



HAL
open science

Space-time structure of heavy rainfall events: application to the Cévennes-Vivarais region.

Daide Ceresetti

► **To cite this version:**

Daide Ceresetti. Space-time structure of heavy rainfall events: application to the Cévennes-Vivarais region.. Ocean, Atmosphere. Université Joseph-Fourier - Grenoble I, 2011. English. NNT: . tel-00551316v1

HAL Id: tel-00551316

<https://theses.hal.science/tel-00551316v1>

Submitted on 3 Jan 2011 (v1), last revised 14 Apr 2011 (v2)

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

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

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : **Océan, Atmosphère, Hydrologie**

Arrêté ministériel : 7 août 2006

Présentée par

Davide CERESSETTI

Thèse dirigée par **Jean-Dominique CREUTIN** et
codirigée par **Gilles MOLINIÉ**

préparée au sein du **Laboratoire d'étude des Transferts en
Hydrologie et Environnement (LTHE, UMR 5564)**
dans l'**École Doctorale Terre-Univers-Environnement**

Structure spatio-temporelle des fortes précipitations: application à la région Cévennes-Vivarais

Thèse soutenue publiquement le **21 Janvier 2011**,
devant le jury composé de:

M. Paolo BURLANDO

Professor ETH Zürich, Rapporteur

Mme Efi FOUFOULA-GEORGIU

Professor Minnesota University, Rapporteur

M. Michel DÉQUÉ

Professeur CNRM Toulouse, Examineur

M. Christian LANTUÉJOUL

Professeur École de Mines, Paris, Examineur

M. François RENARD

Professeur Université de Grenoble, Examineur

M. Remko UIJLENHOET

Professor Wageningen University, Examineur

M. Jean-Dominique CREUTIN

Directeur de Recherche, Université de Grenoble, Directeur de Thèse

M. Gilles MOLINIÉ

Maître de Conférences, Université de Grenoble, Co-directeur de Thèse

M. Stéphane GIRARD

Chargé de Recherche, INRIA Rhône-Alpes, Membre Invité



Contents

I	Acknowledgments	v
II	Keywords	vi
III	Abstract	vii
IV	Introduction	ix

Part I Context

1	General context	1
1.1	Purpose of the thesis	1
1.1.1	Rainfall measurement uncertainties	2
1.1.2	Scale-invariance of heavy space-time rainfall	3
1.1.3	Scale-invariance of rainfall fields in space-time	3
1.1.4	Multi-scale evaluation of the impact of meteorological events	3
1.2	Geographical context	4
1.3	Meteo-hydrological context	4
1.3.1	Climatic features	5
1.4	Hydrological context	7
1.4.1	Ardèche basin	8
1.4.2	Cèze basin	8
1.4.3	Gard basin	9
1.4.4	Vidourle basin	9
1.5	Measurement network	9
1.5.1	Point-rainfall data	10
1.5.2	Radar-estimated rainfall fields	14
1.5.3	Disdrometer data	17
2	Uncertainties in the extreme rainfall measurement	19
2.1	Introduction	20
2.2	Role of the sampling frequency on the statistics of extremes	22
2.2.1	Moving-window and fixed-window sampling	22
2.2.2	Uncertainties associated to the measurement resolution	22
2.3	Tipping-bucket rain gauge measurement	26
2.4	Ground measurement network for spatial estimations: limits	42
2.5	Conclusion	43

Part II
Theoretical background

3	Statistics of extreme point-rainfall intensities	51
3.1	Introduction	52
3.2	Frequency and return period: definitions	52
3.3	Extreme Value Theory	53
3.4	The role of independence in Extreme value analysis	54
3.5	Block-maxima analysis	55
3.6	Analysis of exceedances: POT	59
3.6.1	GPD fitting on synthetic series	59
3.6.2	GPD fitting on real series	62
3.7	Comparison Block-Maxima - POT	63
3.8	Point process: an unified framework for extreme analysis	67
3.9	Stationarity of the rainfall series	68
3.10	Conclusion	69
4	Geostatistics	71
4.1	Introduction	71
4.2	Stationarity of random functions	72
4.3	Theoretical variogram	72
4.4	Sample variogram	73
4.4.1	Anisotropic variogram	75
4.4.2	Variogram of an intrinsic random function	75
4.4.3	Indicator variogram	76
4.4.4	Climatological variogram	79
4.4.5	Variogram models	79
4.5	Interpolation of point data	82
4.5.1	Kriging	83
4.5.2	Simple kriging	85
4.5.3	Ordinary kriging	86
4.5.4	Universal kriging	87
4.6	Conclusion	88
5	Scale invariance and self-similarity	91
5.1	Introduction	91
5.2	Scaling in geometry: fractals	92
5.3	Fractals in nature	93
5.4	The origin of scaling in nature: turbulence	95
5.5	Generalization of the fractal concept: Multifractals	97
5.5.1	Scaling of the Generalized Structure Function ζ	97
5.5.2	Moment scaling analysis	98
5.6	Towards an unified multifractal formalism	104
5.7	Multifractal spectrum estimation	104
5.7.1	Moment-based estimation	104
5.7.2	Wavelet Estimation	107
5.8	Scale-invariance of spatial fields	108
5.9	Scale-invariance of time series	111
5.9.1	Long-range correlation: the first evidence of long-memory systems	112
5.9.2	Scaling range of rainfall time series	113

5.9.3	Point-rainfall scale-invariant models	114
5.10	Scale invariance of space-time rainfall	117
5.10.1	The “Frozen Turbulence” hypothesis	117
5.10.2	The concept of dynamic scaling	118
5.11	Scale invariance of extreme point-rainfall: IDF curves	120
5.11.1	IDF scaling	121
5.12	Application of Multiplicative Cascades	123
5.12.1	Other downscaling techniques	124
5.12.2	Multiplicative Cascades	125
5.12.3	Bare and dressed quantities	126
5.12.4	Canonical and Micro-canonical cascades	126
5.12.5	Discrete and continuous cascades	127
5.12.6	Cascade Implementation	127
5.12.7	Implementation of the multiplicative cascades	130
5.13	Conclusion	131
6	Spatial Rainfall Extremes	133
6.1	Introduction	133
6.2	From point to spatial maxima: ARF	134
6.2.1	Background	134
6.2.2	Geostatistically-based ARF approaches	137
6.2.3	Stochastic approaches	139
6.2.4	Scale-invariant ARF	140
6.3	Max-stable spatial-maxima modeling	143
6.4	Conclusion	146

Part III

Results

7	Heavy tails of rainfall distributions	149
7.1	Introduction	150
7.2	Article: Scaling properties of heavy rainfall at short durations: a regional analysis	151
8	A scale-invariance Intensity-Duration-Frequency model	167
8.1	Introduction	167
8.2	Article: Intensity-Duration-Frequency curves in a GEV scale-invariant framework	169
9	Space-time scaling of rainfall events: the September 2002 storm	185
9.1	Introduction	185
9.2	Outline of the study	186
9.3	Data	187
9.4	Scale Invariance: Generalized structure Function	188
9.5	Scale invariance in Space-Time	189
9.5.1	Connections with Geostatistics	190
9.5.2	Space-time scale-invariance	190
9.5.3	Evaluation of dynamic scaling	192
9.6	Final remarks	192
9.7	Conclusion and perspectives	194

10 Qualification of Meso-scale meteorological simulations	195
10.1 Introduction	195
10.2 Severity Diagrams	196
10.3 Rainfall intensity diagrams: an indicator of the true model resolution	198
10.4 Article. Severity diagrams: a new approach for the multi-scale evaluation of extreme rainfall events	200
10.5 Severity diagrams and ensemble simulations	221
10.5.1 Severity Diagram of observed fields	221
10.5.2 AROME ensemble 1: Variability of Boundary Conditions	221
10.5.3 AROME ensemble 2: Variability of Initial Conditions	222
10.6 Conclusion	222

Part IV

Conclusion and Perspectives

11 Conclusion	229
12 Perspectives	233

I Acknowledgments

Je tiens tout d'abord à exprimer ma profonde gratitude à Gilles Molinié et Jean-Dominique Creutin qui m'ont encadré de façon remarquable. Je suis très heureux d'avoir pu travailler avec eux. Au-delà de leurs qualités professionnelles, je les remercie particulièrement pour la confiance et la liberté d'explorer qu'ils m'ont accordées et pour leur qualités humaines extraordinaires.

Je souhaite également remercier les directeurs du laboratoire Jean Dominique Creutin et Thierry Lebel et le directeur de l'équipe Sandrine Anquetin qui m'ont accueilli, respectivement, dans le laboratoire LTHE et dans l'équipe ASP.

J'exprime de plus mes sincères remerciements à P. Burlando et E. Foufoula-Georgiou qui ont accepté de rapporter sur ma thèse et à M. Déqué, C. Lantuéjoul, F. Renard, R. Uijlenhoet qui ont accepté de faire partie du jury. Merci à Stéphane Girard d'avoir accepté l'invitation en tant que membre invité.

Je tiens à remercier M. Charles Obled pour les discussions passionnantes et ses conseils bienveillants qui m'ont beaucoup aidé. Ses capacités pédagogiques, son expérience remarquable et sa disponibilité à la discussion ont été une aide précieuse. Merci à Eric Barthelemy de m'avoir permis d'effectuer des enseignements en hydraulique malgré mes modestes connaissances du français. Je suis reconnaissant à Efi Foufoula Georgiou pour m'avoir invité au Laboratoire SAFL de Minneapolis: les discussions avec elle et Mohammad Ebtehaj resteront des moments de science inoubliables. Merci à Vitek Krajewski de m'avoir invité à donner un séminaire à l'Université du Iowa. Je remercie le groupe MISTIS de l'INRIA Alpes Grenoble pour m'avoir accueilli, en particulier Stéphane Girard, Laurent Gardes, Caroline Bernard-Michel, Eugen Ursu et Julie Carreau. Je remercie aussi M. Daniel Duband pour avoir soutenu mes recherches. Merci au collègues avec qui j'ai partagé le bureau: Gilles et Benoît, Brice et Pierre-Emmanuel, Angélique, Olivier et Matthieu, ainsi que les collègues du bâtiment A: Nan, Abdel, Stephanie, Moussa et Joris, Isabelle, Celine, Arona, Laurence, Marc et la délégation du bâtiment G, dont Theo, Renaud, Aurelien, Laurent, Eric, Jeremy. Je remercie également le secrétariat du LTHE: Odette, Martine, Elif, Joseph, Synthia, Claire, Ghyslaine pour l'aide dans les tâches administratives et leur disponibilité.

Enfin, je souhaite adresser un remerciement à ma famille, qui à également contribué à ce résultat: mes parents et ma sœur, Pierangelo et Susanna avec Mao and Felipe, et bien sûr Alessandra, qui a été une présence fondamentale et avec qui j'ai partagé ces trois années de joie, colère, bonheur et stress, voyages et expériences.

II Keywords

analysis application approach arf average based behavior
cascade case computed correlation curves daily data defined density
determination deviation different distance
distribution dynamic echelle empirical equation
estimation events example extremes factor
field figure function gage generalized
geostatistics hydrology intensity interpolation invariance km kriging
limited line linear main maxima maximum measurement method
model moment multi-fractal network number observations
order parameters period plot point possible precipitation
present probability process properties provided radar rain
rainfall random range region research resolution
results return sample scale-invariance scaling
section series simple size space-time spatial
statistical stochastic storm structure study temporal therefore used
value variability variance variogram years

III Abstract

Ce travail de thèse concerne la caractérisation de la structure spatio-temporelle des fortes précipitations dans la région Cévennes-Vivarais. La région est soumise à des événements de pluie catastrophiques dont la magnitude gouverne les conséquences à différentes échelles de temps et despace. La détermination de la probabilité d'occurrence des orages est problématique à cause du caractère extrême des ces événements, de leur dimension spatio-temporelle et du manque de données pluviométriques aux échelles d'intérêt. Nous proposons d'adopter des approches d'invariance d'échelles afin d'estimer la fréquence d'occurrence de ces événements. Ces approches permettent d'extrapoler la distribution de la pluie à haute résolution à partir de données d'intensité pluvieuse à plus faible résolution. La paramétrisation de ces modèles étant fortement dépendante de l'incertitude de la mesure, nous avons d'abord caractérisé l'erreur commise dans la mesure de la pluie par un réseau de pluviomètres à augets. Nous avons ensuite exploré le comportement des pluies extrêmes dans la région d'étude, identifiant les gammes d'invariance d'échelles des extrêmes. Dans cette gamme d'échelles, nous présentons un modèle régional Intensité-Durée-Fréquence qui prend en considération l'hétérogénéité spatiale des extrêmes dans la région. Étant donné que le réseau pluviométrique ne permet pas de détecter les propriétés d'invariance d'échelle spatiale des champs de pluie, nous avons adopté une méthode semi-empirique pour modéliser des intensités de pluie intégrés sur des surfaces données (pluie surfacique) sur la base du concept de la mise en échelle dynamique (dynamic scaling). Cette modélisation permet la construction d'un modèle régional Intensité-Durée-Fréquence-Surface. Enfin, nous avons appliqué ce modèle à la construction des diagrammes de sévérité pour trois événements marquants en région Cévennes-Vivarais, afin d'identifier les échelles spatio-temporelles critiques pour chaque événement. Grâce aux diagrammes de sévérité, nous avons pu évaluer, pour ces mêmes événements, la performance d'un modèle météorologique de méso-échelle.

Abstract

The thesis is devoted to the characterization of the space-time structure of heavy rainfall events in the Cvennes-Vivarais area (France). The region is prone to catastrophic storms whose magnitude governs social and economic consequences at different space and time scales. The magnitude of an event cannot be univocally related to a probability of occurrence. The determination of the occurrence probability of storms is problematic because of their extreme character, of their complex space-time development and of the lack of rainfall data at the spatial and temporal scales of interest. We propose to adopt scale-invariant approaches in order to estimate the heavy rainfall frequency assessment. These approaches allow to extrapolate the high resolution rainfall distribution based on low resolution rainfall intensity data. The model estimation being heavily dependent of the data accuracy, the first step consists in the characterization of the error committed in the point and spatial rainfall estimated by tipping-bucket raingage networks. We then explore the extreme rainfall behavior in the region, detecting the range where extremes are scale-invariant. In this range, we present a regional Intensity-Duration-Frequency model for point rainfall maxima taking into account the heterogeneity of extremes in the region. We demonstrate that the rainfall network does not allow to detect scale-invariant properties of extreme rainfall fields, and then we adopt a semi-empirical method based on the concept of dynamic scaling to build regional Intensity-Duration-Area-Frequency curves. Finally, we apply this model for the determination of the severity diagrams for three significant storms in the Cvennes-Vivarais region, with the aim to identify the critical space-time scales of each event. Based on severity diagrams, we then evaluate, for the same events, the performances of a mesoscale meteorological model.

IV Introduction

Les événements extrêmes peuvent avoir un impact majeur sur la vie quotidienne. Bien que rares par définition, ils impliquent un grand nombre de personnes et des biens, causant des dégâts importants en termes socioéconomiques et de vies humaines. Qu'ils soient inondations, sécheresses, vagues de chaleur ou de froid, ouragans, séismes, incendies ou tsunamis, ils mettent en danger la société. Ces dernières années, la communauté a ressenti la sensation d'augmentation de la fréquence de ces événements. Cela est probablement du, au moins en partie, à la couverture médiatique accrue. Toutefois, la fréquence de certains phénomènes tels que les inondations, les sécheresses, les ouragans et les incendies, peuvent avoir été intensifiés par les effets de l'activité de plus de 6 milliards de personnes. Le réchauffement climatique se fait sentir dans plusieurs régions du globe, causant la désertification, la fonte des calottes polaires ou des glaciers. L'augmentation générale de la température est souvent désignée comme possible raison des l'incrément de la fréquence des phénomènes de précipitations intenses dans les régions tempérées. Dans des régions comme les Cévennes, où la situation géographique favorise l'occurrence d'événements de pluie intense en raison de la proximité de la mer Méditerranée et de l'influence de la topographie, il y a eu ces dernières années une forte incidence de phénomènes des pluies considérées comme "extrêmes". L'objectif que nous nous fixons est de quantifier la fréquence de ces phénomènes qui semblent être des "monstres météo-hydrologiques".

Afin d'imaginer l'évolution des événements de pluie intense dans des scénarios de climats futurs, il faut d'abord diagnostiquer l'intensité et la fréquence des pluies intenses de ces dernières décennies. Pour cela, nous procédons à une analyse détaillée des intensités de pluies horaires et journalières enregistrées maintenant depuis plusieurs décennies dans la région Cévennes-Vivarais. Des telles analyses devraient être effectuées en considérant non seulement la pluie ponctuelle telle qu'elle est enregistrée par les pluviomètres mais aussi son extension spatiale sur des zones correspondantes à des bassin hydrographiques. De la même manière que l'échelle spatiale, l'échelle temporelle est d'une importance capitale pour la mesure de l'intensité de la pluie: les régions méditerranéennes sont soumises à des événements de forte intensité qui durent quelques dizaines de minutes, tandis que les zones de montagne reçoivent des intensités plus faibles mais qui s'étalent sur plusieurs heures, avec des cumuls élevés sur 24 heures.

Cette thèse est consacrée à examiner la question du changement d'échelle des précipitations extrêmes dans l'espace espace-temps. Un des principaux aspects théoriques qui peuvent être utilisés pour expliquer, comprendre et modéliser les phénomènes extrêmes à différentes échelles est le concept d'invariance d'échelle. Dans les plages d'échelles où cette hypothèse est confirmée, ils existent des techniques pour reconstruire un phénomène à une échelle différente de celle des mesures.

Le potentiel des approches qui traitent des phénomènes d'invariance d'échelle est énorme et d'un point de vue applicatif beaucoup de problèmes peuvent être abordés du point de vue de l'invariance d'échelle. Bien que d'un grand potentiel, le concept d'invariance d'échelle souffre de la méconnaissance de la physique qui lui est sous-jacente. Dans cette thèse, les fondements théoriques et les méthodes d'applications de l'invariance d'échelle sont discutées et appliquées dans le contexte de la région Cévennes-Vivarais. L'objectif est de diagnostiquer et décrire aussi précisément que possible l'invariance d'échelle des intensités de pluies extrêmes.

