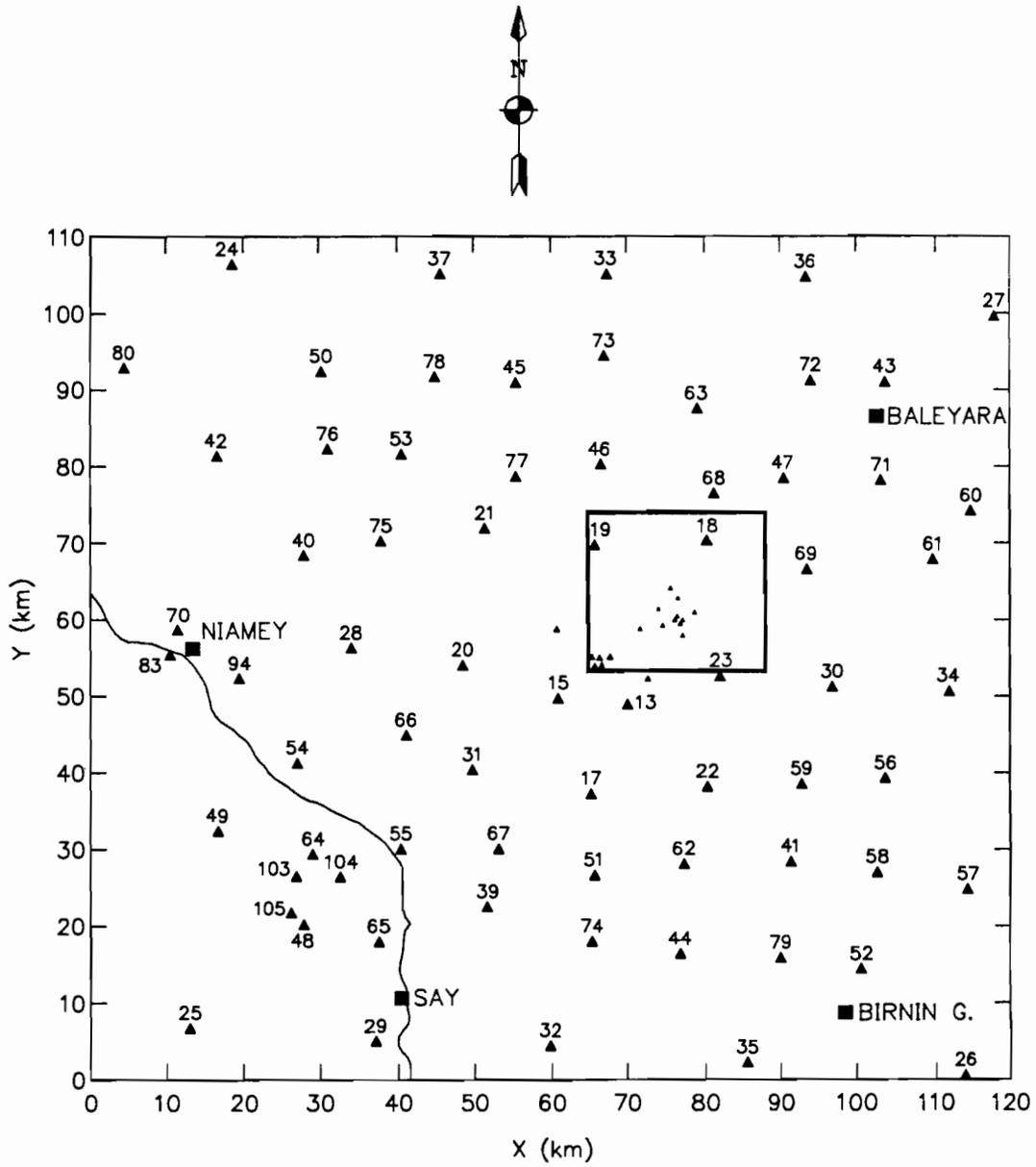


Figure 1

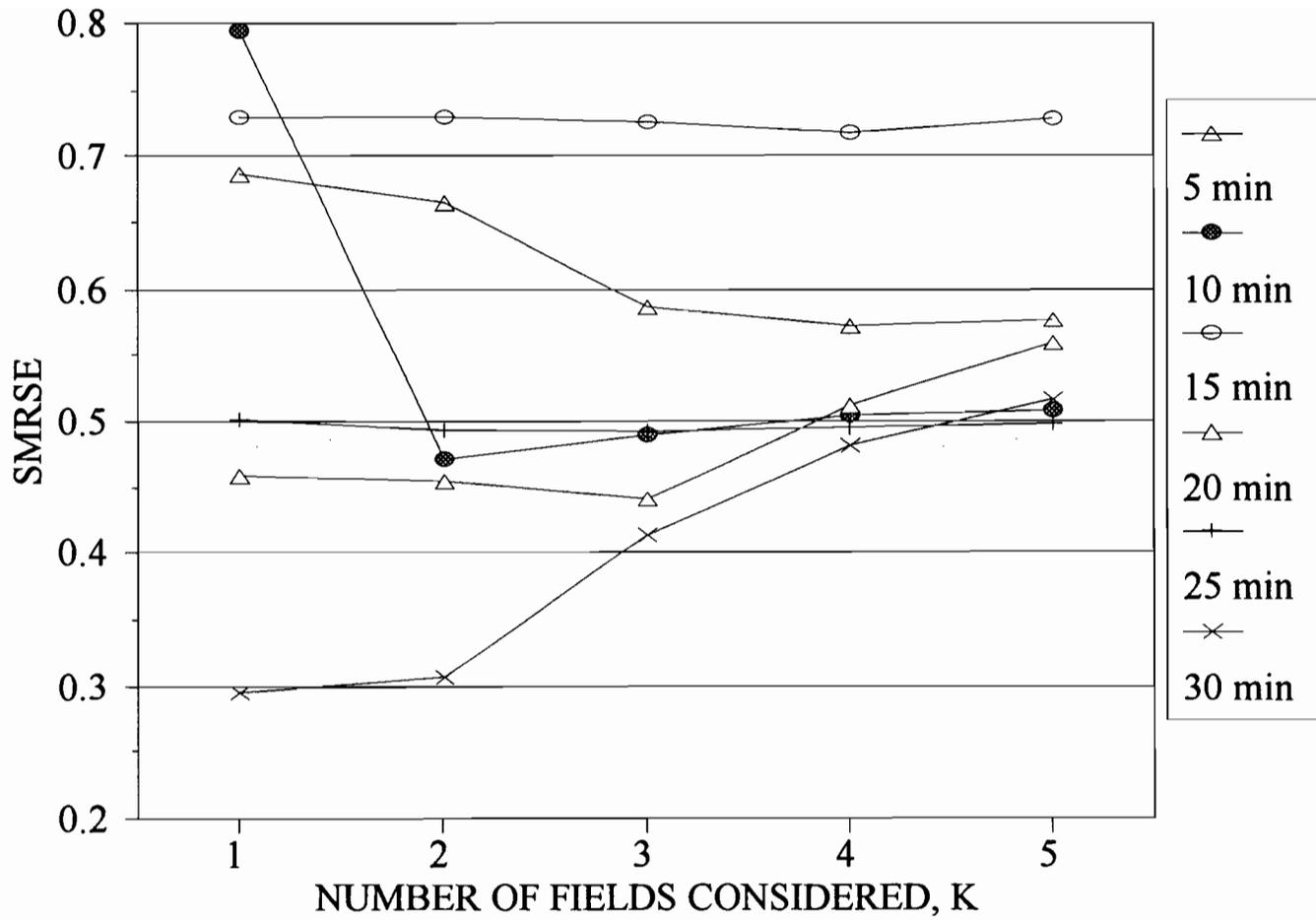


Figure 2a

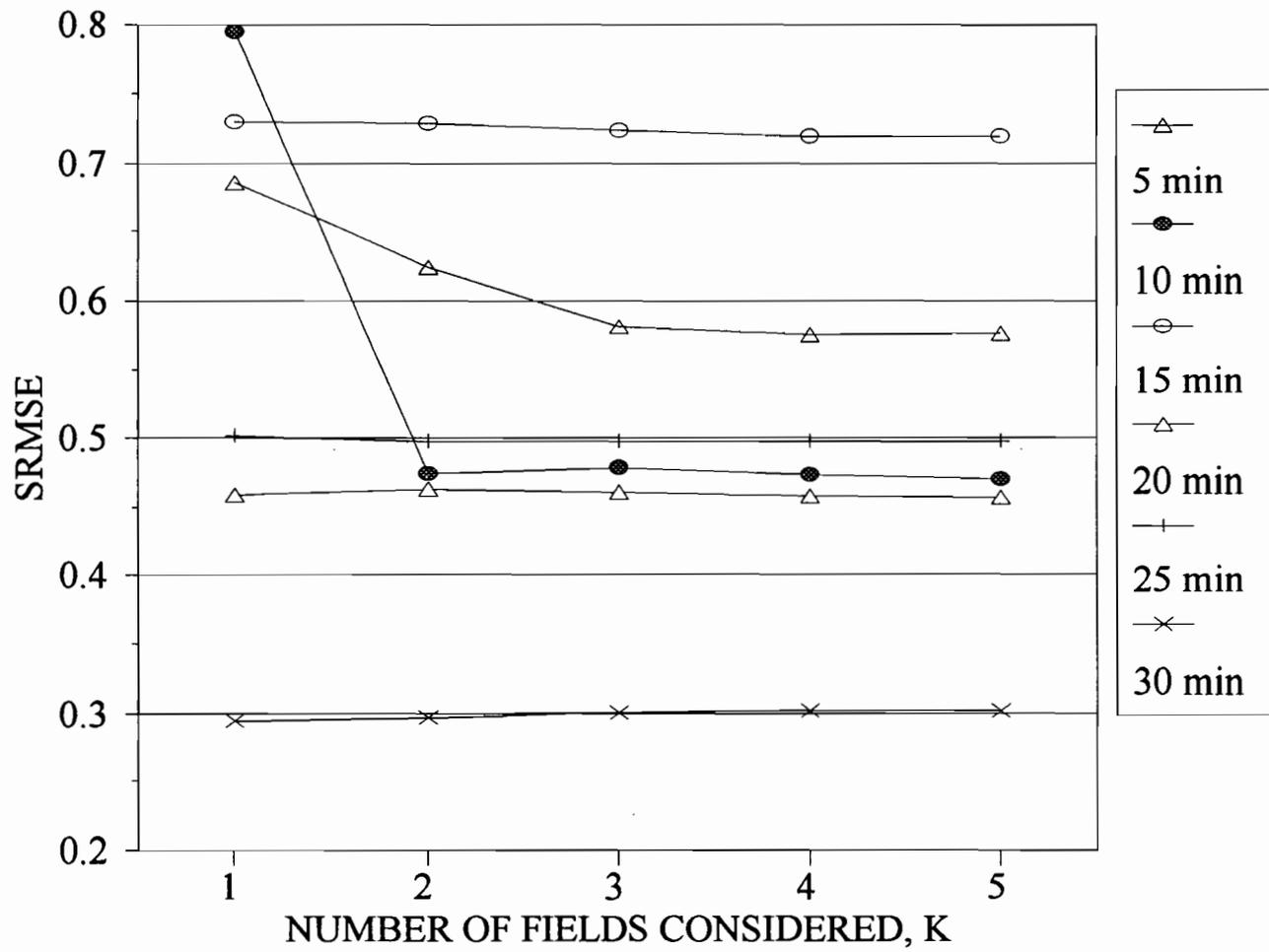


Figure 2b

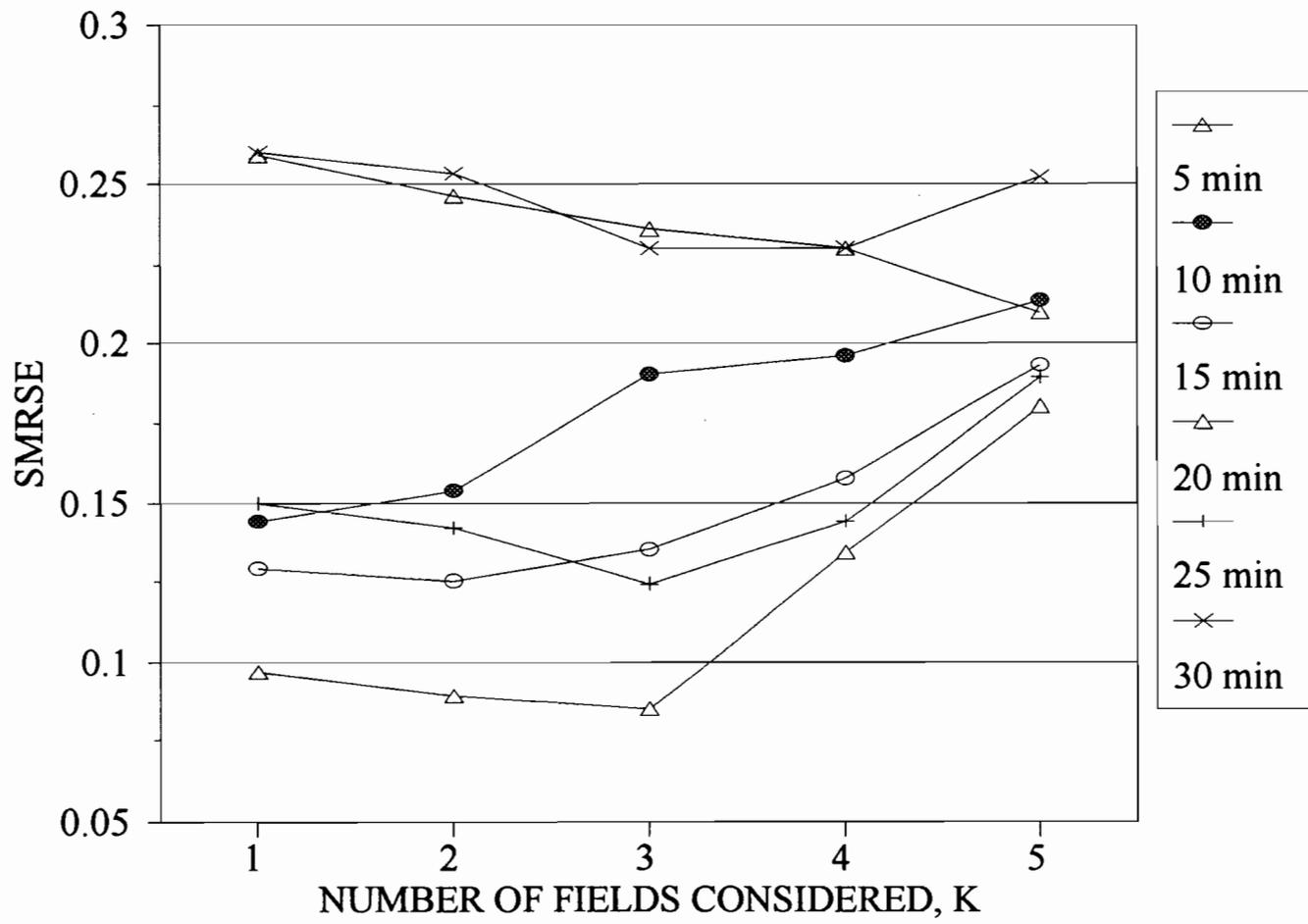


Figure 3a

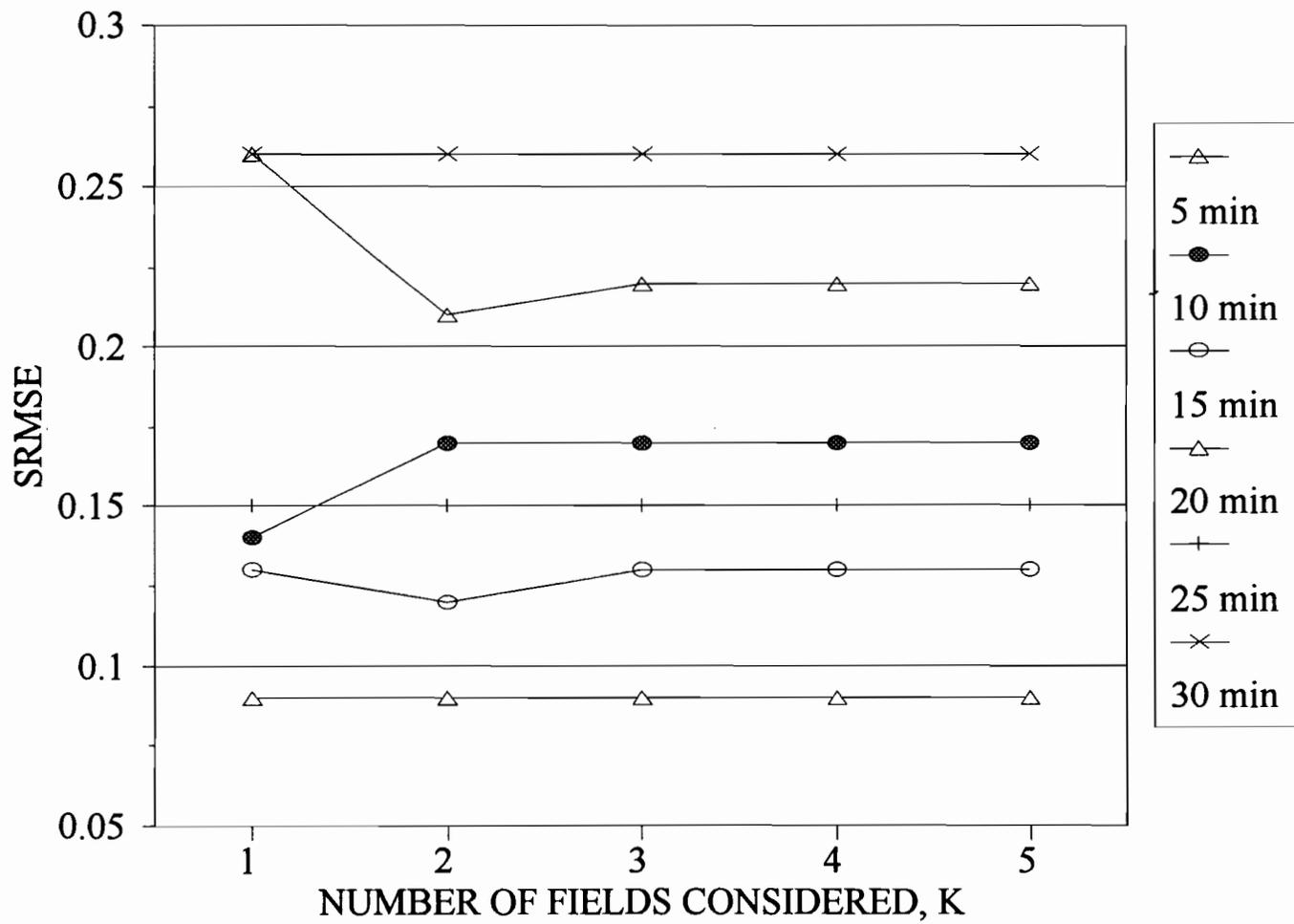


Figure 3b

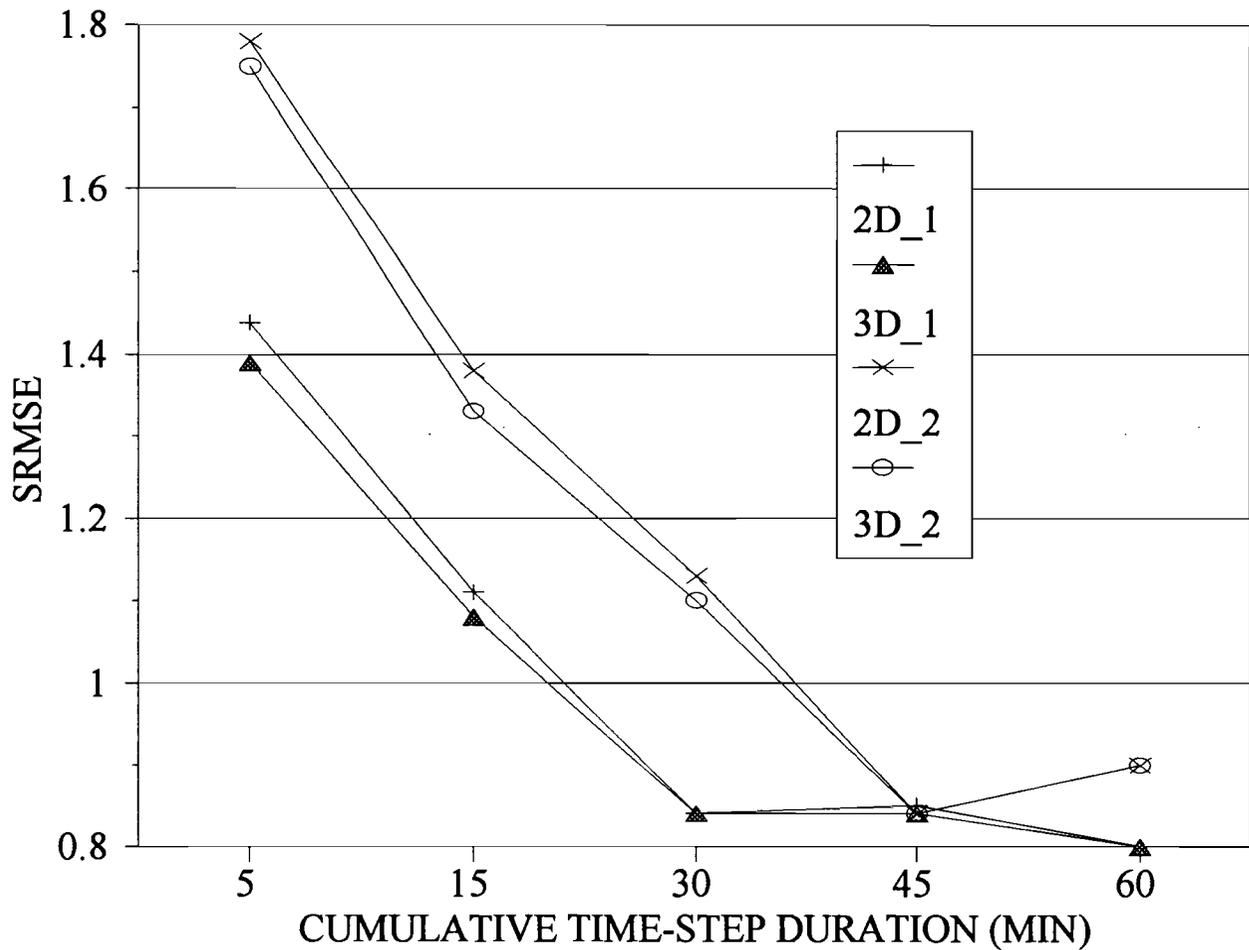