La thèse est structurée de la façon suivante. La **Partie I** donne une description générale du contexte de la thèse. Dans le **Chapitre 1**, nous présentons les questions scientifiques qui ont motivé cette thèse. Ensuite, nous décrivons le contexte géographique, les principales entités hydrologiques et les caractéristiques hydro-météorologiques marquantes de la région d'étude, avant de décrire les caractéristiques principales des réseaux de mesures utilisés pour cette étude. Dans le **Chapitre 2** nous nous focalisons sur la quantification de l'erreur commise en mesurant la pluie par le biais d'un réseau de mesure au sol. Bien que le principe de la mesure de la pluie par pluviomètres à augets basculants soit ancien (le premier date de 1662) et qu'il ait fait l'objet d'études très précises qui aboutissent régulièrement à des améliorations, nous montrons que la question de l'**incertitude** liée à ces mesures est encore d'actualité. Une première partie analyse l'incertitude dans l'estimation de la pluie ponctuelle: l'erreur due à l'échantillonnage peut être conséquente si la résolution de mesure est proche de l'échelle d'analyse. Une autre partie de l'erreur est due à la façon dont l'instrument de mesure (pluviomètre à auget) est conçu. La deuxième partie traite de l'évaluation de l'erreur d'estimation commise lors de la mesure de la lame d'eau spatiale maximale à partir de données pluviométriques.

La **Partie II** est le noyau de la thèse et présente un état de l'art des techniques qui sont utilisées dans ce travail. Le **Chapitre 3** est consacré à la **théorie des valeurs extrêmes** et ses conséquences pour la modélisation des extrêmes. L'objectif est de donner au lecteur un aperçu du problème de l'estimation de la fréquence d'occurrence des extrêmes. Nous montrons que quelle que soit la méthode, la sélection d'un échantillon d'extrêmes est un délicat compromis entre robustesse et biais de l'estimation. Le **Chapitre 4** décrit les concepts de base de l'**analyse géostatistique**, pour l'interpolation des données et l'analyse structurelle de champs n -dimensionnels. L'état de l'art de l'**invariance d'échelle** est présenté dans le **Chapitre 5**. Nous décrivons les premières approches théoriques et des techniques plus innovantes pour estimer les propriétés d'invariance d'échelle des champs géophysiques. Les résultats principaux dans l'analyse de champs 1D, 2D et champs dans l'espace espace-temps sont listés. Les méthodes de **désagrégation** par invariance d'échelle connues sous le nom de "**cascades**" sont introduites en fin de chapitre. Un des objectifs de la thèse étant la caractérisation de la pluie spatiale, le **Chapitre 6** reporte une série de méthodes empiriques ou semi-empiriques pour l'estimation d'une lame d'eau extrême à partir de mesures ponctuelles.

La **Partie III** présente les principaux résultats obtenus pendant la thèse. Premièrement, Ils sont présentés selon un parcours scientifiquement cohérent. Premièrement, nous étudions les fréquences d'occurrence des intensités extrêmes de pluies ponctuelles. Elles sont modélisées par des fonctions de densité dites à **queues lourdes** (**Chapitre 7**). La connaissance des distributions de ces intensités nous a permis d'établir des relations d'invariance d'échelle des précipitations extrêmes. Ces relations nous ont permis de proposer un modèle de **courbes Intensité-Durée-Fréquence** compatible avec les différentes typologies d'invariance d'échelle des extrêmes dans la région d'étude (**Chapitre 8**). Du côté de l'étude de la pluie spatiale, le **Chapitre 9** montre que la pluie spatiale est invariante d'échelle à plusieurs pas d'agrégation temporelle. De plus, on montre qu'une relation entre les échelles spatio-temporelles peut être définie en accord avec le comportement des flux turbulents ("**dynamic scaling**").

La détermination des courbes de réduction surfaciques des pluies (**Areal Reduction Factors**) a permis d'étendre le modèle d'invariance d'échelle ponctuel à la pluie spatialisée, obtenant les courbes

***Intensité-Durée-Fréquence-Surface (IDAF)** pour la région, basées sur le concept d'invariance d'échelle dynamique. Avec ce modèle, il est possible de déterminer la fréquence de tous les événements de pluie dans la région par rapport à leur intensité, qu'elle soit ponctuelle ou surfacique. La fréquence des intensités de pluies à différentes échelles spatio-temporelles est un diagnostic de la sévérité des orages. Nous avons indentifié 3 événements qui sont à l'origine des crues rapides. Nous avons calculé leur sévérité grâce d'une part aux mesures et d'autre part aux simulations de l'intensité de pluie par le modèle météorologique méso-échelle MesoNH. La comparaison de la sévérité simulée et observée par l'utilisation des **Diagrammes de sévérité (Chapitre 10)** constitue une approche innovante pour la qualification des simulations numériques. L'usage de ces diagrammes a ensuite été testé sur des simulations d'ensemble. L'objectif est la de mieux comprendre comment les perturbations dans le conditions initiales/au contour du modèle puissent être choisis pour obtenir un ensemble statistique fiable. L'usage des diagrammes de sévérité pour la qualification des simulations d'ensemble met en valeur l'indication synthétique que cet outil fournit. En effet, il fournit dans un seul diagramme une indication de la justesse des simulations sur une large gamme d'échelles spatio-temporelles et même une indication sur la justesse de la localisation de l'événement.*

*La partie finale de la thèse, **Partie IV**, montre les conclusions principales du travail (**Chapitre 11**) et en décrit les perspectives (**Chapitre 12**).*

Introduction

Extreme events can have a major impact on everyday life. Although rare by definition, they involve a lot of people and goods, causing extensive damages from the socioeconomic point of view and in terms of human lives. Taking the form of floods, droughts, heat or cold waves, hurricanes, earthquakes, fires, tsunamis, they endanger the society. In the recent years, a widespread discussion concerns the feeling that these events are becoming more frequent. This is due, at least in part, to the increase of the media attention. However, the frequency of phenomena such as floods, droughts, hurricanes and fires, may have been intensified by the effects of the activity of over 6 billion people. **Global warming** is observed in many regions of the globe, causing desertification, melting of polar icecaps or glaciers, and it is often designated as a likely reason for the increase of extreme precipitation frequency in temperate regions. The Cévennes-Vivarais region, where the proximity to the Mediterranean Sea and the rough topography favour the occurrence of heavy precipitation events, was submitted, in the last decades, to a number of rainfall events that can be reasonably referred to as “extremes”. is ideal for the precipitation development due to the proximity to the Mediterranean Sea and where the topography favors the stabilization of rainfall phenomena, there has been in recent years a high incidence of phenomena rainy considered “extremes”. Our purpose is to quantify the frequency of these phenomena that could appear as ”**meteo-hydrological monsters**”.

In order to imagine the evolution of heavy rainfall events in a context of **future climate scenarios**, one must firstly determine intensity and frequency of heavy rainfall events occurred in the past. For this, we will perform a detailed analysis of hourly and daily rainfall intensities observed in the last decades in the Cévennes-Vivarais region. This analysis must consider point rainfall recorded by rain gauges as well as the rainfall falling over surfaces of different sizes: more than the rainfall amount recorded at one gage, the spatial extent of an event is the actual responsible of basin floods. Similarly to the the spatial scale, time scale is of fundamental importance: the Mediterranean areas are more subject to intense events lasting few dozen minutes, while mountainous regions may be submitted to stationary phenomena with weak intensity but significant rain accumulations over 24 hours.

This thesis is devoted to the characterization of the **scaling of extreme precipitations** in space and time. One of the main theoretical aspects that can be used to explain, understand and model the extreme events at different scales is the concept of “**scale invariance**”. In the ranges where this assumption is confirmed by the empirical analysis, there are techniques to reconstruct a phenomenon at scales not covered by direct measurements.

The potential of scale-invariant approaches is enormous and, in many practical problems, the adoption a scale-invariant point of view can be useful. Despite this potential, the scale-invariance concept suffers of the lack of knowledge relative to the underlying physics. In this thesis, the theoretical foundations and application methods of scale invariance are discussed and applied in the Cévennes-Vivarais region. The goal is to provide a description, as detailed as possible, of the relationships between extreme precipitations and scale invariance.

The thesis is structured as follows. **Part I** is devoted to the description of the context of the thesis. In **Chapter 1** we present the scientific questions at the origin of the thesis. Afterwards, we introduce the region of study, detailing the meteo-hydrological context with the description of the main

climatic features and the main hydrological entities. At last, we describe the measurement network used throughout the thesis. **Chapter 2** is devoted to the quantification of the error committed in measuring rainfall with a ground measurement system. Even if the first tipping-bucket raingage dates back to 1662 and several studies have been conducted on its mechanism leading to constant improvements, we show that the issue of **uncertainties** related to the rainfall measurement still exists. A first part focuses on the uncertainties of the point rainfall estimation: the estimation error can be significant if the scale of analysis is close to the measurement resolution. Another error source is due to the mechanical structure of the tipping-bucket rain gauge. The second part deals with the evaluation of the estimation error in the spatial rainfall amount estimated from point measurements.

Part II is the core of the thesis and presents a detailed state of the art of the techniques used in the thesis. **Chapter 3** presents the **extreme value theory** and its consequences for the extreme modelling. The aim is to provide the reader with a sensibility to the issue of extreme values estimation; we show that, for a correct modeling of the extremes of limited samples, a compromise between robustness and bias of the estimation must be found. **Chapter 4** describes the basic concepts of **geostatistical analysis**, used for the data interpolation and for the structural analysis of n-dimensional fields. The state of the art of the **scale-invariance** is given in **Chapter 5**. We describe the earlier theoretical approaches and the innovative techniques to estimate the scale-invariant properties of geophysical fields. The main results in the analysis of 1D, 2D and space-time fields are listed. The scale-invariant **disaggregation methods** known as “**cascades**” are presented at the end of the chapter. Since one of the objectives of the thesis was the characterization of spatial rainfall, in **Chapter 6** we report the empirical and semi-empirical methods for the estimation of an extreme rainfall depth over a surface from point measurements, namely **Areal Reduction Factor (ARF)**.

Part III presents the results that have been found during the thesis. First of all, we study the occurrence frequency of point-rainfall extreme intensities and verify if extremes exhibit **heavy-tails**. We also determine if heavy rainfall intensities exhibit scale-invariance (**Chapter 7**). These findings allow us to propose (**Chapter 8**) a scale-invariant **Intensity-Duration-Frequency** model that can accommodate, in each sub-region, the different behavior of extremes. From the point of view of spatial rainfall, the empirical computation of **Areal Reduction Factors** allows to fit a statistical scale-invariant model of **Intensity-Duration-Frequency-Area (IDAF) curves**, based on the concept of “**dynamic scaling**”. In addition, we show that spatial rainfall is scale-invariant at various temporal aggregation scales (**Chapter 9**). With the IDAF model we can determine the frequency of any spatial or point rainfall event within the region. The determination of the occurrence frequency of a storm at different space-time scales allows to draw **Severity Diagrams**, representing the maximum return period of a storm at all space-time scales. For 3 events that originated flash-floods in the Cévennes-Vivarais region, we computed the severity diagrams of the event as observed by the rain gauge network. To test the performance of the MesoNH meteorological model, we compared the severity diagrams of the observed event with those obtained for the MesoNH meteorological model simulations **Chapter 10**. This innovative approach for the multi-scale evaluation of numerical simulation has also been tested for **ensemble simulation**. In this case, the aim is to better intervene in the choice of the initial/boundary conditions of the model in order to obtain a statistically reliable ensemble. The use of severity diagrams for the ensemble simulation evaluation give emphasis to the diagnostic capacity

offered by this intuitive indicator. In a single diagram, it gives elements on the extreme behavior of the storm in a large range of space-time scales, offering additional indications on the capacity of the model to predict the correct storm location.

The final part of the thesis, **Part IV**, reports the main conclusions of the work (**Chapter 11**) and describe some of the short-term and long-term perspectives **Chapter 12**.

Part I

Context

1 General context

Résumé

*Ce premier chapitre est dédié à la présentation des problématiques que l'on aborde dans la thèse. Dans la première section, nous détaillons les objectifs de la thèse. Le chapitre donne ensuite les éléments fondamentaux du **contexte géographique** et **hydro-météorologique** de la région d'étude (Cévennes-Vivarais): description des principaux bassins hydrologiques et caractéristiques climatiques de base de la région (précipitation annuelle, mensuelle et intermittence pluviométrique, prédisposition aux évènements extrêmes) sont détaillées. En conclusion de ce chapitre, nous décrivons sommairement les **bases de données** utilisées dans l'étude, couvrant plusieurs échelles d'espace et de temps.*

1.1 Purpose of the thesis

The present thesis has been developed at the LTHE, Laboratoire d'Études des Transferts en Hydrologie et Environnement de Grenoble, in the framework of the French project MEDUP "Forecast and projection in climate scenario of Mediterranean intense events: Uncertainties and Propagation on environment".

This thesis is developed to deal with the following points:

- i. define the uncertainties in the rainfall measurement due to the measurement network in the region;
- ii. verify the presence of scale-invariance of heavy rainfall events (point and spatial rainfall) in the region with the aim to improve the extreme events modeling at any spatial and temporal scale;
- iii. explore the scale-invariant properties of space-time rainfall fields with the aim to model the fine-scale variability of the rainfall process;
- iv. quantify the impact of meteorological events over a region at each temporal and spatial scales and the uncertainties associated with the ensemble forecasts.

In the following paragraphs, we describe these items.

1.1.1 Rainfall measurement uncertainties

Before the XX century, rainfall was measured at monthly to daily resolution. The limited communication facilities prevented the construction of reliable rain gauge networks. With the industrial revolution, the knowledge of the hydrological balance of basins began to be of interest for the hydroelectric production, and a series of (mainly daily, seldom hourly) gages was installed even in mountainous regions. In the 70s, the massive urbanization, developing faster than even before, caused significant changes in soil occupation and modifications in the drainage networks. Urban flash-floods began to occur, characterized by very fast response times preventing the alert system to efficiently work. To better observe these phenomena, high-resolution rainfall measurement devices (e.g. 1 hour to 6 min resolution) began to spread in metropolitan regions; most of them are automatic, such as tipping-bucket rain gauges (Figure 1.1). In the early 80s, the first radar installations allowed to make the preliminary calibrations and comparison between radar measures and ground rainfall.

At the present time, except rare cases, series of automatized measures sufficiently long to perform statistics without sampling size issues are not available, due to the technological limits of measurement devices before the half of XX century. Even when long rainfall series are available, the rainfall measurement at a point is submitted to a series of uncertainties that we will quantify in Section 2.

The estimation of spatial rainfall is usually done by interpolating point data. Therefore, the interpolation error sums to the point-rainfall uncertainty. The interpolation error is related to the spatial sampling and to the correlation structure of the rainfall field.

The study of the measure uncertainties (Section 2) is fundamental for determining the accuracy of the measurement of point and spatial rainfall and by consequence for characterize the reliability of the extreme analysis.

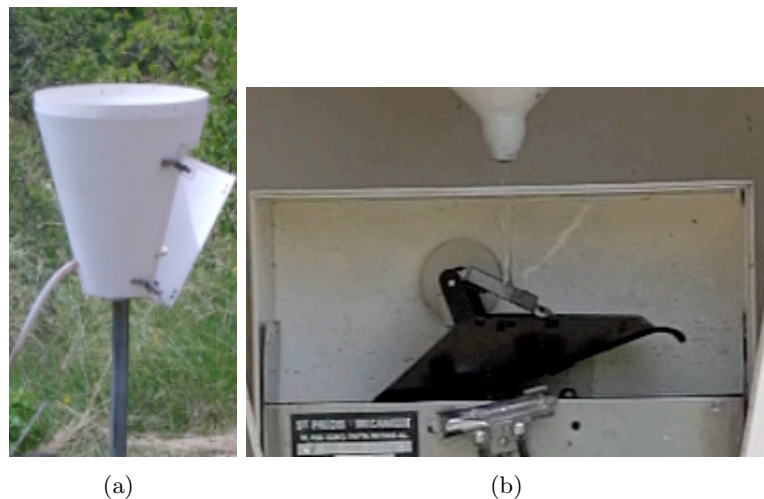


Figure 1.1: a: An operational rain gauge device located in Le Pradel, Ardèche. b: the tipping bucket system inside the rain gauge to record rainfall intensities.

1.1.2 Scale-invariance of heavy space-time rainfall

The Cévennes-Vivarais region is submitted to intense flash-flood events causing severe social and economic damages. We would like to describe the way in which extremes decrease with the involved surface and with the accumulation time. Scale-invariance is an interesting framework to model processes at ungaged scales. We will test the scale-invariant hypotheses on some reference series to infer the general behavior of rainfall in the whole region. In this thesis we will show evidences of scale-invariance of point and spatial heavy rainfall derived from the analysis of rain gauge and radar data and we will define the validity range of scale-invariance.

Scale-invariant methods provide an innovative framework to infer the behavior of a variable at scales where direct measurements are not available. Climatological studies need the analysis of long series but, for instance, hourly rainfall records are not available except in the last 20-30 years and only in some specific rain gauges. Applying scale-invariance, it would be possible to infer the behavior of extreme hourly rainfall for very large return periods from observations at the daily scale.

1.1.3 Scale-invariance of rainfall fields in space-time

Rainfall is a strongly intermittent process. The scales of analysis have large influence in the rainfall properties. Taking for example point rainfall, if one increases the temporal resolution of analysis, the rainfall process becomes more and more intermittent (the proportion of zeros on the total increases) and singular (in time series, for instance, the ratio between maximum and average rainfall increases). The spatial and temporal scales have similar roles in the variability determination: the increase of the integration window over which the rainfall amount is computed causes field smoothing and variability reduction. Dealing with spatial data, at small spatial scales (few dozen meters) rainfall appears as a very spiky phenomenon, while integrating over windows of increasing size the fields are smoothed and their distributions get narrower.

The knowledge of the relations between rainfall in space and time at different scales is therefore of primary importance to better understand how to interpret a series of point measurements and to provide a reliable estimation of the rainfall depth over a basin, which is the main input of hydrological models.