Figure 4

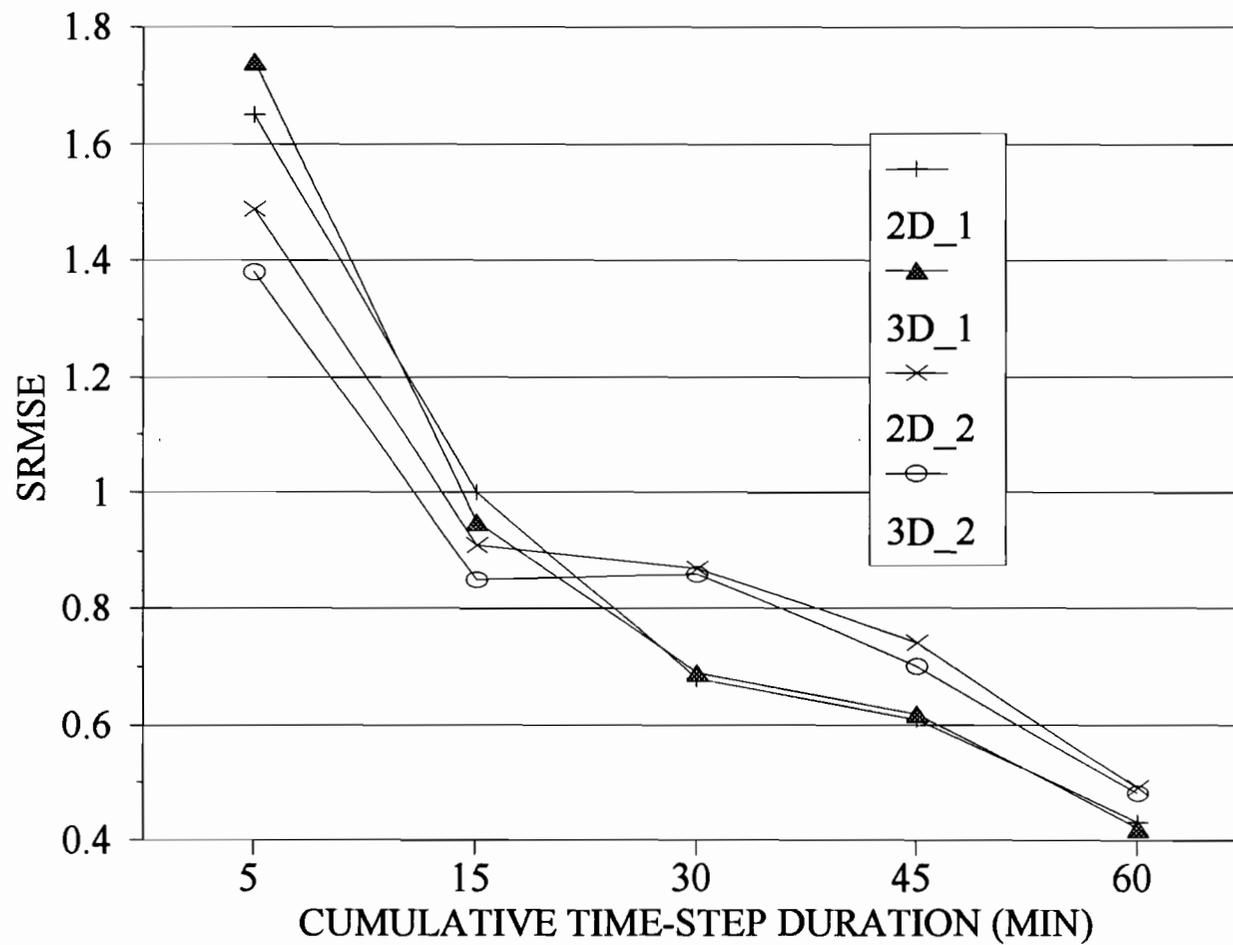


Figure 5

ANNEXE B:

AUTRES RESULTATS

ANNEXE B1:

Complément des résultats sur la classification

Tableau B1.1: Valeurs des paramètres du variogramme (effet de pépite, portée et palier), du CV et de UPA pour les événements pluvieux de 1990.

Pluies	Pépite (mm ²)	Portée (km)	Palier (mm ²)	Moy (mm)	CV	UPA
23/05	75.0	0	75.0	4.0	1.85	0.832
28/05	100	50	500	24.3	0.83	0.657
07/06	-	-	-	7.9	1.67	0.839
12/06	45	50	60	5.0	1.32	0.774
17/06	0	40	140	7.8	1.43	0.785
22/06	0	40	30	3.8	1.43	0.836
24/06	50	40	280	27.2	0.66	0.584
27/06	0	40	200	5.2	1.96	0.880
29/06	60	0	60	3.2	1.96	0.876
03/06	125	0	125	7.8	1.39	0.823
05/07	-	-	-	5.1	2.31	0.896
08/07	0	40	200	5.5	2.15	0.890
09/07	-	-	-	1.5	2.58	0.883
12/07	0	40	200	17.4	0.83	0.684
15/07	0	45	150	2.9	2.84	0.906
15/07	0	60	45	4.9	1.13	0.753
18/07	0	40	90	8.6	1.07	0.752
21/07	120	0	120	7.7	1.39	0.804
21/07	0	40	20	2.7	1.56	0.844
22/07	10	40	40	4.4	1.37	0.727

Tableau B1.1: (suite)

Pluies	Pépite (mm²)	Portée (km)	Palier (mm²)	Moy (mm)	CV	UPA
23/07	180	0	180	18.6	0.75	0.558
27/07	0	50	140	23.7	0.52	0.519
31/07	0	40	150	12.8	0.99	0.733
02/08	0	60	120	4.7	1.81	0.849
04/08	150	0	150	26.6	0.48	0.512
08/08	0	40	120	20.4	0.54	0.579
14/08	5	0	5	1.0	1.96	0.840
17/08	100	50	320	16.7	1.04	0.731
24/08	0	60	240	11.0	1.18	0.733
28/08	0	50	50	3.8	1.57	0.834
29/08	-	-	-	3.9	1.46	0.846
01/09	-	-	-	17.2	1.01	0.758
05/09	-	-	-	7.3	1.43	0.823
07/09	-	-	-	13.0	1.36	0.811
11/09	0	50	100	2.5	2.79	0.908
13/09	60	0	60	15.6	0.54	0.584
17/09	-	-	-	19.5	0.96	0.697

Tableau B1.2: Valeurs des paramètres du variogramme (effet de pépite, portée et palier), du CV et de UPA pour les événements pluvieux de 1991.

Pluies	Pépite (mm ²)	Portée (km)	Palier (mm ²)	Moy (mm)	CV	UPA
14/04	-	-	-	4.6	2.21	0.805
28/04	0	10	30	8.0	1.04	0.680
04/05	0	50	470	28.5	0.73	0.521
16/05	-	-	-	4.6	2.19	0.884
20/05	0	10	22	5.1	1.18	0.701
24/05	-	-	-	28.0	1.04	0.712
25/05a	0	25	16	2.1	1.68	0.811
25/05b	0	10	70	3.1	3.28	0.939
26/05a	0	20	50	10.3	0.97	0.816
26/05b	0	30	800	29.9	0.99	0.788
27/05	0	20	10	1.5	2.11	0.843
28/05	0	25	15	3.3	1.51	0.832
01/06	-	-	-	6.2	1.17	0.726
03/06	10	12	55	14.0	0.58	0.595
07/06	10	10	45	7.9	0.94	0.654
11/06	0	50	130	12.0	0.95	0.718
15/06	0	25	80	9.2	1.28	0.816
20/06	0	50	60	4.6	1.63	0.837
21/06	0	25	3.5	1.7	1.07	0.717
23/06	0	22	6	2.5	0.96	0.747
26/06	0	35	350	27.4	0.66	0.597

Tableau B1.2: (suite)

Pluies	Pépité (mm ²)	Portée (km)	Palier (mm ²)	Moy (mm)	CV	UPA
30/06	0	30	90	7.1	1.91	0.851
04/07	0	35	70	6.6	1.24	0.788
08/07	0	32	140	12.8	0.99	0.726
14/07	-	-	-	18.1	0.92	0.680
17/07	0	50	160	8.6	1.39	0.828
18/07	0	10	2.5	1.0	1.59	0.800
20/07	0	15	110	19.0	0.62	0.622
25/07	0	10	160	14.2	0.82	0.628
26/07	0	30	30	3.0	1.88	0.857
30/07	0	25	80	5.4	1.64	0.865
03/08	0	27	400	37.1	0.54	0.520
06/08	0	40	13	2.7	1.36	0.756
07/08	-	-	-	4.1	1.61	0.824
08/08	0	15	50	14.8	0.56	0.620
10/08	0	12	180	12.3	1.32	0.800
14/08	0	65	35	2.3	2.14	0.872
17/08	-	-	-	19.0	0.55	0.656
18/08	0	30	40	8.3	0.86	0.728
20/08	0	15	65	30.0	0.27	0.381
23/08	-	-	-	5.0	1.86	0.889
26/08	0	30	120	12.9	0.82	0.646
29/08	0	17	140	14.8	0.78	0.687
31/08	0	10	70	11.0	0.86	0.682
02/09	0	11	95	11.8	0.86	0.704
03/09	0	10	80	21.9	0.49	0.504
04/10	0	10	60	3.4	0.85	0.596

Tableau B1.3: Valeurs des paramètres du variogramme (effet de pépite, portée et palier), du CV et de UPA pour les événements pluvieux de 1992.

Pluies	Pépite	Portée (km)	Plier (mm ²)	Moy (mm)	CV	UPA
10/04a	0	25	2	0.9	1.67	0.654
10/04b	0	30	25	3.8	2.05	0.873
11/05	0	30	130	5.0	2.10	0.872
25/05	0	30	125	5.5	1.93	0.768
28/05	0	35	35	2.4	2.39	0.752
01/06	0	50	220	10.9	1.28	0.876
03/06	0	40	7	1.2	2.15	0.849
04/06	20	35	90	11.8	0.82	0.824
06/06	0	25	120	10.7	1.16	0.901
12/06	0	50	80	6.5	1.34	0.871
20/06	0	-	-	6.6	1.37	0.800
23/06	0	35	45	4.0	1.72	0.667
25/06	0	5	5	1.8	2.07	0.805
30/06	0	10	130	22.1	0.67	0.647
07/07	0	12	12	4.2	1.14	0.866
11/07	0	30	40	7.2	1.00	0.771
13/07	0	35	95	10.9	1.11	0.839
18/07	0	12	230	28.2	0.61	0.699
20/07	0	16	35	2.3	2.50	0.853
22/07	0	10	240	30.7	0.54	0.691
27/07	0	23	90	9.3	0.97	0.835
28/07	0	17	55	3.0	2.54	0.881
31/07	0	15	290	34.7	0.52	0.695
04/08	0	20	70	3.5	2.64	0.895
05/08	0	20	30	2.6	2.81	0.861

Tableau B1.3: (suite)

Pluies	Pépîte	Portée (km)	Palier (mm ²)	Moy (mm)	CV	UPA
08/08	0	25	300	22.2	0.78	0.762
10/08	0	50	90	25.4	0.62	0.713
11/08	0	15	30	7.3	0.90	0.761
13/08	5	-	-	3.6	2.04	0.720
15/08	0	28	200	13.4	1.20	0.805
18/08	0	20	9	2.3	1.74	0.787
21/08	0	28	170	46.6	0.33	0.540
22/08	0	15	17	2.0	1.87	0.748
24/08	0	25	25	3.3	1.48	0.842
25/08a	0	25	30	4.3	1.31	0.836
25/08b	0	20	9	1.4	2.08	0.773
25/08c	0	55	125	9.4	1.26	0.869
27/08	0	40	70	8.4	1.01	0.872
28/08a	0	50	2	1.2	1.26	0.756
28/08b	0	15	120	12.7	0.90	0.777
29/08	-	-	-	6.7	1.41	0.901
30/08a	0	15	37	26.1	0.25	0.491
30/08b	20	30	150	12.3	0.93	0.723
02/09	0	-	-	5.7	1.94	0.853
06/09	0	35	100	19.9	0.61	0.626
11/09	0	20	210	18.3	0.87	0.641
14/09	0	15	35	14.7	0.63	0.708
15/09a	0	25	2.5	1.0	1.89	0.802
15/09b	0	-	-	1.2	2.20	0.821

Tableau B1.4: caractéristiques principales des événements pluvieux observés en 1990
(durée D, nombre de stations touchées N, cumul maximum C5, C10, C15, C30, C60 respectivement pour les temps de cumul de 5, 10, 15, 30 et 60 minutes).

Event	N	D (mm)	C5 (mm)	C10 (mm)	C15 (mm)	C30 (mm)	C60 (mm)
23/05	36	150	10.6	20.2	26.7	39.4	42.5
28/05	70	420	14.6	24.1	32.3	56.9	74.5
07/06	56	240	11.5	21.4	32.5	58.6	72.6
12/06	62	310	8.5	15.4	21.5	25.4	28.8
17/06	46	225	11.9	20.6	30.3	50.9	58.8
22/06	56	200	8.0	10.8	13.1	22.6	31.5
24/06	79	365	21.4	35.9	49.3	66.0	70.5
27/06	39	795	12.1	19.6	27.1	34.7	45.1
29/06	33	185	8.5	15.3	22.2	31.8	32.5
03/07	71	280	6.7	13.0	17.6	28.2	48.0
05/07	31	195	13.2	20.4	27.5	44.9	67.0
08/07	42	215	30.0	45.3	52.3	61.5	62.5
09/07	35	170	5.6	10.4	15.5	24.0	29.5
12/07	80	555	13.2	22.2	29.8	42.5	53.0
15/07	36	135	14.7	26.7	35.8	52.6	57.0
15/07	73	180	7.6	14.1	17.1	20.0	21.0
18/07	68	445	13.7	26.1	35.7	47.0	47.0
21/07	50	255	11.1	22.1	31.1	44.0	46.5
21/07	41	355	6.9	10.7	14.4	18.8	19.0

Tableau B1.4: (suite)

Pluies	N	D (mm)	C5 (mm)	C10 (mm)	C15 (mm)	C30 (mm)	C60 (mm)
22/07	61	430	4.4	7.9	10.7	15.0	18.5
23/07	77	345	10.8	18.6	26.6	38.3	43.0
27/07	80	340	11.4	21.0	29.0	46.0	54.1
31/07	70	270	12.0	21.5	29.6	40.6	54.5
02/08	42	230	10.7	15.6	21.0	29.7	34.0
04/08	83	450	13.9	25.0	33.9	48.9	52.5
08/08	86	450	12.0	22.2	31.9	37.4	42.0
14/08	28	175	2.9	4.5	5.3	7.0	7.0
17/08	71	435	15.0	24.2	31.7	45.1	55.1
24/08	73	395	20.1	27.7	33.2	39.2	41.0
28/08	55	300	7.5	11.2	16.1	21.5	25.0
29/08	73	275	7.8	15.5	20.1	27.4	28.5
01/09	85	540	15.9	26.1	35.9	59.2	66.2
05/09	83	315	14.2	24.3	33.1	47.0	48.0
07/09	71	335	10.0	18.9	26.9	47.2	64.3
11/09	33	285	9.0	17.5	25.6	36.1	43.2
13/09	87	290	13.1	24.9	30.8	42.3	45.5
17/09	86	430	14.0	25.8	34.1	52.0	73.5

Tableau B1.5: caractéristiques principales des événements pluvieux observés en 1991
(durée D, N, cumul maximum C5, C10, C15, C30, C60 respectivement pour
les temps de cumul de 5, 10, 15, 30 et 60 minutes).

Event	N	D (mm)	C5 (mm)	C10 (mm)	C15 (mm)	C30 (mm)	C60 (mm)
14/04	6	115	6.3	11.0	14.6	21.9	28.8
28/04	30	295	6.2	12.1	17.0	23.9	34.4
04/05	42	315	8.4	16.5	22.9	37.7	51.2
16/05	36	365	9.7	17.5	22.6	37.7	42.0
20/05	59	200	6.7	11.2	13.3	19.0	25.0
24/05	59	765	10.4	19.4	27.7	47.8	69.4
25/05a	32	135	5.4	9.0	9.7	12.0	12.5
25/05b	34	320	8.6	17.1	25.2	47.3	64.5
26/05a	73	415	6.1	11.0	14.9	29.3	45.0
26/05b	75	1000	11.9	20.1	24.9	47.1	84.5
27/05	32	265	4.6	6.9	8.8	10.0	17.0
28/05	56	320	7.2	12.1	14.0	20.1	22.0
01/06	56	130	11.5	17.1	20.6	21.0	31.5
03/06	82	360	13.6	20.8	25.5	32.9	33.5
07/06	73	170	10.4	16.5	21.4	24.5	24.5
11/06	74	220	15.0	29.4	32.0	42.5	43.2
15/06	74	225	15.7	26.9	37.8	52.1	54.5
20/06	54	410	7.2	13.5	18.3	31.9	41.0
21/06	67	170	2.1	3.2	3.9	5.0	6.5
23/06	83	150	6.1	10.1	11.5	11.5	11.5
26/06	88	470	11.5	22.7	31.9	51.2	73.2
30/06	49	175	16.4	30.7	41.2	47.5	53.0
04/07	73	320	12.2	18.8	25.0	34.3	38.9
08/07	74	300	8.6	16.7	23.3	36.4	47.0

Tableau B1.5: (suite)

Event	N	D (mm)	C5 (mm)	C10 (mm)	C15 (mm)	C30 (mm)	C60 (mm)
14/07	74	325	12.9	23.2	31.0	44.6	52.8
17/07	68	295	10.7	17.3	21.5	36.5	65.0
18/07	47	190	3.7	6.3	6.9	7.0	8.5
20/07	89	495	14.7	24.6	32.2	42.5	60.7
25/07	85	285	12.0	19.3	22.8	34.6	38.0
26/07	40	125	10.3	18.6	25.1	27.5	27.5
30/07	59	280	11.0	20.3	26.0	35.8	39.5
03/08	84	590	11.0	20.4	29.6	51.7	62.3
06/08	57	210	5.3	10.0	14.4	18.5	19.5
07/08	51	225	8.1	13.8	20.1	27.0	30.5
08/08	84	410	8.3	15.8	21.1	32.3	36.5
10/08	77	450	13.5	24.3	32.4	51.4	69.0
14/08	26	205	7.1	13.5	17.4	21.4	21.5
17/08	85	315	15.7	28.6	37.7	46.3	52.0
18/08	78	350	7.4	13.5	20.0	31.2	34.5
20/08	85	335	13.0	20.1	24.9	36.4	44.0
23/08	59	305	9.2	17.9	26.4	41.2	44.5
26/08	85	330	11.6	21.9	29.4	43.3	47.0
29/08	90	380	10.6	18.0	25.3	33.1	36.0
31/08	87	310	6.6	11.7	16.0	26.6	36.8
02/09	89	295	10.8	17.3	23.6	35.6	41.0
03/10	72	310	10.3	18.3	23.8	33.8	42.5
04/10	61	165	1.5	2.4	2.9	4.8	7.5

Tableau B1.6: Caractéristiques principales des événements pluvieux observés en 1992
(durée D, cumul maximum C5, C10, C15, C30, C60 pour les pas de temps
5, 10, 15, 30 et 60 minutes).