1.1.4 Multi-scale evaluation of the impact of meteorological events

One of the main interests of the hydrological community is to quantify the social and economic impact of storms, mainly determining its recurrence interval. The main factor determining the impact of an event is the rainfall depth produced by the storm; the basin response is a subordinate factor. For this reason, we think that a better characterization of floods cannot prescind from a better characterization of storms.

To identify the potential danger engendered by storms, we propose to determine their magnitude at each space-time scale. Establishing the occurrence probability of the event we may be able to determine if the event are common, rare or extreme and at which scales. The pre-requirement is the knowledge of the behaviour of point and spatial rainfall-intensity extremes in the region of analysis.

In this thesis we propose to quantify the impact of heavy rainfall events occurred in the Cévennes-Vivarais region through the use of Severity Diagrams, a tool designed to display the magnitude of an event in a range of spatial and temporal scales. The tool provide all the elements for an objective multi-scale comparison between the extreme features of storms occurred in the same region or in different regions.

We propose an innovative usage of this tool for the qualification of the mesoscale simulation capability to reproduce the space-time structure of three storms occurred within the region of analysis.

1.2 Geographical context

The Cévennes-Vivarais region is located on the south-eastern side of Massif Central. It is easterly bounded by the Rhône river, westerly bounded by the Massif Central, and southerly by the Mediterranean Sea shore. The mountainous part of the region is included in the Cévennes National Park, created in 1970.

The size of the region is 160 km in the E-W direction and 200 km in the N-S direction, for a total surface of 32000 km^2 .

The elevation in the domain raises the 1699 m (Mt. Lozère). The main mountain ridge of the Massif Central is oriented SSW-NNE, and it can be approximately detected in Figure 1.2 by drawing a straight line joining Mont Aigoual (1565m height), Mont Lozère (1699 m), Mont Gerbier de Jonc (1551m) and Mont Mezenc (1753m). A number of fine-scale structures are present, including a series of “shoulders”, perpendicular to the main ridge and delineating the valleys, most of which are oriented NW-SE. These valleys present common features, such as their depth (about 500 meters) and length (in the order of 10 km). Their separation distance (20 km) is a significant feature for the triggering of rainfall events (*Miniscloux et al.*, 2001).

The region presents several rivers, either tributary of the Rhône or directly flowing into the Mediterranean Sea.

1.3 Meteo-hydrological context

In 1999, the the INSU-CNRS (Institut National des Sciences de l’Univers) and by the OSUG (Observatoire de Sciences de l’Univers de Grenoble) created the OHMCV (Observatoire Hydro-Météorologique Cévennes-Vivarais), devoted to the study of extreme rainfall and flash-floods in this region prone to extremely intense events.

The technological development in the last decades is at the base of the recent improvement of the media diffusion of the OHMCV data: most of the measurements and data are available through a portal called SevNol (Système d’Extraction et de VisualisatioN des données de l’OHMCV en Ligne¹). OHMCV (Figure 1.4) is born to gather the measurement networks and to enforce the consultation among the socio-hydro-meteorological communities through a series of workshops and multi-disciplinary projets.

¹<http://sevnol.ohmcv.fr/Sevnol2/?lang=en>

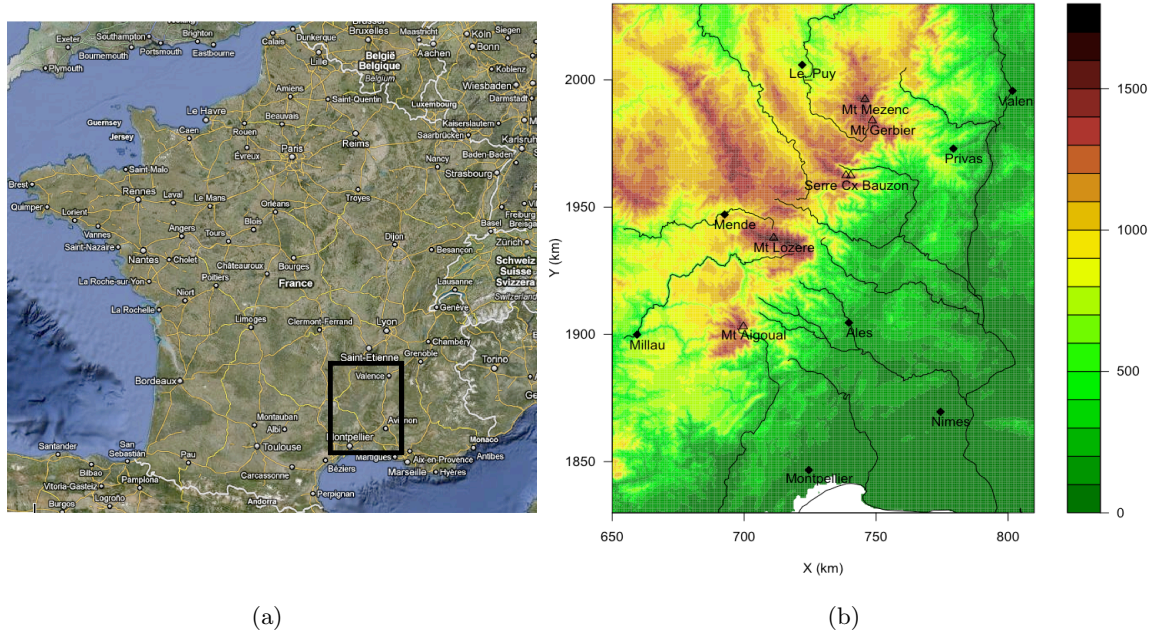


Figure 1.2: a: Location of the Cévennes-Vivarais region. b: Elevation map of the Cévennes region (in meters above sea level), main rivers, mountain peaks and cities.

In the recent years, the international HyMeX (HYdrological cycle in the Mediterranean EXperiment) programme, aiming at a better understanding of the hydrological cycle in Mediterranean Regions, has been defined. Similarly to the AMMA project, involving the African monsoon, it is a long-term project.

HyMeX share with OHMCV the integrated approach regrouping social sciences, economics, meteorology and hydraulics.

LTHE members are involved in the project HyMeX, with important tasks concerning the Working Group 3 on “Heavy rainfall, flash-floods and floods” and the Working Group 5, on “Societal and economic impacts”.

1.3.1 Climatic features

All along the thesis we adopt the thematic cartography to describe the features of the region. The map covers the Cévennes-Vivarais region, whose boundaries have conventionally been fixed in the extended Lambert II projection, as the rectangle of coordinates $X=[650 \text{ km}, 810 \text{ km}]$, $Y=[1830 \text{ km}, 2030 \text{ km}]$. The region is therefore $160 \times 200 \text{ km}^2$.

We can summarize the climatic features of the region by 2 main indicators. The average annual rainfall depth is the first indicator, it is independent of the scale of analysis and it gives a preliminary description of the main climatic patterns of the region. In Figure 1.5 the average annual rainfall, obtained by geostatistical interpolation (see Section 4), is reported. From the map it is evident that the drier sub-regions are located along the Mediterranean shore, while the highest rainfall depth are recorded along the Massif Central mountain ridge, oriented SW-NE, roughly obtainable by joining

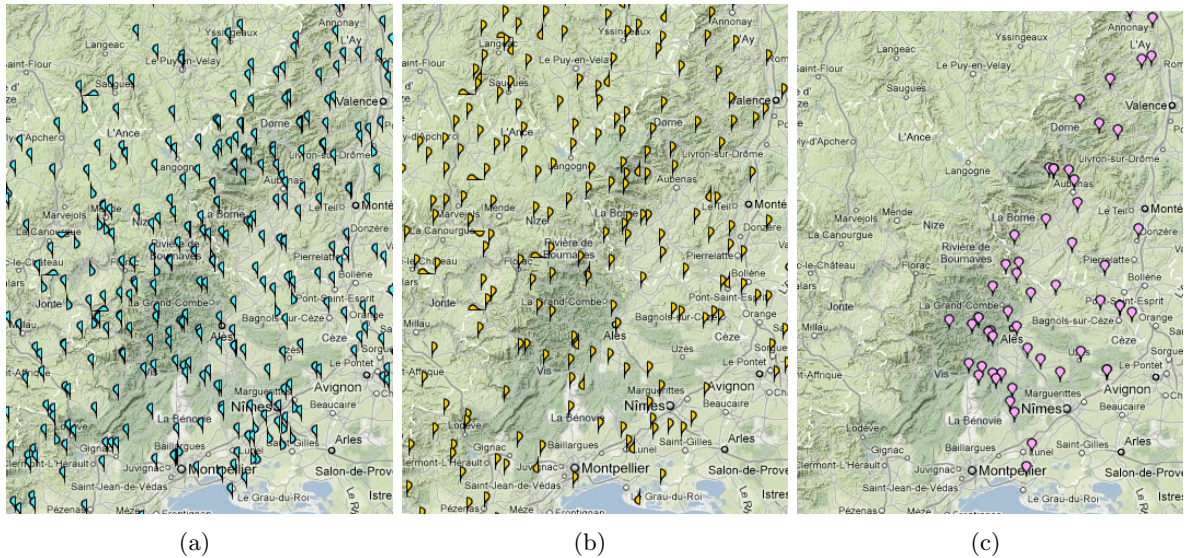


Figure 1.3: The OHMCV survey network. a: hourly rain gauge network. b: daily rain gauge network; c: limnimeter network;. From the SevNol website, OHMCV.

the Mt. Aigoual to the Serre de la Croix de Bauzon. These two locations feature the largest rainfall accumulation, with more than 2 m of equivalent rainfall depth per year.

Another relevant indicator of the climatic behavior of the region is the number of rainy days. From the map in Figure 1.6, we see that the Rhône Valley and the Mediterranean shore feature the driest climate, while over the mountainous region of the Cévennes plateau, in the NW corner, it rains almost 50% of the days.

The climatic pattern take different spatial organization depending on the analyzed season. In Appendix B we draw some maps describing the monthly rainfall regime in the region, with the aim to show some peculiar characteristics of the rainfall regime. In sequence, we show the average monthly rainfall, its proportion with respect to the total annual depth in order to highlight the months in which a dry/wet regime is observed, followed by the montly intermittency (proportion of wet days to the total).

Further information on the climatological behavior in the Cévennes-Vivarais region, with particular interest on the orographic effect on the rainfall regime, can be found in *Molinié et al.* (2010).

1.3.1.1 Flash-Floods

The Cévennes-Vivarais region is naturally prone to very intense storms, for two reasons:

- its position, located at few dozens of kilometers from the Mediterranean Sea, source of warm and humid air masses;
- its complex topography: i) the main mountain ridge oriented 90° to the SSE flux generating heavy precipitation events; ii) narrow parallel valleys evenly separated, ideal conditions for the band convection triggering.

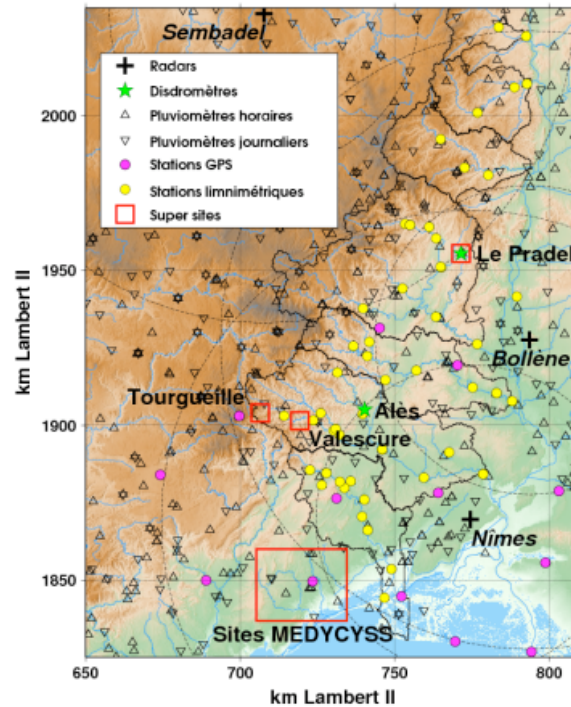


Figure 1.4: Comprehensive vision of the OHMCV survey. The super-site of Le Pradel and the new super-sites defined in the framework of the HyMeX project are reported. From the SevNol website.

The very fast storms that are produced in this zone are known as “Cévenols”. Three particular synoptic conditions have been identified, leading to a generalized southern flow that, depending on the flow incidence, may involve the Aude sub-region (as in the 1999 event) or the northern Gard subregion (as in the 2002 event). With this configuration, the low and warm layers coming from the Mediterranean are advected towards the Cévennes relief. These layers are extremely charged especially during the fall months, enforcing the atmosphere instability from the sea shore to the Cévennes foothills. The complex topography in the relief region is then a factor for the convection triggering and for the moist flux convergence (*Ducrocq et al., 2008*).

In some particular situations, the convection becomes stationary because of several concomitant ingredients: the relief forces the conditionally unstable and moist low-level jet to raise, generating stable systems that remain in the same location until changes in the synoptic conditions occur. A similar situation, not directly related to the orography, is called “cold-pool”, and is generated by the precipitation evaporation. This phenomenon is at the origin of the catastrophic flood occurred in 2002 over the Gard region (*Delrieu et al., 2005*).

1.4 Hydrological context

The hydrographic network of the region is composed, excepted the Vidourle river, by right tributaries of the Rhône River (Figure 1.7-a). The main basins interested by Flash-Floods and covered by the measurement devices of the OHM-CV survey system are shown in Figure 1.7-b.

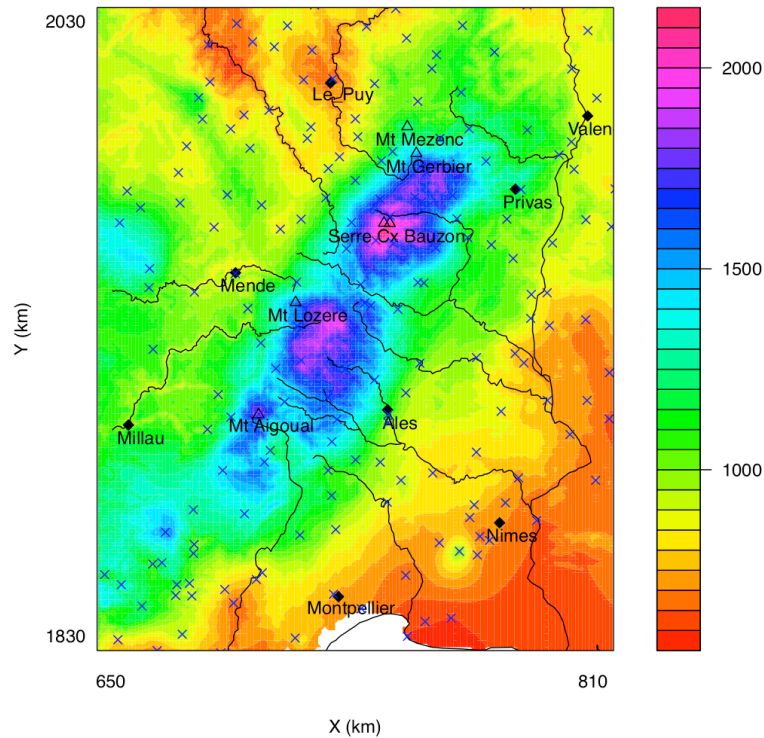


Figure 1.5: Average annual rainfall in the Cévennes-Vivarais region (mm yrs^{-1}).

In the next paragraphs, we briefly describe some physiographical features of the main basins located in the Cévennes-Vivarais region. Most of them have been repeatedly interested by flash-floods events in the last decades, and we report the main floods that have been therein recorded.

1.4.1 Ardèche basin

The Ardèche basin is the larger of the 4 basins interested by the flash-flood phenomena involving the region. With a surface of 2429 km^2 , it raises the elevation of 1700 m. The Ardèche river is 120 km long, with an average flow at the confluence estimated in $65 \text{ m}^3 \text{ s}^{-1}$. In Figure 1.8-a,b the elevation map and the hypsographic map are represented.

In some exceptional floods, such as in 1827, 1890 et 1924, the maximum flow reached more than $7000 \text{ m}^3 \text{ s}^{-1}$.

1.4.2 Cèze basin

The Cèze basin is characterized by a surface of 1329 km^2 . The Cèze river is 128 km long, with an average flow at the confluence estimated in $22 \text{ m}^3 \text{ s}^{-1}$. In Figure 1.9-a and 1.9-b the elevation map and the hypsographic curve are represented, respectively.

The maximum flow recorded at the station of La Roque-sur-Cèze was $2010 \text{ m}^3 \text{ s}^{-1}$, on 1 October 1977.

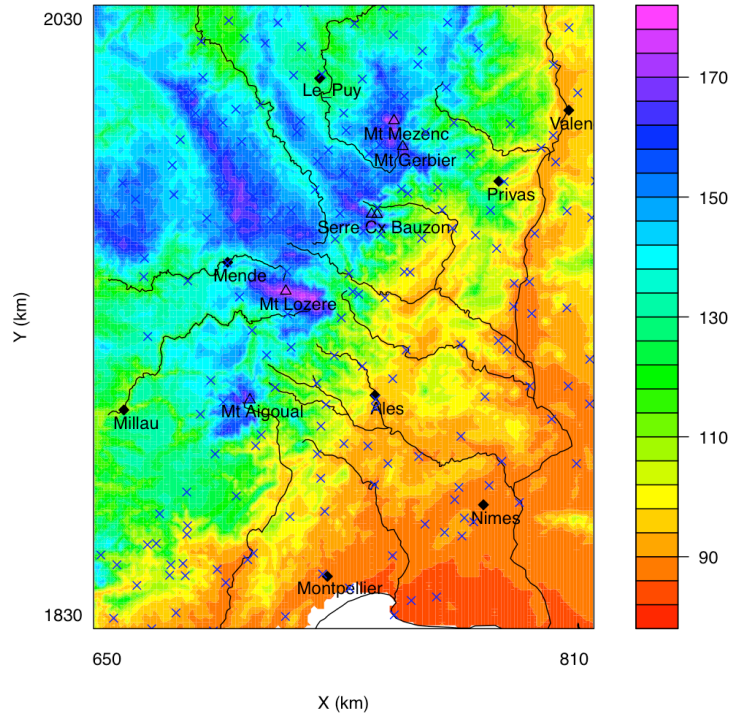


Figure 1.6: Average number of rainy days in one year in the Cévennes Vivarais region.