Pluies	N	D (mn)	C5 (mm)	C10 (mm)	C15 (mm)	C30 (mm)	C60 (mm)
10/04a	21	145	2.1	3.1	4.4	6.5	6.5
10/04b	31	305	7.0	13.1	18.6	26.0	27.0
11/05	45	435	17.3	26.3	31.2	38.4	45.0
25/05	42	205	11.5	18.2	25.0	33.8	51.7
28/05	32	295	7.1	13.0	17.1	27.8	29.2
01/06	78	255	21.3	31.6	38.5	52.2	62.6
03/06	38	165	6.0	10.6	12.4	13.5	13.5
04/06	88	445	9.5	15.8	20.5	37.0	43.5
06/06	93	335	20.4	30.5	40.8	56.9	62.0
12/06	75	200	15.2	27.7	35.3	51.3	51.5
20/06	68	330	11.1	20.8	25.7	31.1	32.5
23/06	38	150	10.7	18.8	23.1	26.0	26.0
25/06	55	200	7.3	12.8	16.6	26.0	30.5
30/06	100	295	28.9	38.0	42.3	47.2	55.1
07/07	91	185	8.6	17.1	18.9	21.5	22.0
11/07	82	225	8.7	14.0	16.5	22.8	27.0
13/07	88	305	8.5	16.0	22.7	29.6	44.0
18/07	103	500	12.6	24.7	34.3	48.8	56.7
20/07	32	230	7.9	13.6	18.4	26.7	29.5
22/07	102	315	15.7	27.6	38.9	58.2	68.8
27/07	98	360	10.6	19.9	30.2	40.5	56.0
28/07	44	200	11.9	23.4	30.3	43.9	56.5
31/07	100	430	13.4	23.6	31.9	47.6	58.0

Tableau B1.6: (suite)

Pluies	N	D (mm)	C5 (mm)	C10 (mm)	C15 (mm)	C30 (mm)	C60 (mm)
04/08	42	270	8.5	16.8	24.3	46.0	67.3
05/08	40	205	11.4	20.7	29.5	40.3	42.5
08/08	98	485	15.9	25.3	35.2	62.6	66.0
10/08	100	525	9.1	17.8	25.2	44.7	56.5
11/08	92	280	10.8	14.7	18.5	27.5	28.0
13/08	34	340	8.4	12.6	14.1	16.7	20.0
15/08	85	330	10.0	19.8	28.0	48.1	74.5
18/08	51	245	5.3	9.2	10.6	13.5	14.6
21/08	104	705	10.5	16.7	22.7	39.5	56.0
22/08	43	125	9.8	12.3	12.9	16.0	18.5
24/08	67	285	8.0	13.0	16.2	24.7	26.5
25/08a	80	250	11.3	16.3	17.5	20.2	26.5
25/08b	34	100	5.4	7.8	9.3	12.5	13.0
25/08c	89	345	7.9	14.3	20.6	28.5	41.4
27/08	97	270	10.2	19.5	28.0	40.7	41.5
28/08a	65	205	3.1	4.8	5.0	5.0	5.5
28/08b	100	420	13.6	25.2	31.2	37.1	37.5
29/08	84	305	7.8	13.2	19.1	28.7	38.0
30/08a	105	485	11.2	20.4	26.4	32.6	35.5
30/08b	88	265	15.2	27.8	33.9	38.0	38.0
02/09	56	500	9.6	18.0	24.4	36.2	47.8
06/09	101	405	12.2	23.4	27.0	33.1	37.0
11/09	88	345	13.6	24.2	33.6	54.7	66.7
14/09	105	400	6.2	11.0	15.9	22.2	25.7
15/09a	49	180	4.8	6.9	7.7	8.0	8.5
15/09b	40	155	6.3	10.6	12.4	13.5	13.5

Tableau B1.7: Groupes des événements pluvieux pour les neuf méthodes de classification étudiées pour l'année 1990. (M1= CDA, M2= CMRCA; M3= CA; M4=SCA; M5= CUPAA; M6=CCVA; M7=CCVUPAA; M8=DUPA et M9= CMRCA+CUPAA)

Event	M1	M2	M3	M4	M5	M7	M6	M8	M9
23/05	2	1	2a	1b	2	1	1	3	1
28/05	3	2	1b	1b	1	3	2	2	2
07/06	-	2	1b	1b	2	1	1	3	1
12/06	1	1	1b	1b	1	3	1	2	1
17/06	1	2	1b	1b	1	3	2	2	1
22/06	1	1	3b	1b	2	1	1	3	1
24/06	3	3	1b	1b	3	2	3	1	2
27/06	1	1	4a	1b	2	1	1	3	1
29/06	2	1	2a	1b	2	1	1	3	1
03/07	2	1	2a	1b	2	1	1	3	1
05/07	-	2	1b	1b	2	1	1	3	1
08/07	1	3	1b	1b	2	1	1	3	3
09/07	-	1	1b	1b	2	4	4	3	1
12/07	1	2	1b	1b	1	3	1	2	1
15/07a	1	2	2a	1b	2	4	4	3	1
15/07b	1	1	1b	1b	1	3	1	2	1
18/07	1	2	2a	1b	1	3	2	2	1
21/07a	2	2	2a	1b	2	3	2	2	1
21/07b	1	1	1b	1b	2	3	2	3	1
22/07	1	1	2b	1b	1	3	1	2	1

Tableau B1.7: (suite)

Event	M1	M2	M3	M4	M5	M7	M6	M8	M9
23/07	2	1	3b	2b	3	2	3	1	2
27/07	1	2	1b	2a	3	2	3	1	2
31/07	1	2	1b	1b	1	3	2	2	1
02/08	1	1	1b	1b	2	1	1	3	1
04/08	2	2	1b	2a	3	2	3	1	2
08/08	1	2	1b	2a	3	2	3	1	2
14/08	2	4	1b	1b	2	3	2	3	1
17/08	3	2	1b	1b	1	3	2	2	2
24/08	1	2	1b	1b	1	3	1	2	1
28/08	1	1	5a	1b	2	1	1	3	1
29/08	-	1	1b	1b	2	1	1	3	1
01/09	-	2	1b	1b	1	3	1	2	2
05/09	-	2	1b	1b	2	1	1	3	1
07/09	-	2	1b	1b	2	1	1	3	1
11/09	1	1	1b	1b	2	4	4	3	1
13/09	2	2	1b	2a	3	2	3	1	2
17/09	-	2	2a	1b	1	3	2	2	2

Tableau B1.8: Groupes des événements pluvieux pour les neuf méthodes de classification étudiées pour l'année 1991. (M1= CDA, M2= CMRCA; M3= CA; M4=SCA; M5= CUPAA; M6=CCVA; M7=CCVUPAA; M8= DUPA et M9= CMRCA+CUPAA)

pluies	M1	M2	M3	M4	M5	M6	M7	M8	M9
14/04	-	-	2a	1b	1	1	2	3	1
28/04	2	6	3b	1b	2	2	1	2	1
04/05	7	3	1b	2b	2	2	1	1	2
16/05	-	3	1b	1b	1	1	2	3	1
20/05	2	6	1b	1b	2	1	1	2	1
24/05	-	1	4a	1b	2	2	1	2	4
25/05a	1	6	3b	1b	1	1	2	3	1
25/05b	2	1	3b	1b	1	3	3	3	1
26/05a	1	6	2a	1b	1	2	2	3	1
26/05b	8	2	1b	1b	1	2	2	2	4
27/05	1	6	2a	1b	1	1	2	3	1
28/05	1	6	2a	1b	1	1	2	3	1
01/06	-	3	3a	1b	2	1	1	2	2
03/06	3	3	1b	2a	2	2	1	2	2
07/06	3	3	1b	1b	2	2	1	1	2
11/06	4	4	1b	1b	2	2	1	2	2
15/06	1	5	2a	1b	1	1	2	3	2
20/06	4	6	1b	1b	1	1	2	3	1
21/06	1	7	1b	1b	2	2	1	2	3
23/06	1	6	1b	1b	2	2	1	2	1
26/06	5	8	1b	1b	2	2	1	1	2
30/06	1	5	2a	1b	1	1	2	3	2
04/07	1	3	3b	1b	1	1	2	2	2

Tableau B1.8: (suite)

Event	M1	M2	M3	M4	M5	M6	M7	M8	M9
08/07	1	3	1b	1b	2	2	1	2	1
14/07	-	9	1b	1b	2	2	1	2	2
17/07	4	10	1b	1b	1	1	2	3	1
18/07	2	6	1b	1b	1	1	2	2	1
20/07	2	9	3b	2a	2	2	1	1	2
25/07	2	3	2a	1b	2	2	1	1	2
26/07	1	3	2a	1b	1	1	2	3	1
30/07	1	3	2a	1b	1	1	2	3	1
03/08	5	1	4a	2a	2	2	1	1	2
06/08	4	6	1b	1b	2	1	2	2	1
07/08	-	6	1b	1b	1	1	2	3	1
08/08	2	6	2a	2a	2	2	1	1	1
10/08	2	8	3b	1b	1	1	2	2	2
14/08	6	6	2a	1b	1	1	2	3	1
17/08	-	5	1b	2a	2	2	1	2	2
18/08	1	6	2a	1b	2	2	1	2	1
20/08	2	3	1b	2a	3	2	1	1	2
23/08	-	3	2b	1b	1	1	2	3	1
26/08	1	9	1b	1b	2	2	1	1	2
29/08	2	3	1b	1b	2	2	1	2	2
31/08	2	6	1b	1b	2	2	1	2	1
02/09	2	3	1b	1b	1	2	1	2	2
03/09	2	3	1b	2a	2	2	1	1	2
04/10	2	7	2a	1b	2	2	1	1	3

Tableau B1.9: Groupes des événements pluvieux pour les neuf méthodes de classification étudiées pour l'année 1992. (M1 = CDA, M2 = CMRCA; M3 = CA; M4 = SCA; M5 = CUPAA; M6 = CCVA; M7 = CCVUPAA; M8 = DUPA et M9 = CMRCA + CUPAA)

Pluies	M1	M2	M4	M5	M6	M7	M8	M9
10/04a	2	2	2a	5	1	2	1	6
10/04b	2	2	2b	2	3	1	3	1
11/05	2	1	1b	2	4	1	3	2
25/05	2	1	1b	1	2	1	2	3
28/05	2	2	1b	1	2	1	2	1
01/06	1	3	1b	2	2	1	3	2
03/06	2	2	1b	2	2	1	3	1
04/06	2	2	1b	2	1	1	3	1
06/06	2	3	1b	2	2	1	3	2
12/06	2	1	1b	2	2	1	3	2
20/06	2	1	1b	1	2	1	2	1
23/06	3	2	3a	3	1	2	2	3
25/06	2	2	1b	1	3	1	3	1
30/06	2	4	2a	5	1	2	1	7
07/07	2	2	1b	2	2	1	3	1
11/07	4	2	1b	1	1	1	2	1
13/07	2	2	1b	2	2	1	3	1
18/07	4	1	2b	4	1	2	2	3
20/07	2	2	2b	2	2	1	3	1
22/07	2	1	1b	4	1	2	2	3
27/07	5	1	1b	2	2	1	3	1
28/07	2	1	1b	2	5	1	3	1

Tableau B1.9: (suite)

Pluies	M1	M2	M4	M5	M6	M7	M8	M9
31/07	2	1	3b	4	1	2	2	3
04/08	5	1	1b	2	5	1	3	1
05/08	2	1	2b	2	6	1	3	1
08/08	2	1	2a	1	1	1	2	3
10/08	4	1	2a	4	1	2	2	3
11/08	2	2	1b	1	1	1	2	1
13/08	4	2	1b	4	2	1	2	1
15/08	2	1	1b	1	2	1	3	3
18/08	2	2	1b	1	2	1	2	1
21/08	2	1	2a	6	7	3	1	4
22/08	2	2	1b	1	2	1	2	1
24/0/	2	2	1b	2	2	1	3	1
25/08a	2	2	1b	2	2	1	3	1
25/08b	2	2	1b	1	2	1	2	1
25/08c	2	2	1b	2	2	1	3	1
27/08	2	1	1b	2	2	1	3	1
28/08a	2	2	1b	1	1	1	2	1
28/08b	2	1	1b	1	2	1	2	3
29/08	2	2	1b	2	4	1	3	1
30/08a	2	1	2a	7	7	3	1	5
30/08b	2	1	1b	4	1	1	2	3
02/09	3	1	1b	2	4	1	3	1
06/09	2	1	2a	8	1	2	1	3
11/09	2	1	2b	5	1	2	1	3
14/09	4	2	1b	4	1	2	2	1
15/09a	2	2	1b	1	2	1	2	1
15/09b	2	2	1b	2	2	1	3	1

Tableau B1.10: Groupe (UPA) et statistiques des événements pluvieux de 1990.

(Moyenne, Moy; Ecart-type, Ect; Probabilité de pluie non nulle, FO).