1.4.3 Gard basin

The Gardon river (also known as Gard) features two main branches: the Anduze Gardon and the Alès Gardon. The basin is characterized by a surface of 2200 km^2 . The river length is 127 km, with an average flow at the confluence of $32 \text{ m}^3 \text{ s}^{-1}$. In Figure 1.9-a and 1.9-b the elevation map and the hypsographic map are represented, respectively.

The basin was submitted to two major floods in 1958 and in 2002, September the 8-9th.

1.4.4 Vidourle basin

The Vidourle basin is the southern of the 4 basins mainly interested by the “Cévenoles” events. Its surface is 1335 km^2 , with an average flow at the estuary of $20 \text{ m}^3 \text{ s}^{-1}$.

A series of catastrophic floods have been recorded: 15 September 1575, 3 July 1684, October 1689, 1-3 October 1723, 18 November 1745, 6 October 1812, 17 September 1858, 1891, 26 September 1907, 16 October 1907, 27 September 1933, 4 October 1958 and 8-9 September 2002.

1.5 Measurement network

The analysis of rainfall at different scales is the main objective of the work. Consequently, the use of different databases (possibly spatially and temporally overlapped) is required.

We will make use of three kinds of data:

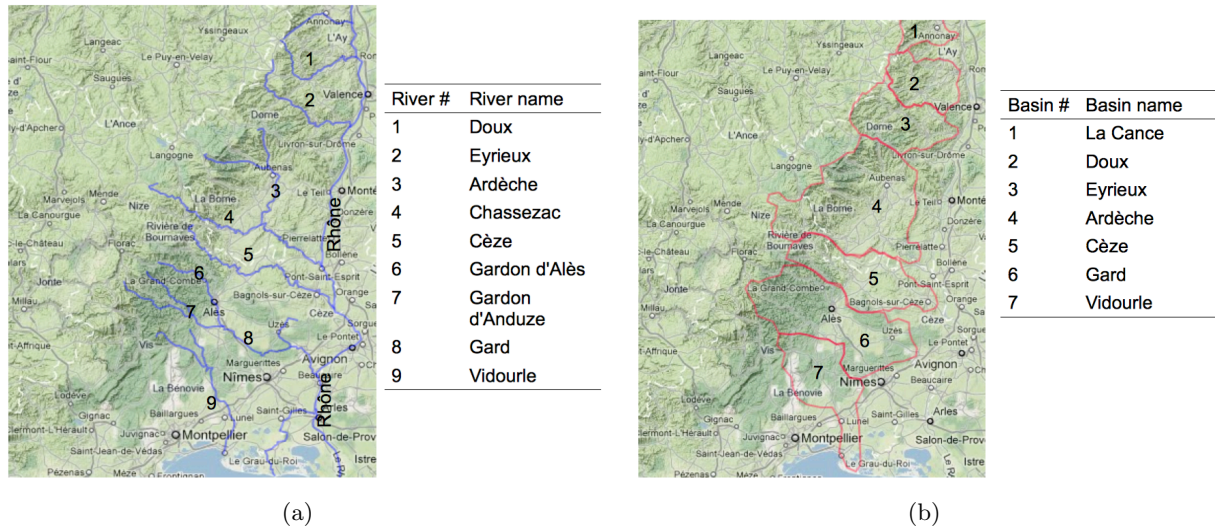


Figure 1.7: a: Main hydrographic network of OHM-CV. b: Main hydrographic basins of the OHM-CV region.

- Point-rainfall data at the rain gauges
- Radar-estimated rainfall fields
- Disdrometer data

In the next subsections, we briefly describe the main properties of each database, observation period and number of available devices.

1.5.1 Point-rainfall data

The Nîmes flood in 1950 highlighted the need of a densely instrumented survey system. Since then, the Cévennes-Vivarais region has been densely gaged, focusing on the measurement of the rainfall produced by deep convective events and on the orographic events occurring over the southeasterly-exposed Massif Central foothills.

In Figure 1.12 we report the locations of the hourly and daily rain gauge stations. Except for rough terrain zones where the density is lower (between Millau and Mount Aigoual for example), the rain gauge density is spatially homogeneous. The rain gauge distribution as a function of the elevation is showed in Figure 1.13, in which the surface of the relative elevation ranges are reported. The rain gauge density per elevation range is approximately constant. The area higher than 1600 m represents a very small proportion of the total surface and, due to accessibility issues and maintenance costs, it is poorly gaged.

In the next subsection, we analyze the daily rainfall database with particular interest to the availability of long series, a mandatory requirement for a reliable estimation of the extreme behaviour of rainfall.

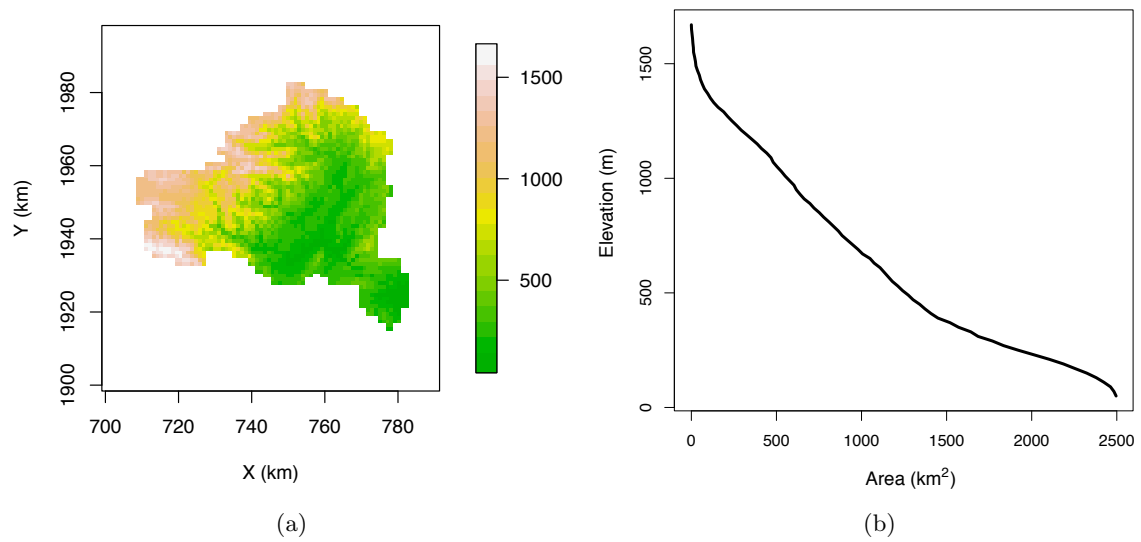


Figure 1.8: Ardèche basin. a: Elevation map. b: Hypsography.

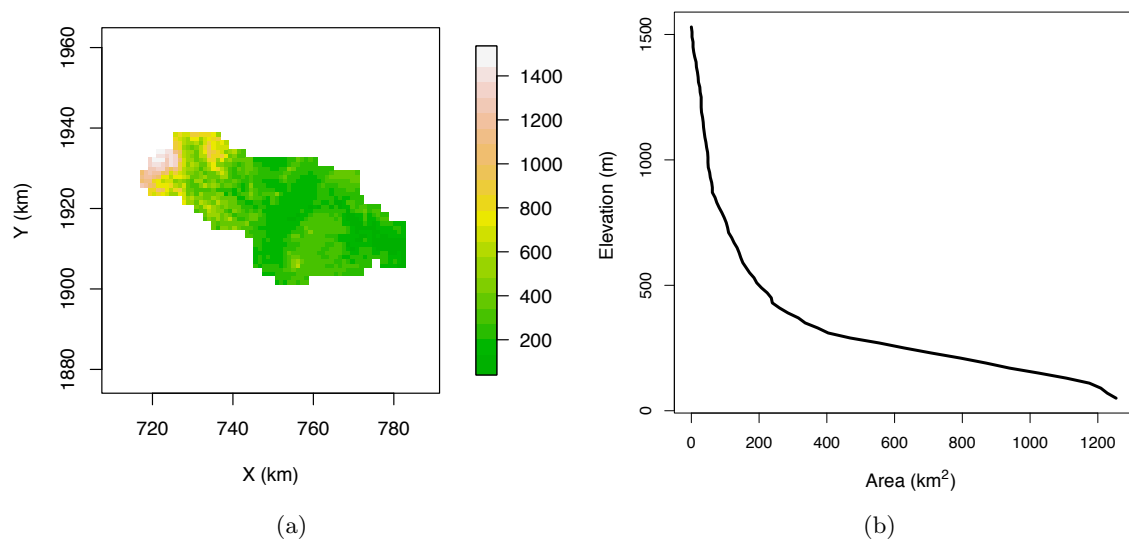


Figure 1.9: Ceze basin. a: Elevation map. b: Hypsography.

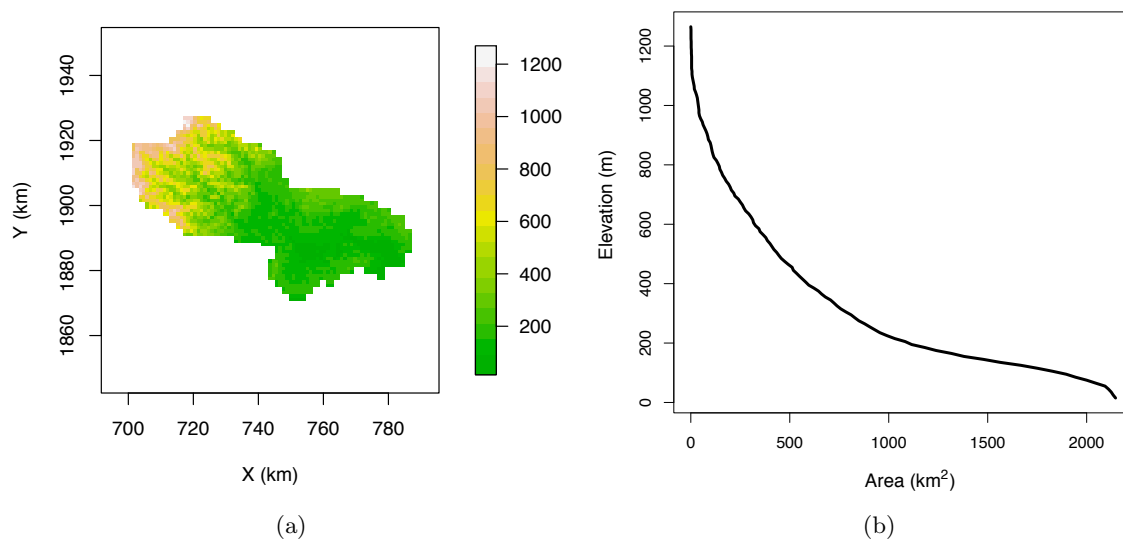


Figure 1.10: Gard basin. a: Elevation map. b: Hypsography.

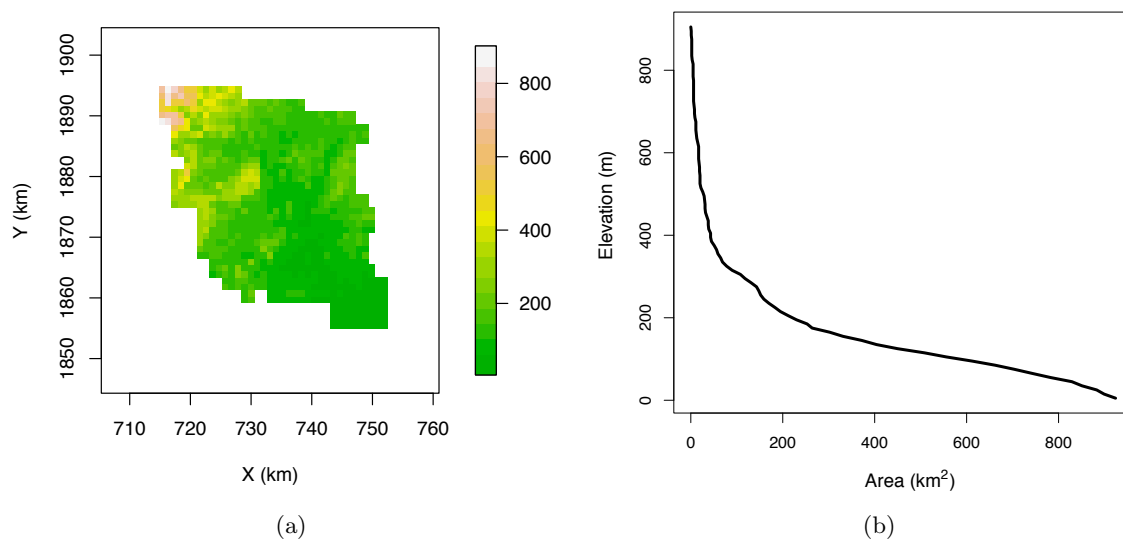


Figure 1.11: Vidourle basin. a: Elevation map. b: Hypsography.

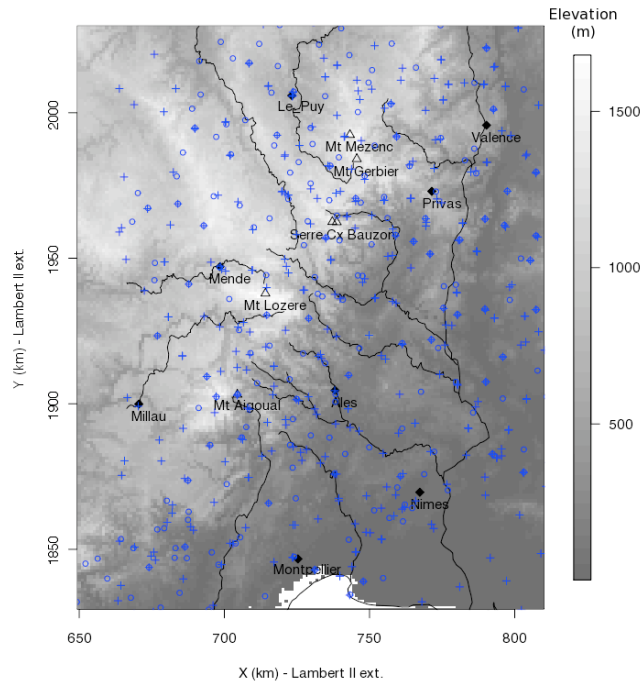


Figure 1.12: Rain gauge networks in the region of analysis: elevation above sea level (gray scale), hydrographic network on the right of Rhône River (solid line), the five highest mountain peaks (triangles), main cities (diamonds), the daily (circles) and the hourly (crosses) rain gauge network.

1.5.1.1 Daily Rainfall

The daily rainfall database is provided by Météo-France and covers the period 1958-2000. About 400 rain gauges are available, and among these, 225 rain gauges feature more than 30 years of data. Some rain gauges are outside the OHMCV window; they are kept to provide reliable estimations at the region boundaries.

The hourly series for the period 2000-2008 have been aggregated to 24 hours in order to extend the daily series to the period 1958-2008. Two stations are gathered when:

- the horizontal distance is lower than 2 km;
- the difference in elevation does not exceed 100 m.

When two close rain gauges were operational at the same time, we take the average of the two stations, increasing the reliability of the measurements. Knowing that elevation has a strong influence in the determination of the pluviometric regime, the vertical distance has been taken as 100 m.

This merging allows to increase the sample size of 75 stations, for which we have 51 years of data.

1.5.1.2 Hourly Rainfall

Three different hourly databases are available; the 3 databases consistently differ among each other.

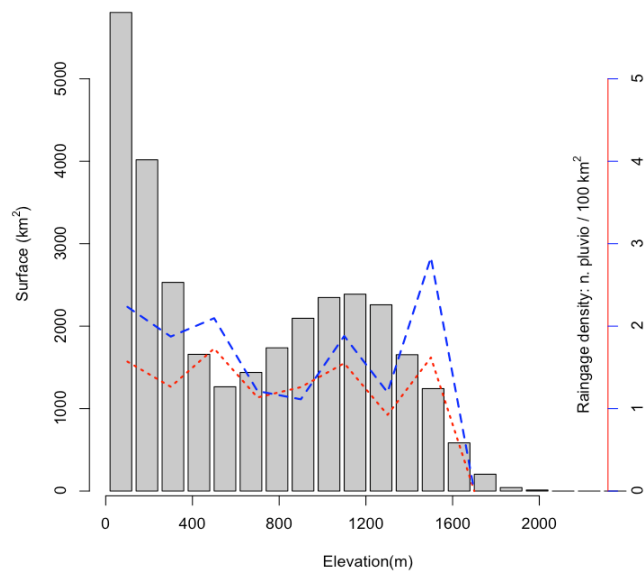


Figure 1.13: Histogram of the surface elevation (grey boxes, left vertical axis); rain gauge density as function of the elevation. The rain gauge density is expressed as the ratio of the rain gauge number to the area of the relative altitude band: long dashed lines for the daily rain gauge network and dashed-line for the hourly one (right axis).

- i. 1972-1992: event-based database in the Cévennes region. The hourly rainfall has been recorded at about hundred rain gauges in a discontinuous manner, mainly in the fall season (more precisely, from August, 15th, to December, the 15th);
- ii. 1993-2000: continuous database provided by OHMCV;
- iii. 2000-2008: continuous database provided by Météo-France.

The database 1972-1992 has been used for long time with the aim to estimate the frequency of rainfall extremes (*Bois et al. (1997); Lebel et al. (1987); Lebel and Laborde (1988)* among others). The rain events occurring during the fall season were recorded, since fall was considered the only season featuring heavy rainfall events. Nevertheless, it is not possible to exclude a priori that some extreme meteorological events could occur in late spring: the frequency estimation could be affected by this assumption (some evidences of this are shown in *Molinié et al. (2010)*).

1.5.2 Radar-estimated rainfall fields

The use of radar imagery for the analysis of the meteorological events developed after the 2nd World War. It allows to characterize the water content within the scanned volume, and this information can be integrated over a vertical profile to obtain a rainfall accumulation. In addition, information on the spatial structure of the storm can be drawn; worthy insights on the rainfall generation mechanisms have been discovered.

The radar measurement of rainfall is based on the following principles:

- the radar reflectivity is measured as the ratio between the energy received by the antenna and the energy emitted by the source.

- the reflection is proportional to the sixth power of the rain droplets diameter.

Due to its tremendous variability range, the reflectivity is usually expressed in decibel, an adimensional ratio of the reflectivity with respect to the reference $z = 1 \text{ mm}^6 \text{ m}^{-3}$:

$$Z = 10 \log_{10} \left(\frac{z(\text{mm}^6 \text{mm}^{-3})}{1(\text{mm}^6 \text{mm}^{-3})} \right) \text{ dBZ} \quad (1.1)$$

The state of the precipitation sensibly affects the reflectivity. It is fundamental to detect if the precipitation is liquid or solid. In addition, to simplify the relation between droplets diameter and water volume, the droplets are usually supposed spherical. Other issues related to the rainfall detection are due to the phase change of precipitation. At high levels in the atmosphere, precipitation appears as snowflakes. Falling to lower levels, the snowflakes melt developing a water coating. Water being an order of magnitude more reflective than ice, large wet snowflakes have large reflectivity (“bright band” effect).

A formula expressing the reflectivity to the rainfall intensity (usually referred to as “Z-R” relationship) is used, conditional on the precipitation type. It has the form:

$$R = aZ^b \quad (1.2)$$

where R is expressed in mm h^{-1} , and a, b are coefficients usually derived from the Drop Size Distribution (DSD), measured with a disdrometer (Section 1.5.3). Some Z-R relations valid for US are provided in Table 1.2.

To obtain a reliable quantitative estimation, a radar/rain gauge merging algorithm must be applied in order to reduce the quantitative errors. The merging is usually performed by means of geostatistical techniques.

The data used in this study derives from the the radar located in Bollène. The radar scans the same volume once every 5 minutes; for each orientation angle the scan is performed at 8 elevation angles. The radar image is scanned over a polar grid, and the transformation to a 2D grid involves three issues: i) the fact to transform a volumetric measure such as reflectivity in a 2D field; ii) the radar scanned volume increases with the distance between target and radar but the 2D grid size is a constant; iii) the beam elevation changes with the distance (and this leads to scan over several elevation angles to compose an image). Table 1.1 reports the main features of the Bollène radar. In Figure 1.14-a the radar locations and covered ranges in the Cévennes-Vivarais region are shown; the Bollène radar is shown in Figure 1.14-a.

The heaviest event ever occurred since the Hydro-Meteorological survey exists is the September 8-9th, 2002 storm. Of this event, we dispose of a reliable radar imagery sequence, derived from the Bollène radar scan. The event is of large interest for evaluating the structure of heavy meteorological events. An example of two instantaneous 2D fields derived from the radar scan of Bollène is shown in Figure 1.15.

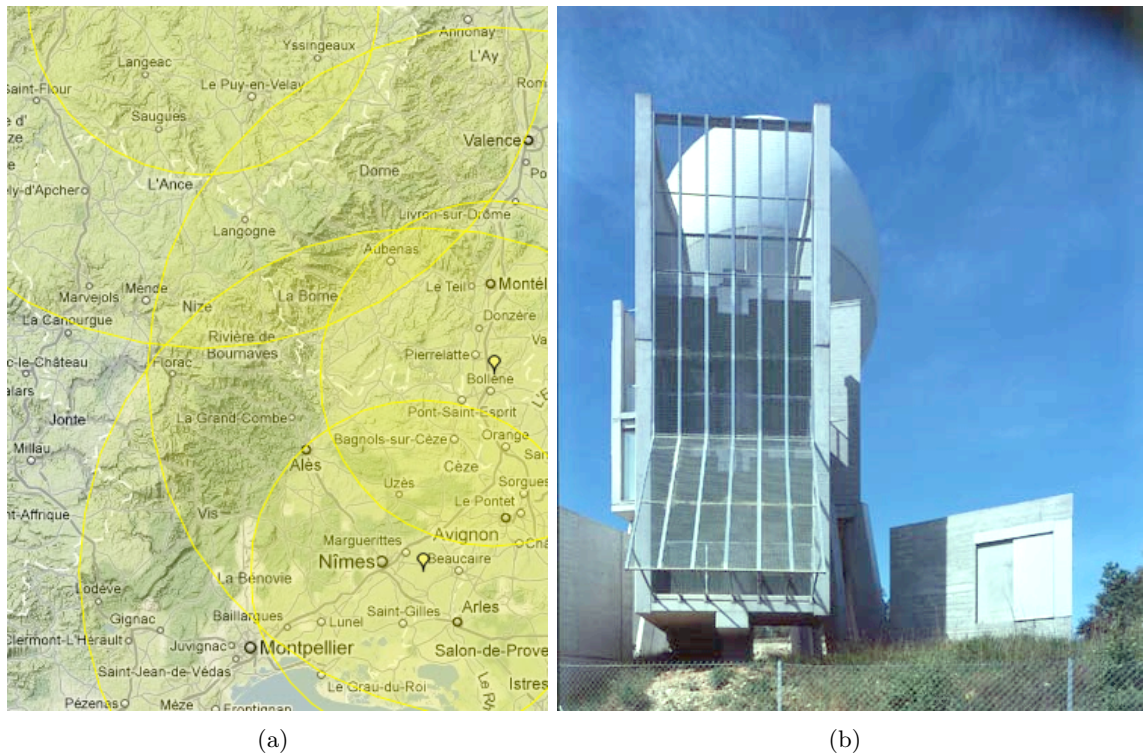


Figure 1.14: a: Radar measurement network in the region of study and relative covered distance. b: Bollène radar.

Table 1.1: Main features of the Bollène radar scan of the event of 2002, 8-9th September.

Radar Name	Bollène
Administration Name	Météo-France
Location (Lambert II)	X=793.658 km; Y=1927.770 km
Location (Polar Coordinates)	Lat=44.32°; Lon=4.76°
Elevation	327 m
Emitted Power	600 kW
Frequency	2.80 GHz
Pulse Length	2 μ s
Power Gain	42.5 dB
Opening 3dB	1.28°
Number of Elevation Angles	8
Revisit time	5 min

Table 1.2: Recommended Z-R relationship from NOAA - US National Oceanic and Atmospheric Administration

Relationship	Recommended for
$Z = 200R^{1.6}$	General stratiform precipitation
$Z = 75 - 130R^2$	Winter stratiform precipitation - Orographic Rain
$Z = 300R^{1.4}$	Summer deep convection
$Z = 250R^{1.2}$	Tropical convective systems

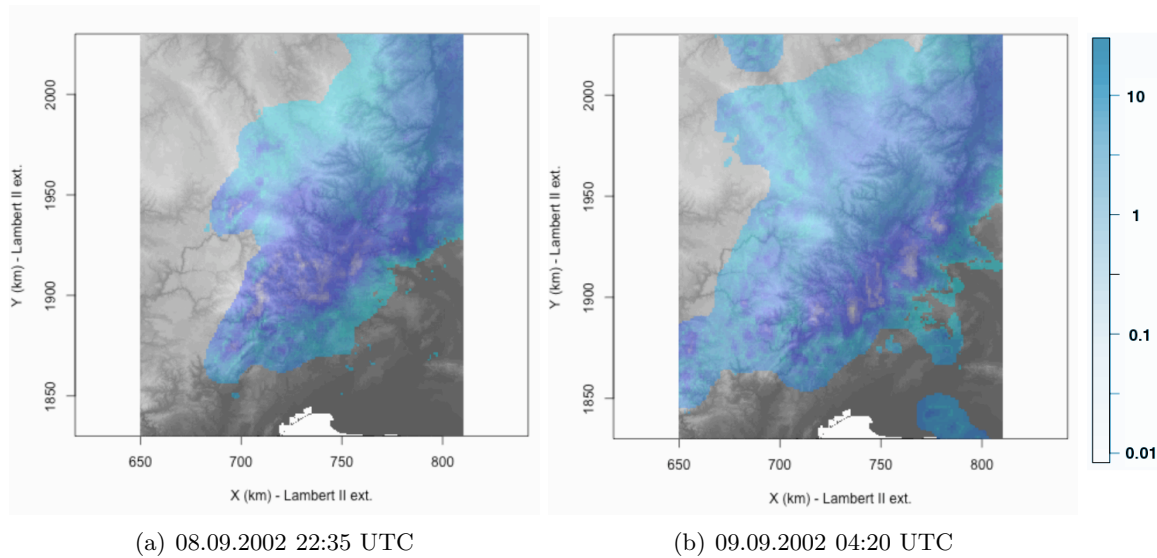


Figure 1.15: Instantaneous radar-rainfall scan (Bollène radar) for the event of 2002, September the 8-9th. a: Instantaneous scan, September the 8th at 22:35 UTC. b: Instantaneous scan, September the 9th at 04:20 UTC. The rainfall intensity values ($mm \cdot h^{-1}$) have been logarithmically transformed for the plot.

1.5.3 Disdrometer data


The rain gauge demonstrated reliable for the measure of point-rainfall at durations in the range 10 min - 24 h (Marsalek, 1981). For smaller durations, the rain gauge can exhibit problems in the measure of the rainfall intensity due to the extreme rainfall variability, and the rain gauge can be efficiently substituted by an optical disdrometer (Krajewski *et al.*, 2006). In addition, the drop size distribution (DSD) provided by disdrometer is of central importance for the radar calibration: the relations of reflectivity factor (Z) and rain rate R (Equation 1.2) significantly change with DSD.

The laser optical disdrometers (Figure 1.16) consists in a laser beam of few cm^2 that scans a volume of few dozens cm^3 . A receiver measures the proportion of the emitted signal that is reflected by the rain drops. This value is proportional to the water volume crossing the beam. The disdrometer can measure the fallen rainfall amount, the intensity as well as the particle size (down to 0.1 mm) and the velocity of precipitation.



Figure 1.16: Laser optical disdrometer for the measure of the Drop Size Distribution (DSD).

Uncertainties in the extreme rainfall measurement



Résumé

Au début du XXe siècle, la mesure du taux de précipitations commence à susciter l'intérêt dans la communauté scientifique internationale. Grâce aux premiers dispositifs électromécaniques, les précipitations peuvent être mesurées automatiquement.

*La thèse repose principalement sur la mesure des précipitations obtenues par **pluviomètres à auget basculant**. Avant d'entreprendre toutes analyses de ces données, il est recommandé de vérifier que la mesure des extrêmes par ce type de dispositif soit fiable tant pour les précipitations ponctuelles que pour l'estimation des précipitations spatiales.*

Le pluviomètre à auget basculant (Figure 2.1) se compose d'un entonnoir de section donnée qui collecte la précipitation pour remplir un des 2 augets. Une fois le premier auget rempli, l'auget tourne, le système d'augets bascule, fermant ainsi un contact électrique. L'information fournie par l'impulsion électrique est enregistrée soit sous forme graphique (avant les années 1980), soit sous forme électronique.

La création des réseaux pluviométriques a été la première étape pour la connaissance de l'étendue spatiale des épisodes de pluies: la collecte d'informations a permis de connaître la hauteur de pluie tombée dans un bassin.

La mesure de pluie au sol est encore la méthode d'estimation la plus fiable pour la mesure des précipitations. Néanmoins, une série de sources d'incertitudes cause un écart entre les précipitations mesurées et les précipitations réelles. Pour être en mesure de détecter et de éventuellement corriger ces biais, chaque pluviomètre doit être caractérisé en termes de:

- **résolution d'échantillonnage** des données enregistrées: l'étude peut donner des résultats erronés si la résolution d'échantillonnage des données est proche de l'échelle souhaitée pour l'analyse. Pour analyser les extrêmes horaires, par exemple, une série de précipitations de résolution de 6 minutes pourrait être suffisante.

- *erreur maximum de la mesure d'intensité*: chaque pluviomètre a une gamme d'intensités de précipitations pour laquelle l'erreur de mesure est inférieure à 1-3%. En dehors de cette plage, la mesure est fautive et ainsi l'analyse des extrêmes.
- *Saturation de l'entonnoir*: L'entonnoir atteint la saturation lorsqu'il reçoit une intensité de précipitations supérieure à une certaine valeur (par exemple, 150 mm h^{-1}): telle intensité peut apparaître considérable à la résolution de l'heure, mais elle est facilement atteinte à la résolution 6-min (Section 2.3). La mise en charge du collecteur peut engendrer une réduction de l'intensité de pluie mesurée qui ne peut pas être corrigée.

Enfin, pour ce qui est de l'évaluation de la sous-estimation spatiale des précipitations et éventuellement de sa correction, il faut aussi connaître la **densité du réseau pluviométrique**.

Le chapitre est organisé comme suit: la Section 2.2 est consacrée à une brève introduction sur la détermination des erreurs dans la mesure des processus ponctuels en fonction de la résolution d'échantillonnage. La Section 2.3 présente une description technique du fonctionnement du pluviomètre à auget: grâce à une simulation numérique qui consiste en un processus de désagrégation des pluies, nous estimons le biais induit par le dispositif de mesure sur l'estimation correcte des précipitations extrêmes et sur les cumuls de précipitation. On propose une méthode pour corriger les données historiques et des directives sur les caractéristiques idéales des nouveaux appareils. La section 2.4 est dédiée à l'analyse des incertitudes liées à l'estimation spatiale de la pluie. Nous montrerons que l'estimation correcte de la hauteur de pluie est fortement biaisée lorsque la densité du réseau pluviométrique est du même ordre de grandeur que la structure de corrélation des phénomènes pluvieux. Il apparaît aussi que, quel que soit le processus d'interpolation, la hauteur maximale des précipitations est systématiquement sous-estimée.

2.1 Introduction

In the beginning of twentieth century the measure of the rainfall rate begins to gain interest in the international scientific community. With the advent of electro-mechanical equipment, the rainfall record could be automatized.

The thesis mainly relies on rainfall measurements obtained through **tipping-bucket rain gauges**. Before to undertake any study on this kind of data, it is recommended to verify the reliability of tipping bucket rain gauge data for the study of spatial and temporal properties of rainfall intensities.

The tipping-bucket rain gauge (Figure 2.1) consists of a collector of known area forcing the precipitation to fill a bucket. Once the first volume full, the bucket rocks. The tipping date is recorded either on a rotating paper (up to 1980) or on an electronic device. Tipping bucket rain gauges are equipped with two twin buckets, so that the second bucket is presented to the flux after the first has tipped.

The rain gauge is still the most reliable estimation method for the ground measurement of rainfall. Nonetheless, a series of uncertainty sources cause the measured rainfall to show deviations with respect to the actual rainfall. To be able to detect and possibly correct these biases, each single rain gauge must be characterized in terms of:

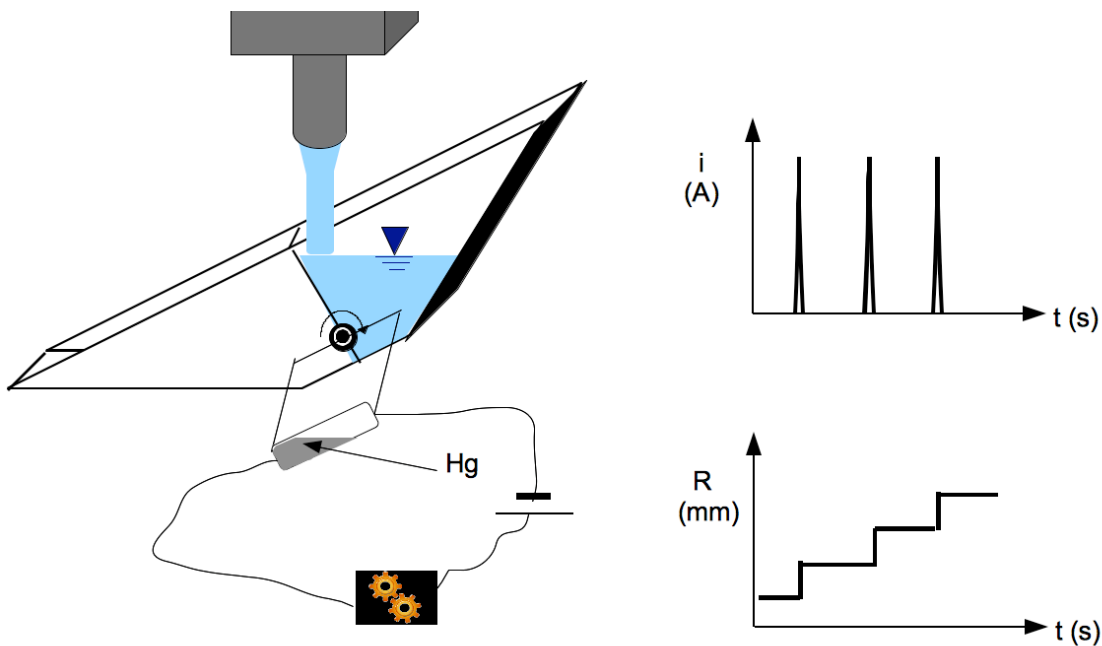


Figure 2.1: Schematic diagram of the electro-mechanical tipping-bucket rain gauge. The right bucket is being filled with rain water. During the tipping, an electrical current impulse (upper right panel) is traduced into a rainfall amount as shown in the lower right panel. The mechanism generating the impulse is the ampoule oscillation, during which the mercury puts in contact 2 electrodes.

- **sampling resolution** of the recorded data: strong underestimations of the extreme rainfall intensities occur if the data resolution is close to the desired scale of analysis. For the analysis of hourly extreme intensities, for example, a rainfall series at the resolution of 6-min could be adequate.
- **maximum sampling error**: each rain gauge has a range of rainfall intensity values for which the bias is lower than 1-3%. Outside this range, the measure is not reliable.
- **saturation of the rain gauge collector**: when receiving very intense rainfall amounts (e.g. higher than 100 mm min^{-1}), the rain gauge collector may saturate preventing the exact measurement of the instantaneous intensity. This rainfall intensity value may appear tremendous at hourly resolution, but it could be easily attained at the 6-min resolution (Section 2.3). A water charge in the collector may appear, laminating the output flow and, therefore, generating underestimation of the measured rainfall intensity.

The points above detailed concern point rainfall. The spatial rainfall estimation is submitted to additional errors. The underestimation committed measuring spatial rainfall underestimation depends on the ratio between **rain gauge network density** and **storm characteristic size**.