Pluies	Moy (mm)	Ect (mm)	CV	FO (%)	Groupe
23/05	4.0 (4.7)	7.5 (9.9)	1.85	51.4	3
28/05	24.3 (22.8)	20.1 (17.7)	0.83	98.6	2
07/06	7.9 (4.2)	13.2 (7.3)	1.67	70.0	3
12/06	5.0 (4.5)	6.6 (6.6)	1.32	77.5	2
17/06	7.8 (8.0)	11.2 (13.3)	1.43	56.8	2
22/06	3.8 (2.7)	5.4 (6.1)	1.43	66.7	3
24/06	27.2 (27.9)	18.1 (19.1)	0.66	94.0	1
27/06	5.2 (4.6)	10.3 (11.6)	1.96	45.9	3
29/06	3.2 (3.4)	6.3 (6.3)	1.96	38.8	3
03/07	7.8 (10.5)	10.9 (14.8)	1.39	83.5	3
05/07	5.1 (4.7)	11.8 (11.5)	2.31	36.5	3
08/07	5.5 (2.4)	11.7 (6.1)	2.15	48.8	3
09/07	1.5 (1.5)	3.8 (3.2)	2.58	40.7	3
12/07	17.4 (15.0)	14.4 (13.3)	0.83	94.1	2
15/07	2.9 (1.2)	8.3 (3.2)	2.84	42.4	3
15/07	4.9 (5.3)	5.6 (6.0)	1.13	86.9	2
18/07	8.6 (11.5)	9.2 (11.5)	1.07	84.0	2
21/07	7.7 (4.8)	10.7 (8.7)	1.39	60.2	2
21/07	2.7 (3.0)	4.3 (4.4)	1.56	49.4	3
22/07	4.4 (2.9)	6.0 (4.3)	1.37	73.5	2

Tableau B1.10: (suite)

Pluies	Moy (mm)	Ect (mm)	CV	FO (%)	Groupe
23/07	18.6 (15.0)	14.0 (12.4)	0.75	95.1	1
27/07	23.7 (20.6)	12.3 (11.3)	0.52	96.4	1
31/07	12.8 (10.9)	12.7 (11.3)	0.99	83.3	2
02/08	4.7 (5.0)	8.4 (8.3)	1.81	49.4	3
04/08	26.6 (24.7)	12.7 (14.7)	0.48	97.6	1
08/08	20.4 (18.8)	11.1 (11.6)	0.54	98.9	1
14/08	1.0 (0.9)	1.9 (2.0)	1.96	32.6	3
17/08	16.7 (16.6)	17.7 (15.6)	1.04	83.5	2
24/08	11.0 (10.1)	13.0 (12.4)	1.18	83.9	2
28/08	3.8 (2.9)	6.0 (4.4)	1.57	62.5	3
29/08	3.9 (3.7)	5.7 (5.1)	1.46	83.9	3
01/09	17.2 (19.5)	17.3 (14.9)	1.01	94.4	2
05/09	7.3 (9.5)	10.4 (12.7)	1.43	91.2	3
07/09	13.0 (10.9)	17.6 (16.0)	1.36	78.0	3
11/09	2.5 (2.9)	6.9 (6.2)	2.79	36.3	3
13/09	15.6 (16.7)	8.4 (9.3)	0.54	95.6	1
17/09	19.5 (13.5)	18.8 (14.9)	0.96	95.6	2

Les valeurs entre parenthèses représentent les statistiques pour un réseau décorrélé.

Tableau B1.11: Groupe (UPA) et statistiques des événements pluvieux de 1991.

(Moyenne, Moy; Ecart-type, Ect; Probabilité de pluie non nulle, FO).

Pluies	Moy (mm)	Ect (mm)	CV	FO (%)	Groupe
14/04	4.6 (4.4)	10.2 (10.8)	2.21	33.3	3
28/04	8.0 (10.4)	8.3 (7.3)	1.04	88.2	2
04/05	28.5 (32.2)	20.7 (19.9)	0.73	97.7	1
16/05	4.6 (2.2)	10.0 (3.2)	2.19	59.0	3
20/05	5.1 (4.4)	6.0 (5.8)	1.18	86.8	2
24/05	28.0 (23.1)	29.2 (25.5)	1.04	79.7	2
25/05a	2.1 (1.6)	3.4 (3.2)	1.68	42.1	3
25/05b	3.1 (4.1)	10.2 (11.7)	3.28	44.7	3
26/05a	10.3 (10.1)	9.9 (11.1)	0.97	96.1	2
26/05b	29.9 (27.6)	29.6 (25.7)	0.99	98.7	2
27/05	1.5 (0.4)	3.1 (1.5)	2.11	42.1	3
28/05	3.3 (1.4)	5.0 (1.9)	1.51	71.5	3
01/06	6.2 (6.4)	7.3 (8.8)	1.17	66.7	2
03/06	14.0 (14.6)	8.1 (7.3)	0.58	97.6	1
07/06	7.9 (5.4)	7.4 (6.4)	0.94	83.0	2
11/06	12.0 (11.5)	11.4 (12.2)	0.95	82.2	2
15/06	9.2 (6.2)	11.8 (9.8)	1.28	81.3	3
20/06	4.6 (5.8)	7.4 (7.9)	1.63	58.7	3
21/06	1.7 (1.2)	1.9 (1.6)	1.07	72.8	2
23/06	2.5 (2.7)	2.4 (2.7)	0.96	90.2	2
26/06	27.4 (27.7)	18.1 (21.3)	0.66	95.7	1

130101

Tableau B1.11: (suite)

Pluies	Moy (mm)	Ect (mm)	CV	FO (%)	Groupe
30/06	7.1 (4.4)	13.4 (7.8)	1.91	53.3	3
04/07	6.6 (5.3)	8.2 (6.7)	1.24	80.2	2
08/07	12.8 (12.6)	12.6 (12.9)	0.99	82.2	2
14/07	18.1 (19.4)	16.6 (16.7)	0.92	83.1	2
17/07	8.6 (9.5)	12.0 (13.8)	1.39	76.4	3
18/07	1.0 (0.9)	1.6 (1.9)	1.59	52.8	2
20/07	19.0 (21.3)	11.9 (15.2)	0.62	98.9	1
25/07	14.2 (13.8)	11.7 (12.2)	0.82	97.7	1
26/07	3.0 (2.2)	5.5 (5.3)	1.88	46.0	3
30/07	5.4 (7.1)	8.9 (10.5)	1.64	68.6	3
03/08	37.1 (32.0)	20.1 (18.9)	0.54	100	1
06/08	2.7 (2.4)	3.6 (4.5)	1.36	67.1	2
07/08	4.1 (5.8)	6.7 (8.6)	1.61	60.0	3
08/08	14.8 (15.2)	8.3 (7.3)	0.56	98.8	1
10/08	12.3 (13.6)	16.3 (19.0)	1.32	90.6	2
14/08	2.3 (2.4)	5.0 (5.5)	2.14	31.0	3
17/08	19.0 (20.2)	10.5 (10.9)	0.55	98.8	2
18/08	8.3 (9.4)	7.1 (7.4)	0.86	91.8	2
20/08	30.0 (31.4)	8.1 (8.5)	0.27	100	1
23/08	5.0 (6.0)	9.3 (10.7)	1.86	67.8	3
26/08	12.9 (16.0)	10.6 (9.8)	0.82	96.6	1
29/08	14.8 (14.5)	11.5 (11.2)	0.78	100	2
31/08	11.0 (7.2)	9.5 (5.4)	0.86	96.7	2
02/09	11.8 (8.6)	10.2 (8.8)	0.86	98.9	2
03/09	21.9 (22.9)	10.7 (10.2)	0.49	97.3	1
04/10	3.4 (3.0)	2.9 (2.3)	0.85	83.6	1

Les valeurs entre parenthèses représentent les statistiques pour un réseau décorrélé.

Tableau B1.12: Groupe (UPA) et statistiques des événements pluvieux de 199².

(Moyenne, Moy; Ecart type, Ect; Probabilité de pluie non nulle, FO).

Pluies	Moy (mm)	Ect (mm)	CV	FO	Groupe
10/04a	0.9 (1.4)	1.6 (1.7)	1.67	37.5	1
10/04b	3.8 (3.8)	7.7 (6.8)	2.05	54.4	3
11/05	5.0 (8.5)	10.5 (13.9)	2.10	46.4	3
25/05	5.5 (5.6)	10.6 (9.6)	1.93	41.6	2
28/05	2.4 (3.6)	5.8 (7.8)	2.39	32.0	2
01/06	10.9 (11.3)	14.0 (11.1)	1.28	81.3	3
03/06	1.2 (1.7)	2.5 (3.3)	2.15	39.2	3
04/06	11.8 (10.3)	9.6 (10.6)	0.82	90.7	3
06/06	10.7 (8.9)	12.4 (9.5)	1.16	94.9	3
12/06	6.5 (7.5)	8.7 (10.3)	1.34	75.0	3
20/06	6.6 (5.7)	9.1 (7.3)	1.37	66.7	2
23/06	4.0 (4.8)	6.9 (7.5)	1.72	37.6	2
25/06	1.8 (1.1)	3.8 (1.6)	2.07	54.5	3
30/06	22.1 (20.0)	14.8 (15.1)	0.67	95.2	1
07/07	4.2 (3.9)	4.7 (4.7)	1.14	87.5	3
11/07	7.2 (6.5)	7.1 (6.7)	1.00	79.6	2
13/07	10.9 (11.6)	12.1 (14.1)	1.11	86.3	3
18/07	28.2 (27.5)	17.1 (17.4)	0.61	98.1	2
20/07	2.3 (2.5)	5.9 (6.5)	2.50	30.5	3
22/07	30.7 (30.1)	16.5 (16.1)	0.54	100	2
27/07	9.3 (7.7)	9.1 (7.2)	0.97	96.1	3
28/07	3.0 (3.6)	7.7 (9.9)	2.54	43.1	3
31/07	34.7 (29.4)	18.0 (16.8)	0.52	99.0	2
04/08	3.5 (3.8)	9.1 (7.6)	2.64	42.0	3
05/08	2.6 (2.8)	7.4 (6.2)	2.81	40.4	3

Tableau B1.12: (suite)

Pluies	Moy (mm)	Ect (mm)	CV	FO (%)	Groupe
08/08	22.2 (25.4)	17.2 (16.5)	0.78	98.0	2
10/08	25.4 (29.7)	15.8 (14.6)	0.62	100	2
11/08	7.3 (6.9)	6.5 (6.9)	0.90	89.3	2
13/08	3.6 (1.7)	7.3 (4.0)	2.04	33.3	2
15/08	13.4 (14.5)	16.1 (21.3)	1.20	80.2	3
18/08	2.3 (2.7)	3.9 (4.4)	1.74	48.6	2
21/08	46.6 (41.5)	15.4 (14.4)	0.33	100	1
22/08	2.0 (2.5)	3.8 (4.1)	1.87	41.3	2
24/08	3.3 (5.0)	4.9 (6.3)	1.48	64.4	3
25/08a	4.3 (5.0)	5.6 (6.3)	1.31	76.9	3
25/08b	1.4 (0.6)	2.8 (1.5)	2.08	32.4	2
25/08c	9.4 (10.1)	11.8 (13.7)	1.26	84.8	3
27/08	8.4 (9.9)	8.4 (8.4)	1.01	93.3	3
28/08a	1.2 (1.0)	1.5 (1.6)	1.26	62.5	2
28/08b	12.7 (10.6)	11.4 (9.2)	0.90	96.2	2
29/08	6.7 (5.6)	9.4 (11.2)	1.41	80.0	3
30/08a	26.1 (26.9)	6.5 (6.3)	0.25	100	1
30/08b	12.3 (10.2)	11.4 (12.0)	0.93	83.8	2
02/09	5.7 (8.2)	11.1 (13.6)	1.94	53.3	3
06/09	19.9 (18.6)	12.2 (12.3)	0.61	96.2	1
11/09	18.3 (15.7)	15.9 (15.8)	0.87	83.8	1
14/09	14.7 (14.0)	9.3 (8.7)	0.63	99.1	2
15/09a	1.0 (0.9)	1.8 (1.9)	1.89	46.7	2
15/09b	1.2 (1.7)	2.7 (2.4)	2.20	37.7	3

Les valeurs entre parenthèses représentent les statistiques pour un réseau décorrélé.

Tableau B1.13: Paramètres des modèles de seuillage aux seuils optimum pour les groupes des méthodes de classification CMRCA, SCA, DUPA et CMRCA + DUPA (en colonne on a pour un groupe donné, le seuil optimum, le coefficient de détermination R^2 , l'ordonnée à l'origine A et la pente B) pour les échantillons, 1990, 1991, 1992, 1990+ 1991 et 1990+1991+1992.