The chapter is organized as follows: Section 2.2 is devoted to a brief introduction on the determination of a proper scale of analysis of point rainfall intensities as a function of the sampling frequency. Section 2.3 presents a technical description of the functioning of the tipping-bucket rain gauge: through a numerical simulation consisting in a rainfall disaggregation process, we will estimate the bias induced by the measurement device on extreme rainfall and rainfall depth measurements.

Section 2.4 extends the analysis of uncertainties to the spatial estimation of rainfall: we will show that the correct estimation of the rainfall depth is strongly biased when the rain gauge density is of the same magnitude order of the correlation structure of the fields or lower. We will demonstrate that, whatever the interpolation process, the maximum rainfall depth is systematically underestimated.

2.2 Role of the sampling frequency on the statistics of extremes

The sampling frequency of the measure has wide importance in the determination of the statistics of extremes. The longest rainfall time series are those at the daily resolution (an operator each day at 6 a.m. recorded the rainfall accumulated during the last 24 h). These series allows to estimate extreme quantiles with a good precision. On the other hand, any analysis on daily extremes conducted using this data is submitted to a systematic underestimation, due to the identity between the scale of the analysis and the sampling frequency. The arbitrary choice of the hour of the day at which the measure is recorded causes the events occurring in the first hour of the day to be cut. In Section 2.2.2, we analyze this underestimation source in detail.

2.2.1 Moving-window and fixed-window sampling

We must define here the concepts of fixed-window record and moving-average record. In Figure 2.2 the sampling schemes for fixed-window and moving-average are shown for the aggregation process of a 6-h series to a 24-h series.

Depending on the information we want to extract from the data, one sampling technique is preferable respect to the other.

- **fixed window:** a window of size equal to the scale of analysis is moved of a step corresponding to the scale of analysis. The fixed window sampling must be used when the aim is drawing statistics on the whole sample, such as average, standard deviation or intermittency. The observations in this case are not overlapping. The daily rainfall measurements at a rain gauge can be seen as a 24-h fixed window sampling.
- **moving average:** a window of size equal to the scale of analysis is moved of a step corresponding to the sampling frequency. On this kind of data, no overall statistics can be computed on the series due to the partial overlapping of the data (high interdependence of data). Nevertheless, the sample is useful for the extraction of maxima/minima over a block of fixed size. In the example of Figure 2.2 it appears that the fixed-window sampling did not detect nor the actual maximum (40) neither the minimum (18) at 24 h, correctly detected with the moving window sample.

2.2.2 Uncertainties associated to the measurement resolution

In this section we aim to quantify the uncertainty associated with the measurement resolution. Very often, the study of the extremes of time series is done at the scale corresponding to the measurement resolution. For example, the daily rainfall maxima are estimated based on daily rainfall series. Even

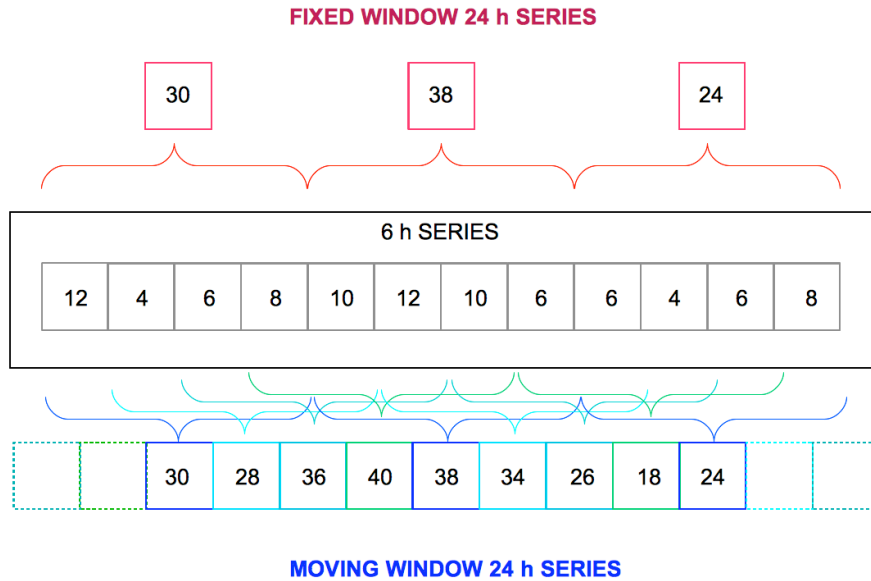


Figure 2.2: Illustration of the fixed-window and moving-window sampling technique for obtaining 24-h data starting from a 6-h database.

though the average and variance of the series are correctly estimated, the statistics concerning extreme observations are systematically underestimated by the fixed window sampling. It is common experience that the extreme rainfall at the duration of 24 h is underestimated of at least 10% if daily time series are used. In the next paragraphs we quantify this underestimation. Section 2.2.2.1 uses the 52-years hourly series of Montpellier to assess the underestimation of the extreme behavior of the series due to resolution issues. In Section 2.2.2.2 we extend the analysis to the whole rain gauge network. The absence of long hourly series prevent this computation. To compute the global underestimation due to sampling issues we reconstruct 4-hour data from daily series, by applying a disaggregation scheme. The results of the two sections are comparable, leading to an average underestimation of extreme quantiles of 13 %.

2.2.2.1 Assessment on real data

We can evaluate the systematic underestimation induced by fixed-window sampling by analyzing the rainfall series of Montpellier, in which hourly rainfall has been measured for over 50 years. By subsequent aggregation of the data, we may determine the underestimation error committed in determining the rainfall intensity maxima at the scale $D=24h$ based on the data at the resolution of $\lambda_{sample} = 1, 2, 4, 8, 12$ and 24 h, respectively.

In Figure 2.3-a we plot the expected value of the annual maxima for $D=24h$ at various sample resolutions, divided by the 24-h maxima evaluated at the 1-hour resolution in order to show the underestimation induced by the sampling resolution. A clear underestimation of the maxima average (higher than 10%) is found when the sample resolution is close to the resolution of analysis. The diamonds show the empirical standard deviation of annual maxima for $D=24h$ as a function of the

sample resolution λ , normalized by their maximum, obtained for $D=1\text{h}$. Even in this case, the lower the ratio between the scale of analysis and the sample resolution, the higher the underestimation.

The first two empirical moments of rainfall intensity maxima (average and standard deviation) are the needed parameters for a first approximation of the extreme rainfall return levels (see details in Section 3.5). We show in Figure 2.3-b that the sampling resolution seriously affects the estimation of these first two moments: the estimation based on daily data leads to an underestimation of 13% of the 100-year return level for the 24-h rainfall.

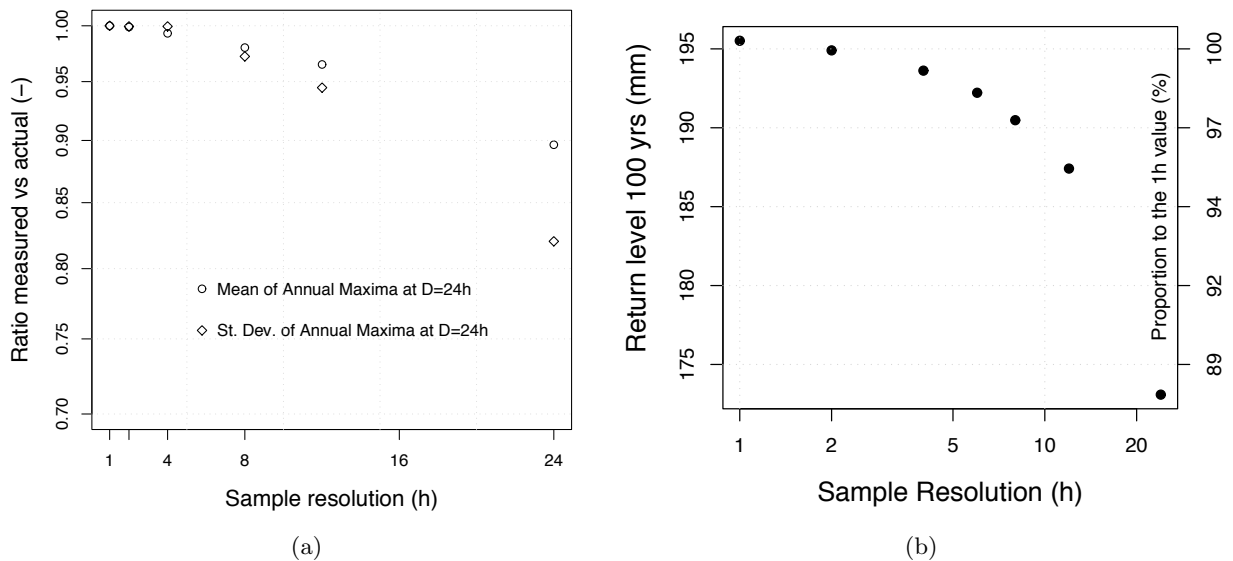


Figure 2.3: a: Effects of sampling resolution on the statistics of daily annual rainfall maxima. White circles: ratio between average daily-rainfall maxima computed with degraded resolution and computed at the sample resolution of 1 h. White diamonds: ratio between standard-deviation of daily-rainfall maxima computed with degraded resolution and computed at the sample resolution of 1 h. b: Effects of the rainfall sampling resolution on the extreme value statistics for the 24-h duration (return level rainfall depth for 100 years).

This is a clear example of how statistics even on apparently well-known variables such as daily rainfall can be perturbed by resolution issues.

2.2.2.2 Assessment thanks to self-similar cascades

Anticipating one of the scale-invariance applications for meteo-hydrology, we use self-similar cascades (Section 5.12) to test the effect of fixed-window maxima extraction at the daily scale. We implemented a disaggregation scheme (*Over and Gupta, 1996; Schmitt et al., 1998*) whose parametrization is based on the scale-invariant series of daily data in the range 1 h - 7 days. Our objective is to generate series at 4-hours resolution and then compute the annual maxima of the daily series obtained using a moving-average scan. The fixed-window and moving-average daily maxima will then be compared in their main indicators, average and standard deviation, in order to quantify systematic underestimation due to the fixed-window data sampling.

The analysis is performed on the 225 daily rain gauge stations featuring at least 30 years of continuous observations in the period 1958-2008.

Figure 2.4 shows a comparison of the fixed-window and moving average sampling methods through a quantile-quantile plot. The graph shows a linear relation between fixed-window and moving-average annual maxima. Table 2.1 gives the regression features, indicating that the correlation coefficient R^2 is very close to the unit. Figure 2.4-b reports the relation between the empirical standard deviation computed by fixed-window and moving-average, respectively. The underestimation of the actual average as well of the actual standard deviation due to the fixed-window sampling is of about 13%. This result is in agreement with the underestimation of rainfall extremes computed on the observed rainfall series of Montpellier 2.2.2.1.

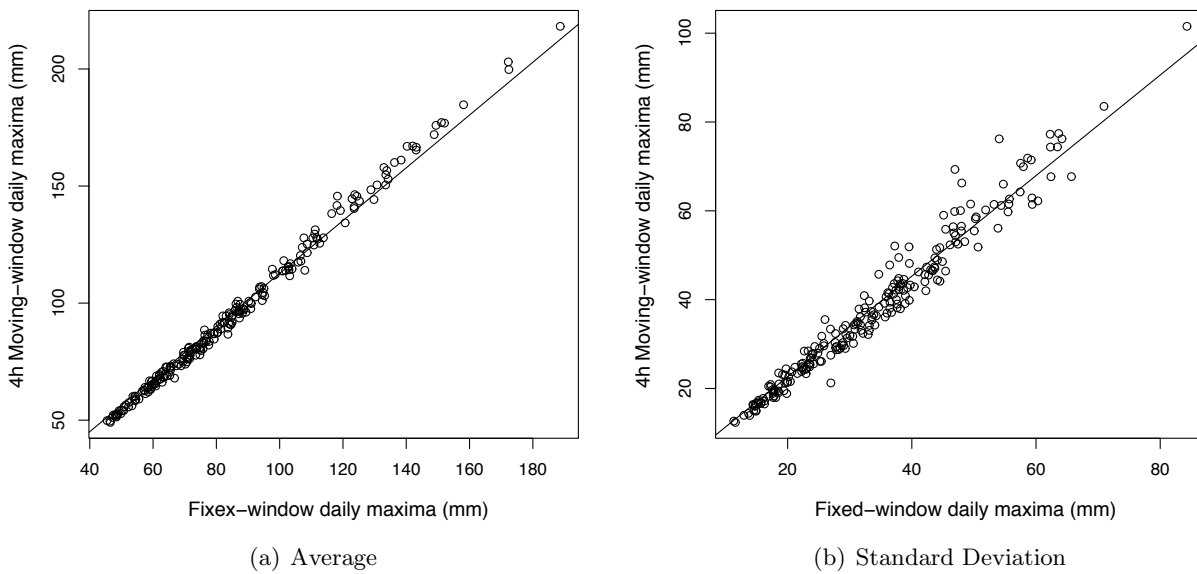


Figure 2.4: Effect of the rainfall sampling resolution on the statistics of annual maxima. a: q-q plot of the average annual-rainfall maxima at the daily scale. b: q-q plot of the empirical standard deviation of annual-rainfall maxima.

Table 2.1: Relation between fixed and moving window daily annual maxima in terms of sample mean and sample variance.

Linear regression formula	R^2
$E[x]_{mov} = (1.127 \pm 0.003) \cdot E[x]_{fix}$	0.999
$s_{mov}^2 = (1.132 \pm 0.006) \cdot s_{fix}^2$	0.993

2.3 Tipping-bucket rain gauge measurement

In preparation for Journal of Hydrology

Optimal design and correction of operational tipping-bucket rain gauges: a numerical study

Ceresetti, D., Molinié, G., Creutin, J.D.

LTHE, Laboratoire d'Etudes des Transferts en Hydrologie et Environnement - (CNRS,UJF,IRD,INPG), Grenoble, France

Abstract

Keywords: tipping-bucket, pluviometer, raingauge, rainfall variability, small-scale, scale-invariance, extreme rainfall, measurement, ground rainfall.

1. Introduction

The tipping bucket rain gauge is a mechanical system for measuring rainfall intensities. The device basically intercepts the rain flux and drives the collected water to monitored tipping buckets recording the time of occurrence of rainfall quanta. The number of tippings in a given period of time (a few minutes to a day) is proportional to the average rainfall intensity.

Despite the apparent simplicity of the system, the measurement of rainfall rates or accumulations is subject to many errors. As pointed out by Molini et al. (2005a), the measurement is basically affected by a deficit of rain water entering the collector due to wind, evaporation (Sevruk, 1972) or drop splash effects, besides the issue of solid precipitation. The evaluation of this kind of errors is out of the scope of the present paper.

Our study focuses on the specific errors linked to the use of tipping-buckets. Several antecedent studies (Calder and Kidd, 1978; Niemczynowicz, 1986; Sevruk, 1987, 1996; Humphrey et al., 1996; Wood et al., 2000; Habib et al., 2001; Ciach, 2003) analyzed the systematic underestimation of rainfall rates due to the principle of measurement of tipping-buckets. They show that for high intensities a correction is needed to reduce the measurement bias resulting from the discretization of the rain flux. The indoor calibration of the gauge for selected rainfall rates is a reliable way to establish the amplitude of this bias. Nevertheless, this type of calibration is not sufficient for the proper correction of tipping bucket errors in real conditions when the rainfall rate varies at short time scale. In order to characterize and subsequently correct the response of a tipping-bucket gauge we propose a numerical simulation approach following the path proposed by Molini et al. (2005b). We use statistical cascades to build high-resolution time series of rainfall rates that we consider to be the true rainfall. To compute the corresponding tipping-times, we simulate the functioning of the tipping-bucket including in the computation the effects of the water storage in the collector. Subsequent aggregations yield series of measured intensities and accumulations. According to the same authors, the lack of a proper correction of the measured rainfall can result in a global underestimation of the extreme quantiles. They estimate that, with a simple linear correction, the 100-years return period rainfall is underestimated by 45-65% and 25-40%, respectively for hourly and daily rainfall.

Email address: jean-dominique.creutin@ujf-grenoble.fr (Creutin, J.D.)

A detailed analysis of the biases induced by the small-scale rainfall variability on the total accumulation and on the estimated rainfall intensity is therefore needed to understand how to correct historical series and how to realize an optimal design of rain gauges.

The paper is structured as follows. In Section 2 we describe the principles of the rainfall estimation by tipping-bucket rain gauge and we detail on the rain gauge features. Section 3 presents the analyzed rainfall series and the simulation method that will be used for the disaggregation. In Section 4 we present three formulas for the correction of the raw rainfall measurements. The results of the correction application to the simulated data series are presented in Section 5. Section 6 provides some further elements that will serve as guidelines for historical data correction and for future rain gauge dimensioning.

2. Tipping-bucket measurement errors

The sampling mechanism of tipping-buckets leads to several errors in the measurement of the rainfall intensity. Concerning small to moderate rainfall rates, the main tangible sampling error is when the active bucket remains partially filled at the end of a rain event. The corresponding rainfall amount is then allocated to the subsequent rainfall event, unless the water stored into the bucket evaporates, in which case the rainfall amount is lost. The resulting error is less than the rainfall quantum corresponding to a full bucket, but it can turn to a few tens of mm in terms of annual accumulation for instance. Concerning high rainfall rates (say more than $20 \text{ mm} \cdot h^{-1}$) the problem of underestimation is more critical and mainly related to the lack of measurement during the tipping of the buckets. Every time the tipping volume is reached, the device rocks under the weight of the full bucket and presents the empty bucket to the water flux coming from the collector. During the corresponding period of time the water flux coming from the collector is partly lost leading to the above-mentioned underestimation. In this study we try to describe and model this last category of problems, related to the measurement of high rainfall rates.

The bias due to the tipping time is a function of the actual rainfall intensity I_a , of the tipping time Δt and of the nominal rainfall depth quantum per tip h_n which depends on the ratio between the tipping volume and the surface of the collector. According to Marsalek (1981) the bias can be expressed as:

$$\frac{I_r}{I_a} = \frac{h_n}{h_n + I_a \Delta t}, \quad (1)$$

where I_r is the recorded rainfall rate.

Usual values of the bucket volume and the collector surface give values of h_n ranging from 0.1 to 0.5 mm. Reducing the collector surface may decrease water losses but this solution degrades unacceptably the precision of the gauge at low and moderate rain rates. Keeping the flow from the collector below a chosen threshold may also limit the water losses. In many rain gauges the size of the bottom hole of the collector is designed to evacuate rainfall rates up to, for instance, $150 \text{ mm} \cdot h^{-1}$. Beyond this value, a water height establishes in the collector and the output flow is laminated. This solution improves the measurement of the rainfall accumulation over long periods but clearly affects the measurement of high intensities at small durations. The tipping time Δt depends on several factors. It is basically related to the distance travelled by the buckets to rock from one position to the other one. It depends on the angular velocity of the buckets during the tilt that is sensitive to the quality of the bearing mechanism and to the pressure of the incoming water flux on the buckets. This last factor depends on the position of the water flux in regard of the barycenter of the bucket. Conscious of this issue, some manufacturers designed a small device located above the bucket that reduces the tipping time by increasing the momentum generated by the water flow (see Figure 1).

Molini et al. (2005a) highlight that, in Equation 1, the actual rainfall is equal to the measured rainfall either if $I_a = 0$ or if tipping time $\Delta t = 0$. While the latter condition is far to be achieved, the former condition implies that only the measurements of null rainfall are unbiased. In other words, all the positive rainfall values are underestimated. We see in Section 4 that a slight modification of Equation 1 allows to set the optimal functioning of the gauge at a rainfall value higher than zero.

In our study the simulated rain gauge is of the tipping-bucket model, with a measurement resolution of 0.2 mm. The reversal time has been estimated at 0.2 s, averaging the experimental values obtained with a video camera able to capture 1200 frames per second. The rainfall falling within a circular surface of 0.1 m² is poured into a container with a conical hollow bottom. The bottom hole diameter is 50 mm long. The geometry of the conical bottom is approximately such as the height equals the diameter. A series of indoor experiments with constant rainfall intensity has been carried out in order to define the minimum rainfall intensity needed to put in charge the collector. According to these measures, for rainfall intensities higher than 150 mm · h⁻¹ the collector starts laminating the flow. For these intensities and above, our numerical simulation uses a time step of 1 s to compute the height of water contained in the collector and the corresponding outflow, determined through the Torricelli's law:

$$Q = C_c A \sqrt{2 \cdot g \cdot h} \quad (2)$$

where Q is the flow in m³ · s⁻¹, A the bottom hole area (m²), g the gravity constant 9.81 m · s⁻², h the water height (m) and C_c an adimensional term expressing the reduction of the output flow area due to the . In our case the term C_c must be empirically estimated since it takes account of the pressure drop due to the pollen filter located on top of the hole. Our empirical estimation yields $A \cdot C_c = 0.173 \text{ cm}$.

3. Data set used and rainfall disaggregation by bounded cascade

The rain gauge data set used in this study is from the station of Montpellier Bel-Air, among the longest hourly rainfall series in France. The hourly series has been obtained by digitization of the analogical pluviographic measurement. The considered period covers the years 1920-1972, in which the station never moved and very rarely was malfunctioning. The rainfall resolution is 0.2 mm. The maximum hourly rainfall ever recorded at this station is 69.7 mm · h⁻¹. The average annual rainfall is 715 mm, corresponding to 0.082 mm · h⁻¹. The variability of rainfall at hourly resolution is high: the ratio between maximum and mean hourly rainfall is 850.

To reproduce the quasi instantaneous rainfall variability we implemented a downscaling method similar to those proposed by Over and Gupta (1996) or Schmitt et al. (1998). These scale-invariant stochastic models for rainfall time series are based on the separate modeling of rainfall intermittency and of rainfall intrinsic variability. As illustrated by Figure 2-a for the series of Montpellier, the scale-invariance of intrinsic rainfall is verified in the range 2 - 100 h. In this range the statistical moments are aligned, but for durations lower than 1-2 hours, the process deviates from scale invariance. Therefore the statistical disaggregation through a self-similar cascade can not be implemented and we are led to model the range in which the process is not self-similar by a bounded cascade in which the rainfall variability depends on the scale (Menabde, 1998).

For a sound downscaling, not only the rainfall depth but also the rain intermittency should be rescaled. As for intrinsic rain variability, rainfall intermittency (showed in Figure 2-b for the rainfall series of Montpellier) is scale-invariant in the range 2-100 hours, but deviates from scale invariance for durations lower than 1 hour.

The cascade model we propose uses the parametrization in Veneziano and Furcolo (2002) fixing $\alpha = 2$, so that the cascade is log-normal with an atom at zero:

$$K(q) = C_\beta(q - 1) + C_1(q^2 - q) \quad (3)$$

where $K(q)$ is the moment scaling function, q the moment order, C_β and C_1 the two model parameters describing the fractal dimension of the support (the intermittency) and the mean codimension of the process, respectively. The variance of the series at each step n of the cascade is related to the variance of the series at the step $n - 1$ through the scale relationship

$$\sigma_n^2 = m^{K(2)} \cdot \sigma_{n-1}^2 \quad (4)$$

where m is the multiplicity of the cascade (i.e. the number of elements at scale n that originate from one element at scale $n - 1$).

The parameters for the series of Montpellier have been determined month by month by applying the method described in Schmitt et al. (1998) and are reported in Table 1.

For each step n of the cascade, we compute the intermittency $P(I_n > 0)$ based on the coefficient C_β as:

$$P(I_n > 0) = P(I_{n-1} > 0) * m^{-C_\beta} \quad (5)$$

Consequently, the ‘‘inner’’ intermittency, i.e. the average proportion of zero observations obtained applying the disaggregation to a positive rainfall value is

$$P(I_n = 0|I_{n-1} > 0)_{inv} = 1 - [P(I_{n-1} > 0) - P(I_n > 0)] \quad (6)$$

where the index *inv* indicates that the relation is valid for scale-invariant intermittency process.

For durations lower than 1 hour we model the deviation from scaling by introducing a correcting term into the inner intermittency definition, depending on the scale d_0 where inner intermittency is supposed to disappear (in our case we fixed $d_0 = 1$ s):

$$P(I_n = 0|I_{n-1} > 0)_{dev} = P(I_n = 0|I_{n-1} > 0)_{inv} \cdot \frac{\log((d/d_0))}{\log((D/d_0))} \quad (7)$$

where D is the larger scale of the cascade. The index *dev* indicate deviations from scaling.

We disaggregated the hourly rainfall series of 52 years recorded in Montpellier down to the 15 second time step following the above showed disaggregation scheme. Figure 3-a and Figure 3-b display the moment analysis of the disaggregated series in terms of intrinsic variability and intermittency, respectively. The time resolution of 15 seconds is too coarse to correctly model the water level in the reservoir under extreme intensities. For this reason, the 15 s rainfall intensity is equally spread to the 1 second resolution needed for the simulation.

We simulated the rain gauge functioning over the 52-years, for a total of 30 different stochastic simulations. We stored a number of variables. The actual rainfall is directly produced by the stochastic disaggregation. We also archived the lost rainfall (i.e. that has been lost during the rocking time), the water height in the collector and the corresponding stored volume.

The measured rainfall intensity is computed from the number of tipplings derived from the tipping bucket simulation. We used the correction formulas described in the next section to transform measured into corrected rainfall rates. The observations are subsequently aggregated, obtaining series at coarser resolutions (typical values are 5, 10, 15, 20, 30, 45, 60 min).

4. Rainfall correction methods

The rain gauge is usually calibrated by fixing an optimal rainfall value for which the rainfall intensity measure is unbiased. Depending on the aim of the measure, one may have interest in fixing a low or high optimal rainfall value. If the tipping bucket volume is chosen as the nominal rainfall increment per tip, the unbiased rainfall value is $0 \text{ mm} \cdot \text{h}^{-1}$, and Equation 1 describes the relation between actual and measured rainfall intensities. In alternative, the tipping bucket volume of the rain gauge can be optimized for a positive rainfall value (the tipping bucket volume can thus be obtained subtracting from the nominal rainfall depth per increment the rainfall volume fallen during the reversal time at the optimum rainfall intensity). Figure 4 shows the basics to set the rain gauge for measuring an optimum rainfall value.

A generalization of Equation 1 can be proposed for positive optimum rainfall intensities:

$$\frac{I_r}{I_c} = \frac{h_n}{(h_n - I_{opt}\Delta t) + I_c\Delta t} \quad (8)$$

where I_{opt} is the optimum value for which the rain gauge measurement is unbiased, I_c is the corrected rainfall and I_r the recorded value. Obviously when $I_c = I_{opt}$ then $I_r = I_c$.

Independently from the optimum rainfall value, a choice must be done: take the raw measurements, affected by significant under/overestimations, or apply a correction algorithm in order to reduce the estimation errors when the rainfall intensity is far from the optimum value. Three correction methods have been applied to the simulated measurements: two of them are empirical and the third one reflects the measurement principle of tipping-buckets, described by Equation 8.

The linear model is the simplest empirical correction, consisting in multiplying the measured rainfall by a constant to obtain the corrected rainfall:

$$I_c = \alpha_1 \cdot I_r \quad (9)$$

where α_1 is an empirical parameter. The second empirical correction is a power-law formula which is often referred to as “dynamic correction” because it better fits with the non linearity of Equation 1 linked to the water loss:

$$I_c = \beta_1 \cdot I_r^{\beta_2} \quad (10)$$

where β_1 and β_2 are empirical parameters. This formula better corrects heavy rainfall intensities respect to the linear one, but in some cases its application induces an underestimation of the total rainfall depth.

The parameters (α_1 , β_1 and β_2) in Equation 9 and 10 are estimated through indoor calibration. Using a reservoir with constant water level, constant water flows can be injected into the rain gauge collector, corresponding to given rainfall intensity values. Each value of rainfall intensity is kept for several minutes. During this time, the flow is measured by the tipping-bucket rain gauge and the output collected into a graduate cylinder. When a significant water amount has flown through the rain gauge, the total water volume into the cylinder is compared with the volume corresponding to the number of tippings of the bucket. The operation is repeated for various intensity levels (e.g. from $10 \text{ mm} \cdot \text{h}^{-1}$ to $150 \text{ mm} \cdot \text{h}^{-1}$ with a step of $10 \text{ mm} \cdot \text{h}^{-1}$) until a sufficient number of calibration points are available.

In Table 2 numerically-determined coefficients for the linear and the dynamic correction for the simulated rain gauge are respectively shown, for two values of the optimum rainfall (0 and $50 \text{ mm} \cdot \text{h}^{-1}$).

The third method respects the physics of the tipping-bucket mechanism, and is obtained by inversion of Equation 8:

$$I_c = I_r \frac{h_n - I_{opt}\Delta t}{h_n - I_r\Delta t} \quad (11)$$

The volume for the bucket tipping h_n and I_{opt} can be only approximately set by manually turning the screws in Figure 4. For a reliable estimation of h_n and I_{opt} it is suitable to apply the indoor calibration similarly to the previously detailed correction methods.

The time step at which the correction is implemented is of primary importance. At the 1-min time step it may be frequent that extreme instantaneous rainfall cause the increase of the water level in the collector, leading to lamination of the output flow. In addition, since the rainfall amount falling in 1-min is often lower than the gauge resolution (0.2 mm), significant sampling errors may occur (Molini et al., 2005a). The time step of 5 minutes has a twofold advantage. On the one side, the probability to have lamination of the extreme intensities compared to the 1-min time step is very low and, at the same time, the correction can be applied discriminating more than 60 discrete values within the range 2.5-150 mm · h⁻¹ (Figure 5-a). The hourly time step is less favorable because the aggregation reduces the average rainfall intensity and gathers heavy and common rainfall event smoothing the rainfall variability. Figure 5-b shows that the sampling is dense for intensities higher than 20 mm · h⁻¹ although very few observations are recorded in this range.

Before to move to the study of the simulations, it is necessary to analyze the relative error of the raw measurements in an idealized case of constant rainfall rate and to quantify the improvement given to the estimation by each of the correction algorithms presented above. By definition, only one rainfall intensity value yields unbiased measurements. Figure 6-a and 6-b display the measurement bias of low to high rainfall rates obtained by calibrating the optimum rainfall value at 0 mm · h⁻¹ and at 50 mm · h⁻¹, respectively.

The application of Equation 11 leads to a perfect correction of the raw measurements (in the ideal case of $I = const \leq 150 \text{ mm} \cdot \text{h}^{-1}$ and accumulation time $T \rightarrow \infty$). The measurement error (for $I = const \leq 150 \text{ mm} \cdot \text{h}^{-1}$) is in the order of 6% imposing the rain gauge optimum at 0 mm · h⁻¹ (dots in Figure 6-c). Setting (as in the most of European gauges) an optimum rainfall value of 50 mm · h⁻¹ reduces the maximum raw measurements error to about 4% (dots in Figure 6-d). Significant improvement of the maximum raw measurement errors are obtained applying the linear and the dynamic correction (the corrections yield maximum errors of 1.5 %, and 0.5%, respectively). The problem with the dynamical correction when the optimum rainfall value is higher than 0 is that all the measurements lower than the optimum are overestimated (Figure 6-d, solid line). Since weak rainfall intensities represent the majority of the total observations, the risk is to have incongruous total rainfall depth estimated by tipping-bucket device respect to the graduate cylinder. The solution is to apply the correction in Equation 9 and 10 only when the optimum rainfall is set to zero.

The linearly and dynamically corrected curves reported in Figure 6-c,d do not show substantial differences imputable to the optimum rainfall choice: the correction provides equal results whatever the choice of the optimum rainfall value.

From the graphs in Figure 6-c,d it is clear that an optimal calibration strategy may differ as a function of the usage of the device. If one is mostly interested in rainfall accumulations and uses linear or dynamic correction, a null optimum rainfall value is recommended. If one is only interested in extreme intensities, one may either fix a zero threshold applying one of the correction formulas or fix a positive threshold using the formula in 11. If no correction is planned, it is better to fix a high optimum rainfall value. If one is interested either in rainfall accumulations or in extreme intensities, the use of the physically based correction formula in 11 is recommended; the optimum rainfall choice does not affect the accuracy of the measure.

Up to this point, the functioning of the rain gauge has been tested with constant rainfall intensities. The response of the device to the natural rainfall variability can be very different from the ideal case of constant

rainfall rates. This point is explored in the next section.

5. Analysis of simulation results

The performed simulations allow to analyze raw and corrected tipping-bucket measurement both in term of “instantaneous” rainfall rates including extreme intensities and long term rainfall accumulations. We will analyze each of these measurements in detail.

We recall that, when the instantaneous rainfall intensity is higher than 150mmh^{-1} , the rain gauge starts laminating the output flow, with a twofold consequence: the correction formulas shown in Section 4 do not yield correct results and the collector lamination delays the output flow. Both effects lead to underestimating the actual rainfall intensity. Our simulation shows that an average of about 0.5 event per year presents instantaneous rainfall intensity higher than $150\text{mm} \cdot \text{h}^{-1}$. Figure 7-a reports the results for a randomly chosen simulation, and Figure 7-b shows the whole set of simulations.

5.1. Total rainfall accumulation

The sum of tipping-bucket measurements over a long period (i.e. a year) is sometimes used for estimating the total rainfall. Usually, the water ejected by the tipping buckets is collected to measure the total rainfall amount (e.g. by weighting or reading a graduate cylinder). Due to the mechanical errors mentioned above, the total collected volume often differs from the sum of the rainfall quantum observed by the tipping-bucket device.

Table 3 displays the average annual-rainfall amounts obtained with and without correction. In order to understand these results we must keep in mind that most of the cumulated rainfall intensities are weak. The accumulation of raw measurements gives an excellent estimation of annual accumulations when the gage is calibrated for low intensities ($I_{opt} = 0\text{mm} \cdot \text{h}^{-1}$). The linear correction, resulting in a constant factor higher than the unit (Table 3), produces an overestimation of the total rainfall amount. For low I_r values, the power-law relationship (Table 3) of the dynamical correction (Equation 10) is close to a linear relationship with slope lower than 1 (Figure 5-a,b). The application of the physically-based correction formula reported in Equation 11 leads to the most accurate total.

5.2. Extreme rainfall

In order to examine the case of large return-period rainfall rates, we assume that the rainfall maxima can be modeled with the 2-parameter Gumbel distribution. We estimate, for durations of 5, 10, 15, 20, 30, 45 and 60 minutes, the 100 year return period rainfall based on a sample of extracted annual maxima. The average results for 30 simulations are shown in Table 4 and in Figure 8. The three correction methods have substantially equivalent skills in correcting the measured maximum. The best result is provided by the rigorous correction, and the dynamic correction provides significant improvements respect to the linear one. The use of correction methods is thus recommended in the case of $I_{opt} = 0\text{mm} \cdot \text{h}^{-1}$, because the raw measurements are systematically underestimated. The choice of a positive rainfall optimum improves the raw measurement, without affecting the correction method skills.

Analyzing the error in the estimation of extremes we can state that, as expected, the strong singularities are smoothed by the temporal aggregation and the error decreases with the accumulation period.

6. Conclusive remarks

The paper dealt with the uncertainty of rainfall measurement using a tipping-bucket rain gauge. We used a statistical disaggregation of hourly data measurements in order to reproduce the fine-scale variability of the rainfall process. This rainfall disaggregation down to the scale of 15 seconds allowed us to show the influence of the small-scale rainfall variability on rain gauge measurement accuracy. The functioning of a tipping-bucket gauge, including the flow lamination by the collector, has been numerically reproduced.

For the station that we have analyzed, located in Montpellier (France), we found that in average once every two years the collector enters in charge due to instantaneous rainfall intensities over $150 \text{ mm} \cdot \text{h}^{-1}$. This has consequences on the rainfall measurement at very high temporal resolutions (less than 5 minutes) and leads to underestimated rainfall intensities. This lead to say that, designing a new rain gauge, the first and most important step is to determine the maximum rainfall intensity that can occur at the desired temporal resolution in order to obtain a correct dimensioning of the collector area and of the bottom hole. The main problem with historical rain gauge series is that the devices were designed to measure rainfall at a scale and later the attention moved to finer resolutions, for which these problems may occur.