	Total	CMRCA		SCA		DUPA			CMRCA+UP A	
		1	2	1b	2a,b	1	2	3	1	2
1990	17.5 0.97 2.07 0.344	10 0.97 -0.2 0.300	17.5 0.95 2.53 0.344	17.5 0.98 1.6 0.378	17.5 0.99 6.6 0.248	20 0.92 8.2 0.264	20 0.99 3.1 0.406	15 0.94 1.3 0.398	15 0.95 1.6 0.301	20 0.90 8.3 0.268
1991	17.5 0.92 2.49 0.363	7.5 0.97 0.9 0.161	17.5 0.94 3.12 0.308	17.5 0.94 1.6 0.419	30 0.98 12.4 0.398	30 0.95 8.5 0.483	20 0.96 2.7 0.464	10 0.90 0.7 0.244	7.5 0.95 1.1 0.164	20 0.88 4.6 0.370
1992	20 0.95 2.61 0.408	10 0.97 0.7 0.230	20 0.94 2.81 0.405	17.5 0.95 2.1 0.375	20 0.92 1.7 0.422	25 0.97 1.0 0.475	17.5 0.97 1.5 0.405	15 0.96 1.6 0.370	10 0.90 1.2 0.230	20 0.91 2.1 0.417
1990-91	17.5 0.94 0.355 2.28	15 0.94 1.90 0.304	20 0.93 2.70 0.417	17.5 0.95 1.68 0.398	30 0.95 12.9 0.432	30 0.94 10.4 0.430	20 0.96 2.8 0.441	15 0.88 1.2 0.383	15 0.92 1.7 0.308	20 0.88 5.1 0.349
1990-91- 92	17.5 0.94 2.1 0.368	12.5 0.94 1.5 0.253	20 0.93 3.5 0.377	17.5 0.94 1.8 0.393	30 0.88 10.4 0.436	30 0.91 9.2 0.454	20 0.96 2.6 0.436	15 0.91 1.3 0.361	15 0.91 1.9 0.300	20 0.92 4.2 0.390

ANNEXE B2

Variogramme d'un champ intermittent

ANNEXE B2

**LIEN ENTRE LES VARIOGRAMMES CONDITIONNEL ET NON CONDITIONNEL
D'UN PROCESSUS INTERMITTENT**

Soit γ_I le variogramme du processus intermittent $R(x,y,t)$, γ le variogramme conditionnel coorespondant ($R(x,y,t) > 0$) et γ_c le variogramme croisé entre la zone pluvieuse et la zone non pluvieuse. Pour une classe de distance h , ces variogrammes sont donnés par

$$\gamma_I(h) = \frac{1}{2N_T(h)} \sum_{i=1}^{N_T(h)} (R(i) - R(j))^2 \quad (1)$$

où $N_T(h)$ est le nombre total de couples de points (i,j) appartenant à la classe de distance h ,

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} (R(i) - R(j))^2 \quad (2)$$

où $N(h)$ est le nombre total de couples de points (i,j) appartenant à la classe de distance h avec $R(i) > 0$ et $R(j) > 0$

$$\gamma_c(h) = \frac{1}{2N_c(h)} \sum_{i=1}^{N_c(h)} (R(i) - R(j))^2 \quad (3)$$

où $N_c(h)$ est le nombre total de couples de points (i,j) appartenant à la classe de distance h avec $R(i) = 0$ et $R(j) > 0$ ou $R(j) = 0$ et $R(i) > 0$.

Notons $N_0(h)$ étant le nombre de couples de points (i, j) appartenant à la classe de distant h avec $R(i)=0$ et $R(j)=0$, ainsi on a

$$N_T(h) = N(h) + N_c(h) + N_0(h) \quad (4)$$

et

$$\gamma_I(h) = \frac{1}{2N_T(h)} \left[\sum_{\substack{i=1 \\ N_0(h)}}^{N_T(h)} (R(i) - R(j))^2 + \sum_{i=1}^{N_c(h)} (R(i) - R(j))^2 + \sum_{i=1}^{N_0(h)} (R(i) - R(j))^2 \right] \quad (5)$$

Le dernier terme dans cette somme est égal à zéro, ainsi l'équation (5) se ramène à

$$\gamma_I(h) = \frac{N(h) \gamma(h) + N_c(h)}{N_T(h)} \quad (6)$$

ou

$$\gamma_I(h) = \alpha(h) \gamma(h) + \beta(h) \gamma_c(h) \quad (7)$$

avec

$$\alpha(h) = \frac{N(h)}{N_T(h)} \quad \beta(h) = \frac{N_c(h)}{N(h) + N_c(h) + N_0(h)} \quad (8)$$

Les paramètres $\alpha(h)$ et $\beta(h)$ sont fonction seulement du degré d'intermittence du process $R(x, y, t)$. Dans le cas d'un processus continu, $\alpha(h)$ et $\beta(h)$ prennent respectivement les

valeurs de 1 et 0. Dans le cas où le variogramme $\gamma(h)$ possède σ_r comme palier, et en considérant que $\alpha(h)$ et $\beta(h)$ deviennent indépendant de h au delà de la portée du variogramme, l'équation (7) devient

$$\gamma_I(h) = \alpha\sigma_r + \beta\gamma_c(h) \quad (9)$$

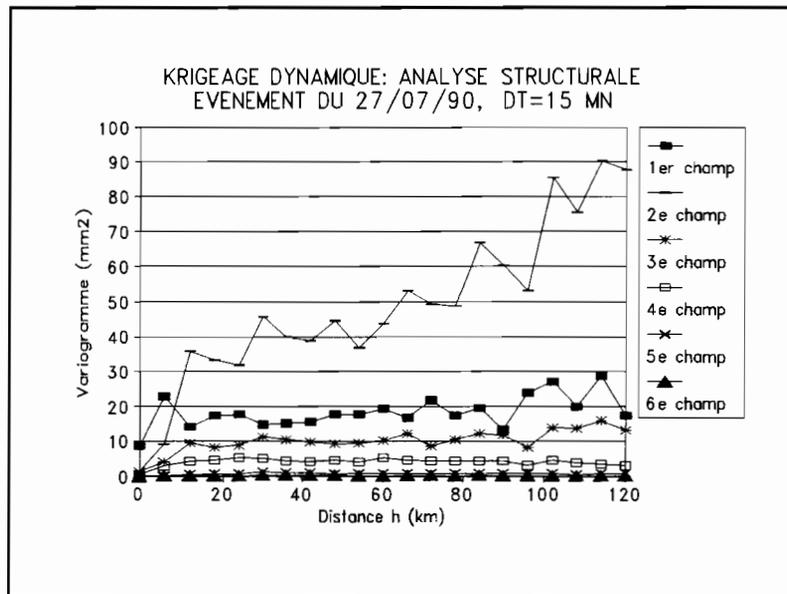
Ceci veut dire que le comportement du variogramme γ_I au delà de la portée r est entièrement déterminé par celui du variogramme croisé γ_c .

Ainsi pour des champs continus ou ayant une intermittence négligeable ($N_c(h)$ est petit par rapport à $N_T(h)$ pour toutes les distances) γ_I est égal à γ et l'interpolation de champ en considérant γ ne cause aucun problème. Par contre dès que $N_c(h)$ devient de plus en plus grand, γ_I est la composition de deux variogrammes. La modélisation de γ_I par les modèles connus de variogramme n'est pas très évidente, ainsi d'autres méthodes d'estimation autres que celles reliées au variogramme sont à considérées dans ce cas.

ANNEXE B3:

Résultats complémentaires au krigeage dynamique

(a)



(b)

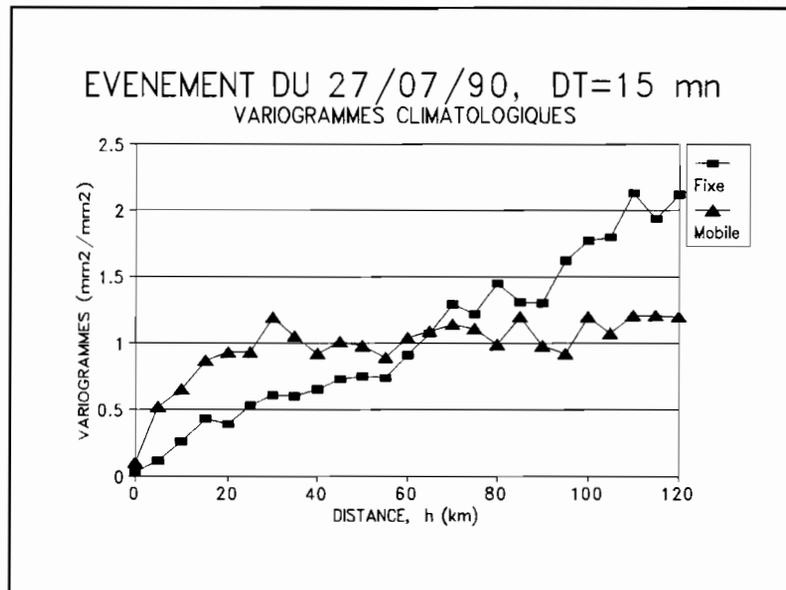


Figure B3.1: Variogrammes des champs pour l'événement du 27/07/90, a) repère mobile; b) variogrammes climatologiques

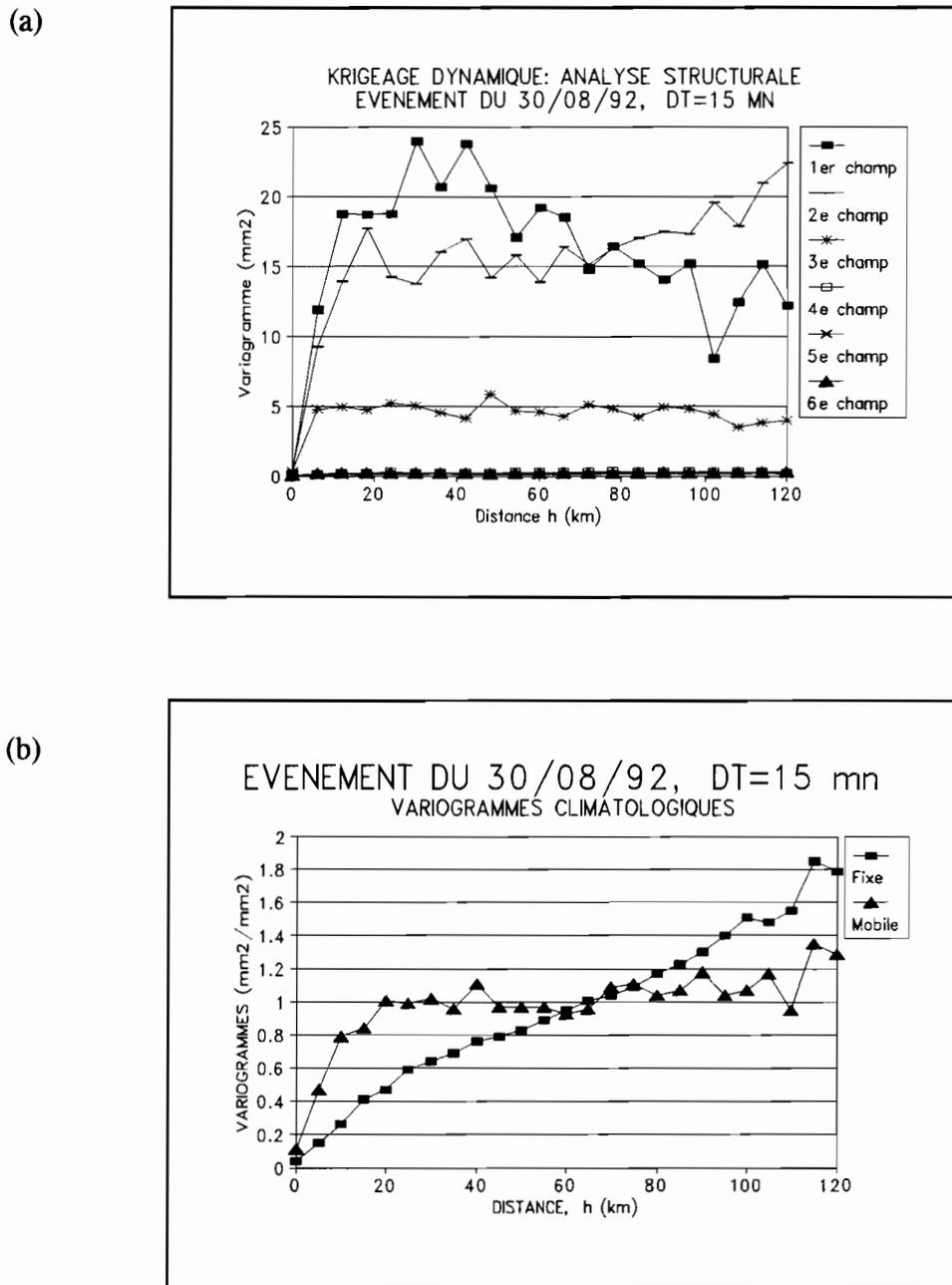
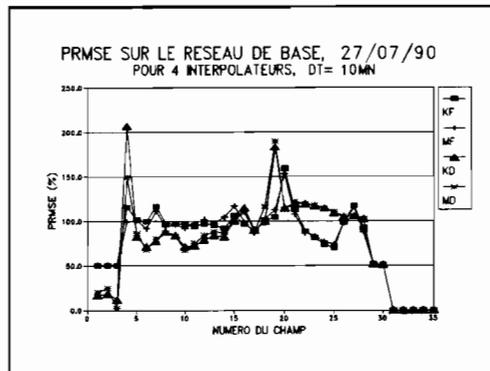
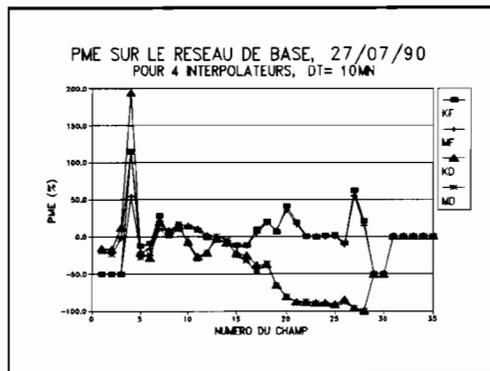


Figure B3.2: Variogrammes des champs pour l'événement du 30/08/92, a) repère mobile; b) variogrammes climatologiques

(a)



(b)



(c)

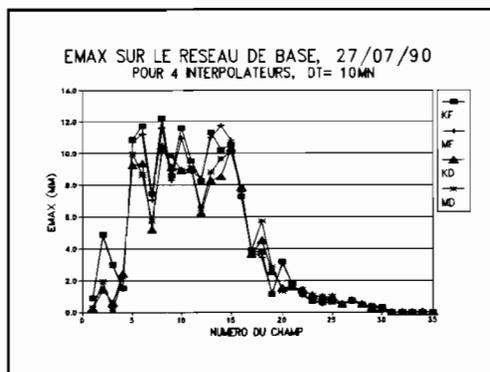
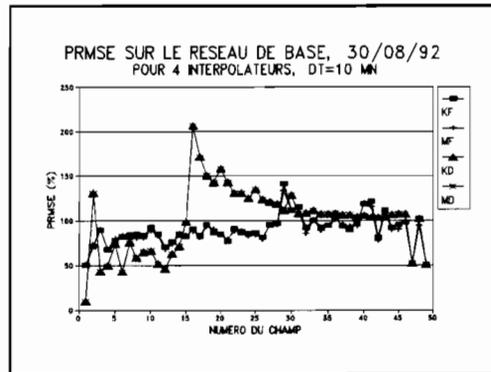
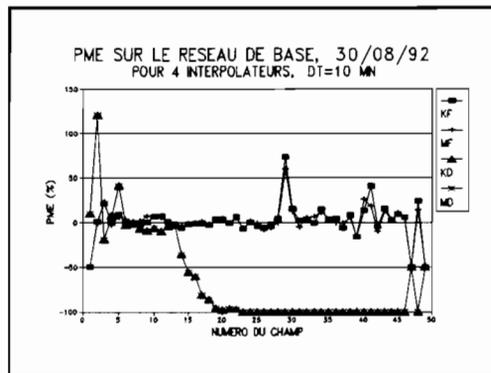


Figure B3.3: Erreurs de validation ponctuelle pour l'événement du 27/07/90 sur le réseau de base, a) PRMSE
b) PME et c) Emax

(a)



(b)



(c)

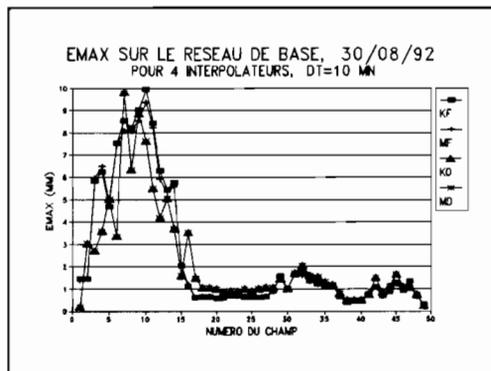


Figure B3.4: Erreurs de validation ponctuelle pour l'événement du 30/08/92 sur le réseau de base, a) PRMSE
b) PME et c) Emax

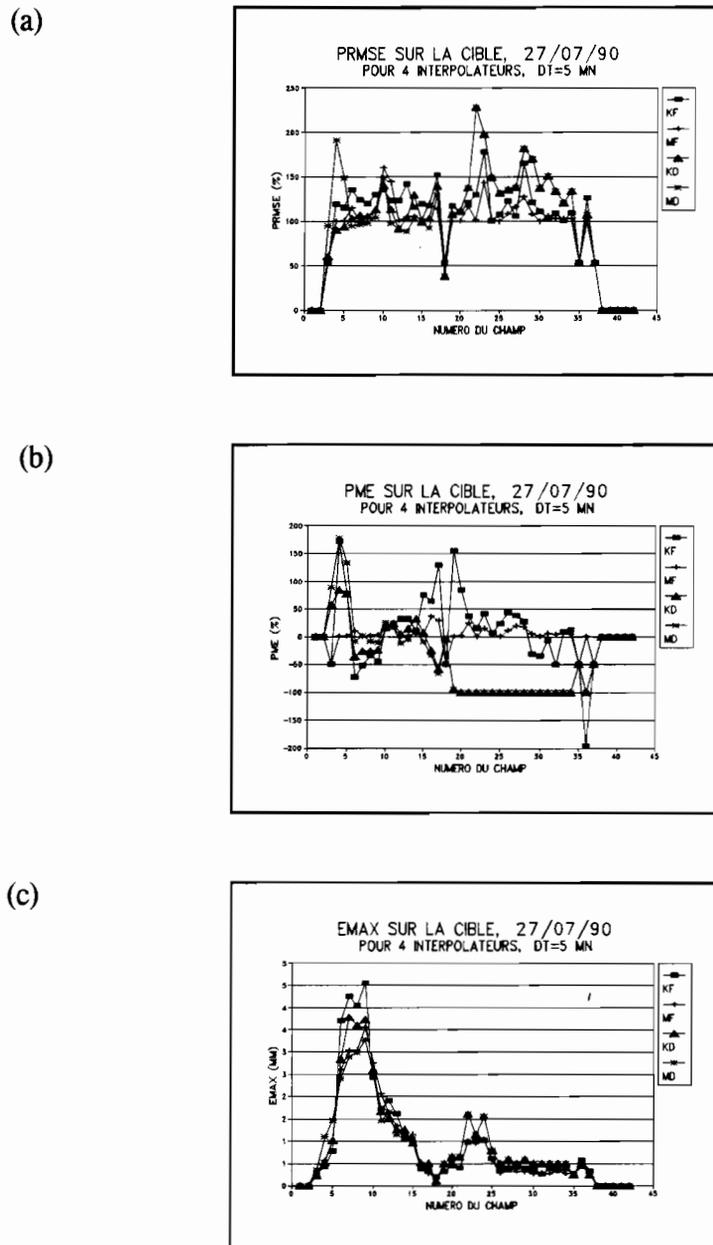
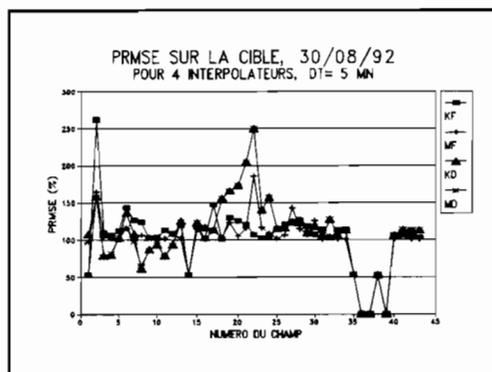
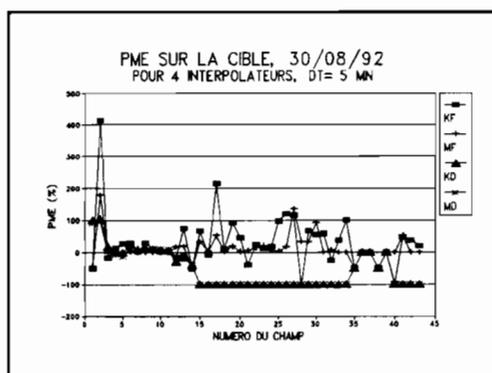


Figure B3.5: Erreurs de validation ponctuelle pour l'événement du 27/07/90 sur la cible a) PRMSE, b) PME et c) Emax

(a)



(b)



(c)

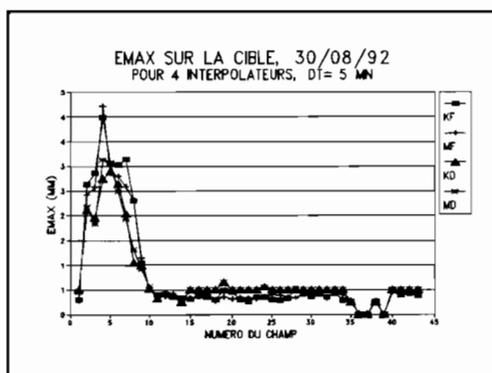
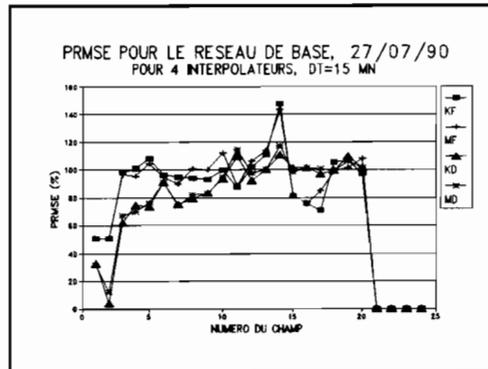
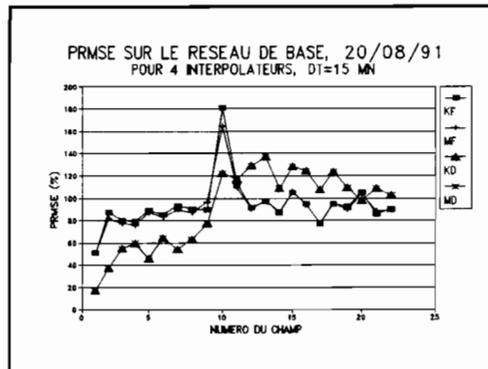


Figure B3.6: Erreurs de validation ponctuelle pour l'événement du 27/07/90 sur la cible a) PRMSE, b) PME et c) Emax

(a)



(b)



(c)

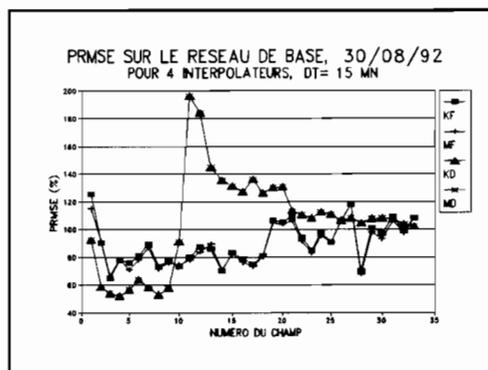
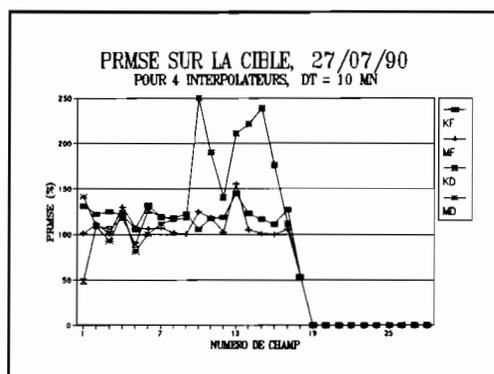
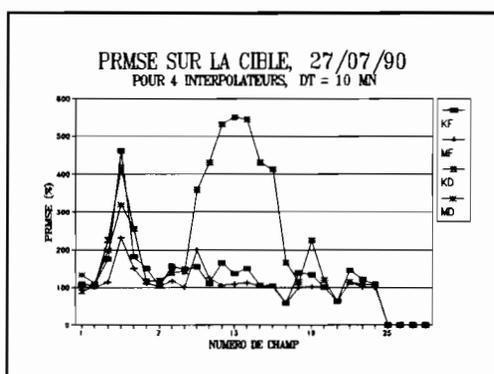


Figure B3.7: Erreurs de validation ponctuelle (PRMSE) sur le réseau de base pour le pas de temps de cumul de 15 min a) 27/07/90 b) 20/08/91 et c) 30/08/92

(a)



(b)



(c)

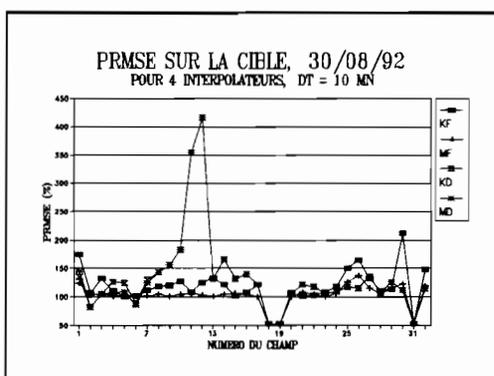


Figure B3.8: Erreurs de validation ponctuelle (PRMSE) sur la cible, $\Delta t = 10$ min

a) 27/07/90, b) 20/08/91 et c) 30/08/92.

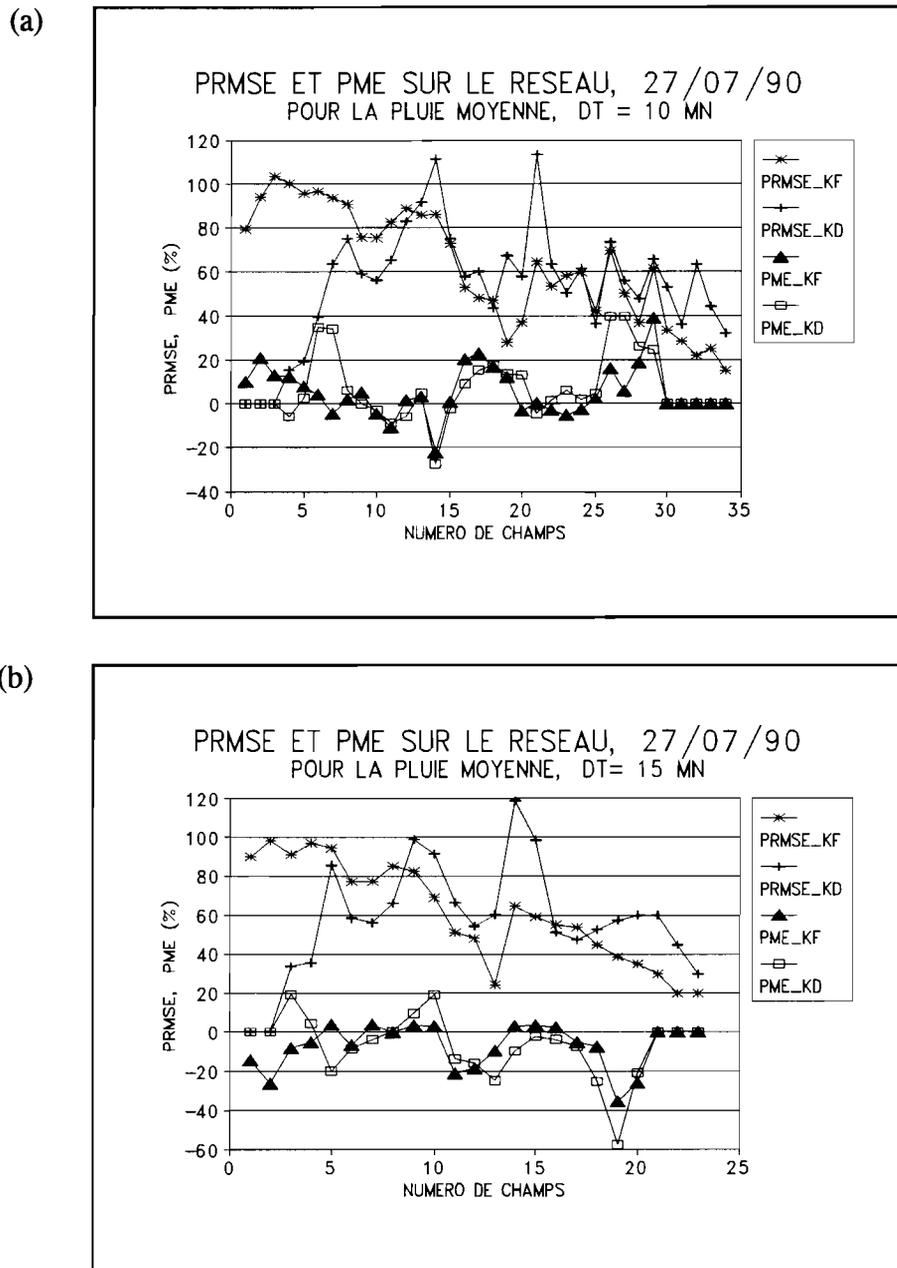


Figure B3.9: Erreurs de validation (PRMSE et PME) pour la pluie moyenne pour l'événement du 27/07/90, a) $\Delta t = 10$ min et b) $\Delta t = 15$ min.

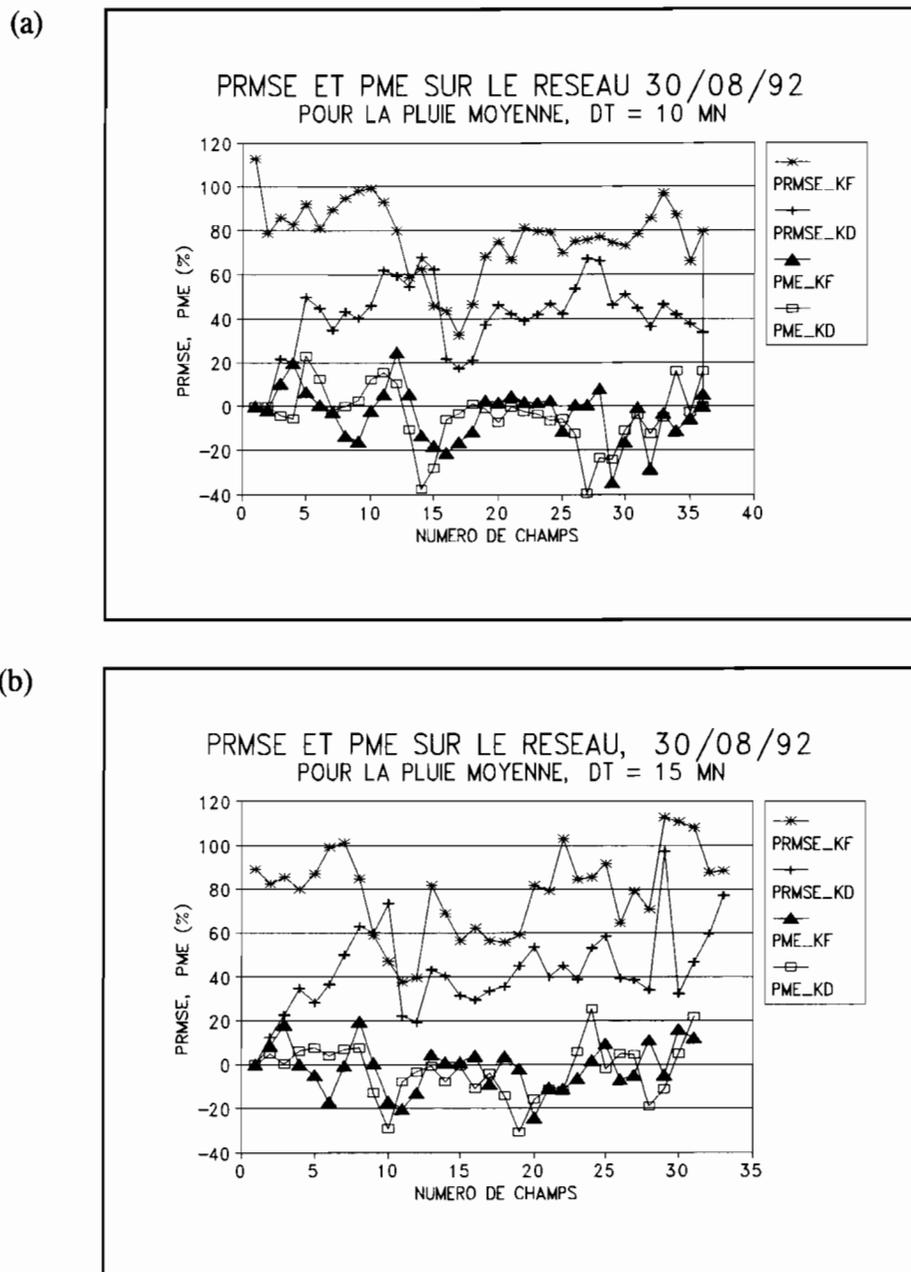
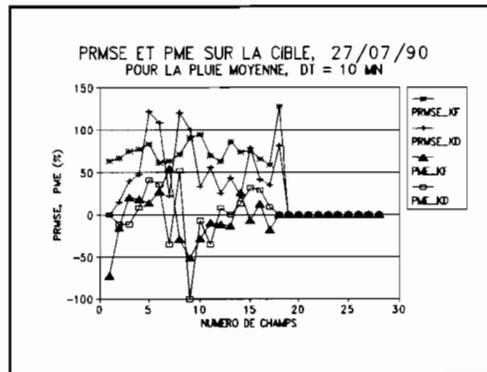
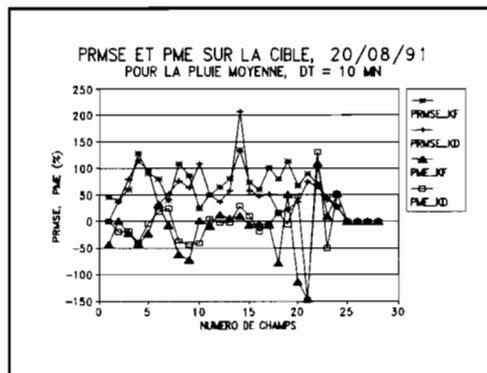


Figure B3.10: Erreurs de validation (PRMSE et PME) pour la pluie moyenne pour l'événement du 32/08/92, a) $\Delta t = 10$ min et b) $\Delta t = 15$ min.

(a)



(b)



(c)

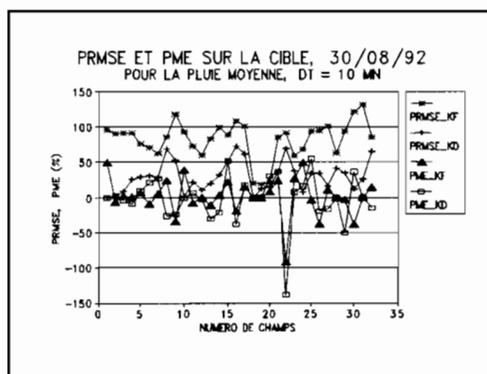


Figure B3.11: Erreurs de validation (PRMSE et PME) pour la pluie moyenne sur la cible $\Delta t = 10$ min a) 27/07/90, b) 20/08/91 et c) 30/08/92.

ANNEXE B4

Estimation des vitesses d'approche par krigeage dual

ANNEXE B4

DYNAMIQUE DE SYSTÈME PLUVIEUX: ESTIMATION DES VITESSES D'APPROCHE PAR KRIGEAGE DUAL

1 INTRODUCTION

La dynamique de la pluie a été jusqu'à maintenant considérée dans la plupart des études comme étant caractérisée par une vitesse constante. Cette hypothèse n'est valide que pour des événements pluvieux sensiblement idéaux. Des méthodes d'estimation de la vitesse moyenne de l'orage à partir des données pluviographiques ont ainsi été développées (Diskin, 1990; 1987; Johnson et Bras, 1979; Niemczynbicz et Dahlblom, 1984). Mis à part la méthode de Marshall qui est basée sur l'analyse de la fonction de corrélation spatio-temporelle des champs de pluie, les autres méthodes sont caractérisées par l'analyse spatiale des temps d'observation d'une des caractéristiques de l'hétérogramme de pluie (début de la pluie, milieu de la pluie, ..etc). De part même la nature complexe du phénomène de pluie, l'hypothèse d'une vitesse de déplacement constante sur une grande distance est irréaliste et ceci s'observe bien sur les données radar. L'orage a généralement une vitesse d'approche différente d'un point à l'autre. Pour le modèle du krigeage dynamique se basant sur les deux repères: mobile et fixe,

seule l'hypothèse d'une vitesse d'approche variable permet, pour des événements pluvieux réels, de passer d'un repère à un autre sans perdre d'informations positionnelles. Une technique d'estimation de vitesse d'approche en chacun des points du réseau de mesure est ici proposée. Tout comme les techniques d'estimation de la vitesse moyenne, cette nouvelle méthode est aussi basée sur l'analyse spatiale du temps d'observation d'une des caractéristiques de l'hyétogramme. La technique du krigeage sous sa forme duale est considérée.

2 ESTIMATION DE LA VITESSE DE L'ORAGE

Si $t(x,y)$ est le temps d'observation d'une des caractéristiques de l'hyétogramme de pluie, les composantes $V_x(x,y)$ et $V_y(x,y)$ de la vitesse au point de coordonnées (x,y) sont données par

$$\frac{1}{V_x(x,y)} = \frac{\partial t(x,y)}{\partial x} \quad (1)$$

$$\frac{1}{V_y(x,y)} = \frac{\partial t(x,y)}{\partial y} \quad (2)$$

Ainsi la détermination des composantes du vecteur vitesse passe par la modélisation de la fonction $t(x,y)$ caractérisant le temps d'arrivée d'une caractéristique de l'hyétogramme. Cette modélisation peut se faire par les techniques de regression (cas de la vitesse constante) ou par krigeage. Au

niveau des méthodes d'estimation de la vitesse, considérées jusqu'à date, on suppose que $t(x,y) = a_0 + a_1x + a_2y$ qui est l'équation d'un plan. Ainsi à partir des équations (1) et (2), on voit que dans ce cas, la connaissance des constantes a_1 et a_2 permet de déterminer la vitesse moyenne de l'orage. Ces valeurs peuvent être obtenues, par la méthode de regression (Diskin, 1990; 1987) ou par la méthode combinatoire de trois points (connaissant trois points, on resout, un système de trois équations à trois inconnues; on fait la moyenne par la suite). Dans le cas du krigage, c'est l'espérance de $t(x,y)$ et non $t(x,y)$ qui sera donnée par $a_0 + a_1x + a_2y$. Sous la forme duale du krigage, $t(x,y)$ est donnée par

$$t(x,y) = m(x,y) + \sum_{j=1}^N b_j K(|X - X_j|) \quad (3)$$

avec $m(x,y) = E[t(x,y)]$ étant la dérive de $t(x,y)$, $K(|X - Y|)$ la covariance entre les points X et Y ($|X - Y|$ est la distance euclidienne entre les deux points X et Y) et b_j ($j=1, \dots, N$) étant les coefficients du krigage dual. N est le nombre de stations de mesure. Comme $m(x,y) = a_0 + a_1x + a_2y$, $t(x,y)$ est donné par

$$t(x,y) = a_0 + a_1x + a_2y + \sum_{j=1}^N b_j K(|X - X_j|) \quad (4)$$

où a_0 , a_1 , a_2 et b_j ($j=1, \dots, N$) sont la solution du système du krigeage dual suivant:

$$\sum_{j=1}^N b_j K(|X_i - X_j|) + a_0 + a_1 x_i + a_2 y_i = t_i \quad i=1, \dots, N \quad (5)$$

$$\sum_{j=1}^N b_j = 0$$

$$\sum_{j=1}^N b_j x_j = 0$$

$$\sum_{j=1}^N b_j y_j = 0$$

Ainsi les composantes du vecteur vitesse en (x, y) sont données par:

$$V_x(x, y) = \frac{1}{a_1 + \sum_{j=1}^N b_j \frac{\partial}{\partial x} K(|x - x_j|)} \quad (6)$$

$$V_y(x, y) = \frac{1}{a_2 + \sum_{j=1}^N b_j \frac{\partial}{\partial y} K(|x - x_j|)} \quad (7)$$

Si h est la distance euclidienne, les dérivées partielles de $K(X - X_i)$ par rapport à x et y sont données par

$$\frac{\partial K(h_i)}{\partial x} = \frac{x - x_i}{h_i} K'(h_i) \quad (8)$$

$$\frac{\partial K(h_i)}{\partial y} = \frac{y-y_i}{h_i} K'(h_i) \quad (9)$$

où $K'(h)$ est la dérivée de la fonction de covariance par rapport à h et h_i est donnée par

$$h_i = \sqrt{(x-x_i)^2 + (y-y_i)^2} \quad (10)$$

La mise en oeuvre de la méthode n'est possible que pour des $K(h)$ dérivables. La validation de cette nouvelle méthode peut se faire par le biais des données radar.

École Polytechnique de Montréal
C.P. 6079, Succ. Centre-ville
Montréal (Québec)
H3C 3A7