We verified that the small-scale variability does not significantly affect this underestimation of extreme rainfall for accumulation periods higher than 5 min. Different from what Molini et al. (2005b) found, our analysis did not show large deviations in the estimation of extreme quantiles. This result is partly due to the lamination effect of the collector, which, on the other hand, affects the measurement of very high intensities for accumulation periods lower than 5 min. Another reason is the location of the analyzed rain gauge station: rain gauges located in regions with different climatic features may lead to higher instantaneous rainfall intensity and, consequently, to significant deviations.

It is well known that the rain gauge calibration may improve the raw measurement of the rainfall intensity. For this reason, most of the European rain gages are set to an optimal rainfall value that is higher than $0 \text{ mm} \cdot \text{h}^{-1}$. We demonstrated that the choice of an optimum intensity value at $50 \text{ mm} \cdot \text{h}^{-1}$ instead of $0 \text{ mm} \cdot \text{h}^{-1}$ leads to the decrease of the maximum raw-measurement error from 6% to about 4%.

To furtherly reduce the measurement error, the application of correction formulas is a necessary step. Having series at very-high resolution (for instance 1 or 5 min), the correction can be directly applied on data. In cases where hourly data is the only available series, the correction can be applied on a synthetic series generated at finer scale by applying cascade disaggregation of the low-resolution series.

Three methods have been used for the correction of rainfall observations: the standard empirical corrections (a linear and a so-called dynamic correction) and a physically-based correction. The empirical corrections work well for heavy rainfall observations, but they provide underestimation of the annual rainfall amount. The inversion of the physically-based formula proposed by Marsalek (1981) is coherent with the principle of the tipping-bucket mechanism and it actually demonstrated to be the best correction method for either the maximum intensities and total accumulations. The number of parameters to be estimated does not increase, but the equation is slightly more complex. The advantage is that, when the parameters α , β_1 and β_2 of the linear and dynamic transformations (Equation 10) have no physical meaning, h_n and i_{opt} of Equation 11 are intrinsic properties of the device (namely, the tipping-bucket volume and the rainfall intensity for which the measure is unbiased).

Extracting samples of maxima from the true rainfall series, raw measurement and corrected series, we compared the 100 year return period rainfall at various accumulation periods. The comparison revealed that the raw measurements can give underestimation in the order of 5% in the rainfall for $T_R = 100\text{y}$ for $D=5$ min, decreasing with the accumulation duration. Any of the correction methods provide a sensible improvement of the extreme-rainfall estimation, but the rigorous correction (Equation 11) is the only that, in addition, yields a reliable total rainfall amounts.

From a methodology view point our conclusion is that the analysis of the small-scale variability of rainfall is a necessary step for having reliable high-resolution ground rainfall measurements. This step concerns both historical records and new rain gauges design.

Concerning historical records, the raw hourly data could be corrected by applying the correction in Equation 8 to a 5-min disaggregated series obtained from the original hourly data. The direct correction at the hourly scale is not recommended, leading to errors higher than the measure bias. To perform a reliable correction, the knowledge of the rain gauge features (e.g. optimum rainfall value) is mandatory. Concerning the design of new rain gauges, one must decide if a correction of data will be undertaken or not. If no correction is planned, the best optimum depends on the aim of the measurement. If other measures of rainfall depth are available, a relatively high optimum (in the order of 50mm h^{-1}) allows to obtain low biases in the raw measurements of extreme rainfall. This however generates a systematic overestimation of low rainfall intensities. If corrections are planned, the choice of the optimum rainfall value does not affect the correction. The best is to fix the optimum as the (supposed) value of average positive rainfall, so that the sum of the underestimations and over-estimations is balanced. The best correction method is the application of Equation 11 directly obtained by inversion of the Equation 1 at the time scale of 5 min. The two other methods yield acceptable errors in the estimation of rainfall extremes but give spurious results in term of accumulated rainfall.

Extending the results to the rainfall network of the OHMCV (Hydro-meteorological Cévennes-Vivarais observatory), it seems that further analyses should be conducted especially in zones prone to deep convective events, in which the maximum rainfall depth may exceed 100 mm in 1 hour, differently from what happens in Montpellier, where the maximum hourly intensity recorded in over 50 years of data is in the order of 70 mm.

A short-term perspective of the work is to perform a sensitivity analysis of the different disaggregation methods (namely the disaggregation method proposed by Molini et al. (2005b)), to verify the reliability of the results shown in this paper.

References

- Calder, I., Kidd, C., 1978. A note on the dynamic calibration of tipping-bucket rain gauges. *Journal of Hydrology* 39, 383–386.
- Ciach, G., 2003. Local random errors in tipping-bucket rain gauge measurements. *Journal of Atmospheric and Oceanic Technology* 20, 752–759.
- Habib, E., Krajewski, W., Kruger, A., 2001. Sampling errors of tipping-bucket rain gauge measurements. *Journal of Hydrologic Engineering* , 159–166.
- Humphrey, M., Istok, J., Lee, J., Hevesi, J., A.L., F., 1996. A new method for automated dynamic calibration of tipping-bucket rain gauges. *Journal of Atmospheric and Oceanic Technology* 14, 1513–1519.
- Marsalek, J., 1981. Calibration of the tipping bucket rain gauge. *Journal of Hydrology* , 343–354.
- Menabde, M., 1998. Bounded lognormal cascades as quasi-multiaffine random processes. *Nonlinear Processes in Geophysics* 5, 63–68.
- Molini, A., Lanza, L., La Barbera, P., 2005a. The impact of tipping-bucket rain gauge measurement errors on design rainfall for urban-scale applications. *Hydrological processes* , 1073–1088.
- Molini, A., Lanza, L., La Barbera, P., 2005b. Improving the accuracy of tipping-bucket rain records using disaggregation techniques. *Atmospheric Research* , 203–217.
- Niemczynowicz, J., 1986. The dynamic calibration of tipping-bucket rain gauges. *Nordic Hydrology* 17, 203–214.
- Over, T., Gupta, V., 1996. A space-time theory of mesoscale rainfall using random cascades. *Journal of Geophysical Research* 101, 26319–26331.
- Schmitt, F., Vannitsem, S., Barbosa, A., 1998. Modeling of rainfall time series using two-state renewal processes and multifractals. *Journal of Geophysical Research* 103, 23181–23193.
- Sevruk, B., 1972. Evaporation losses from storage gauges: Distribution of precipitation in mountainous areas, in: *Proceedings of Geilo Symposium 1972*, pp. 96–102.

- Sevruk, B., 1987. Point precipitation measurements: why are they not corrected?, in: Water for the Future: Hydrology in Perspective - Proceedings of the Rome Symposium, April 1987., pp. 477–486.
- Sevruk, B., 1996. Adjustment of tipping-bucket precipitation gauge measurements. Atmospheric Research 42, 237–246.
- Veneziano, D., Furcolo, P., 2002. Multifractality of rainfall and scaling of intensity-duration-frequency curves. Water Resources Research 38, 1306.
- Wood, S., Jones, D., Moore, R., 2000. Static and dynamic calibration of radar data for hydrologic use. Hydrology and Earth System Sciences 4, 545–554.

Table 1: Monthly parameters C_β and C_1 of the beta-lognormal model for the series of Montpellier Bel-Air. The parameter α , Levy’s stability index, is fixed to 2 in order to have log-normal distribution of the weights.

Parameter	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
C_β	0.43	0.46	0.41	0.50	0.51	0.59	0.67	0.62	0.56	0.47	0.46	0.43
C_1	0.035	0.036	0.040	0.050	0.059	0.062	0.093	0.059	0.079	0.059	0.050	0.041

Table 2: Rain gauge correction formula for the linear correction (Equation 9) and dynamic correction (Equation 10) for optimum values of 0 and 50 $mm \cdot h^{-1}$.

Optimum rainfall value	Linear correction	Dynamic correction
$i_{opt} = 0 \text{ mm } h^{-1}$	$I_a = 1.032 \cdot I_r$	$i_a = 0.924 \cdot I_r^{1.024}$
$i_{opt} = 50 \text{ mm } h^{-1}$	$I_a = 1.018 \cdot I_r$	$i_a = 0.909 \cdot I_r^{1.024}$

Table 3: Comparison of actual and measured (raw and corrected) rainfall maximum for $I_{opt} = 0 \text{ mm } \cdot h^{-1}$ and $I_{opt} = 50 \text{ mm } \cdot h^{-1}$. Meas: raw measurements; Lin: linear correction; Dyn: dynamic correction; PhyB: Physically-based correction (Equation 11). The result is the average of 30 stochastic simulations of 52-year series.

Actual raindepth ($mm \text{ y}^{-1}$)	$I_{opt} = 0 \text{ mm } h^{-1}$			
	Meas.	Lin.	Dyn.	PhyB
715	713 (-0.3%)	734 (+2.7%)	684 (-4.3%)	713 (-0.3%)
Actual raindepth ($mm \text{ y}^{-1}$)	$I_{opt} = 50 \text{ mm } h^{-1}$			
	Meas.	Lin.	Dyn.	PhyB
715	722 (+1%)	735 (+2.8%)	683 (-4.5%)	714 (-0.1%)



Figure 1: Detail of a tipping-bucket mechanism for rainfall measurement. As it can be seen in a standard recording rain gauge, a small device deployed above the bucket is designed to drive the water flow beyond the barycenter of the system, helping the filled bucket to rock.

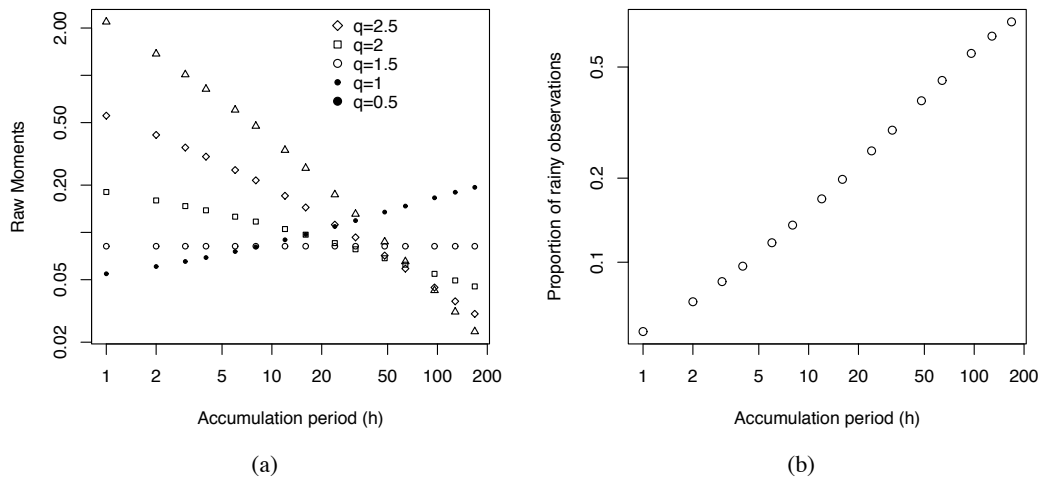


Figure 2: Hourly rainfall series of Montpellier (1920-1972). a: Statistical moments for scales in the range 1 hour - 1 week. b: Rainfall intermittency in the range 1 hour - 1 week.

Table 4: Comparison of true versus observed (and corrected) maximum rainfall intensity ($mm \cdot h^{-1}$) for durations ranging from 5 to 60 min. The two calibration scenarios ($I_{opt} = 0 mm \cdot h^{-1}$ and $I_{opt} = 50 mm \cdot h^{-1}$) are successively reported. The result is the average of 30 stochastic simulations of 52 years series.

Accum. period (min)	Rain intensity (mmh^{-1})	$I_{opt} = 0 mm h^{-1}$			
		Measured	Linear	Dynamic	Physically-based
5 min	156.4	151.0	155.8	155.6	155.4
10 min	127.2	123.4	127.3	126.7	126.4
15 min	108.9	106.1	109.4	108.6	108.4
20 min	103.9	101.3	104.5	103.6	103.5
30 min	92.4	90.2	93.0	92.0	92.0
45 min	80.6	78.7	81.2	80.2	80.2
60 min	70.0	68.5	70.7	69.6	69.6
Accum. period (min)	Rain intensity (mmh^{-1})	$I_{opt} = 50 mm h^{-1}$			
		Measured	Linear	Dynamic	Physically-based
5 min	156.5	153.1	155.8	155.6	155.4
10 min	127.2	125.1	127.4	126.7	126.5
15 min	108.9	107.5	109.4	108.5	107.5
20 min	103.9	102.7	104.5	103.5	103.5
30 min	92.4	91.4	93.1	92.0	92.0
45 min	80.6	79.8	81.3	80.2	80.2
60 min	70.0	69.5	70.7	69.7	69.7

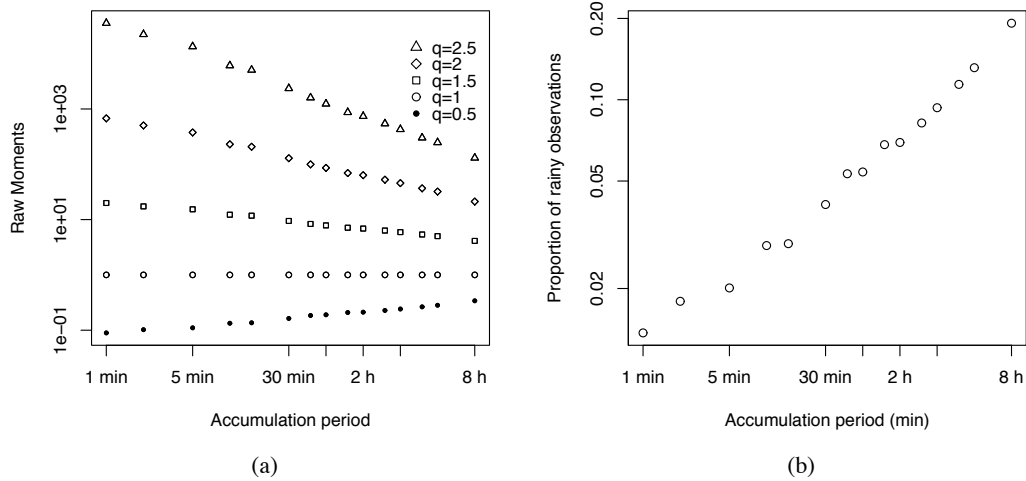


Figure 3: 15-sec disaggregation of the rainfall series of Montpellier for the simulation #30 (period 1920-1972). a: Statistical moments for scales in the range 15 sec - 1 day. b: Rainfall intermittency in the range 15 sec - 1 day.

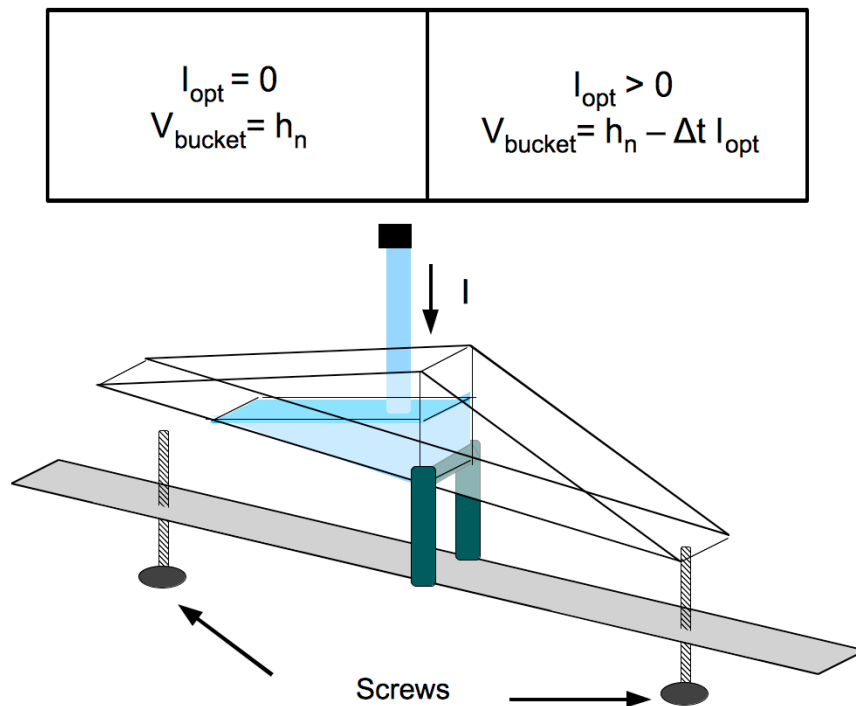


Figure 4: Scheme of the tipping bucket device with indications on the bucket volume to assign in order to get unbiased measurements of the optimum rainfall value.

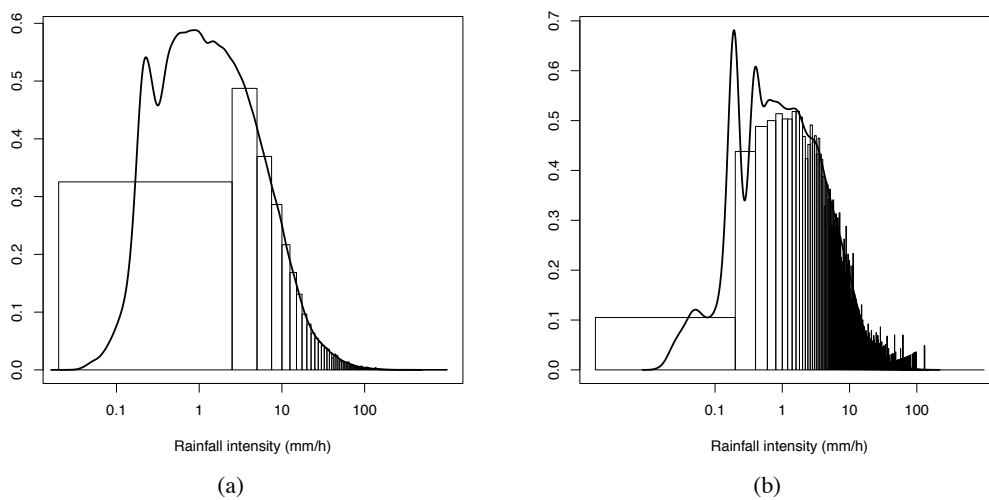


Figure 5: Discretization of the rain gauge measurement due to the resolution of the rain gauge tipping-bucket device: comparison between rainfall frequency distribution (density line) and discretized rainfall as seen by the rain gauge (histogram). a: 5 min, b: 60 min.

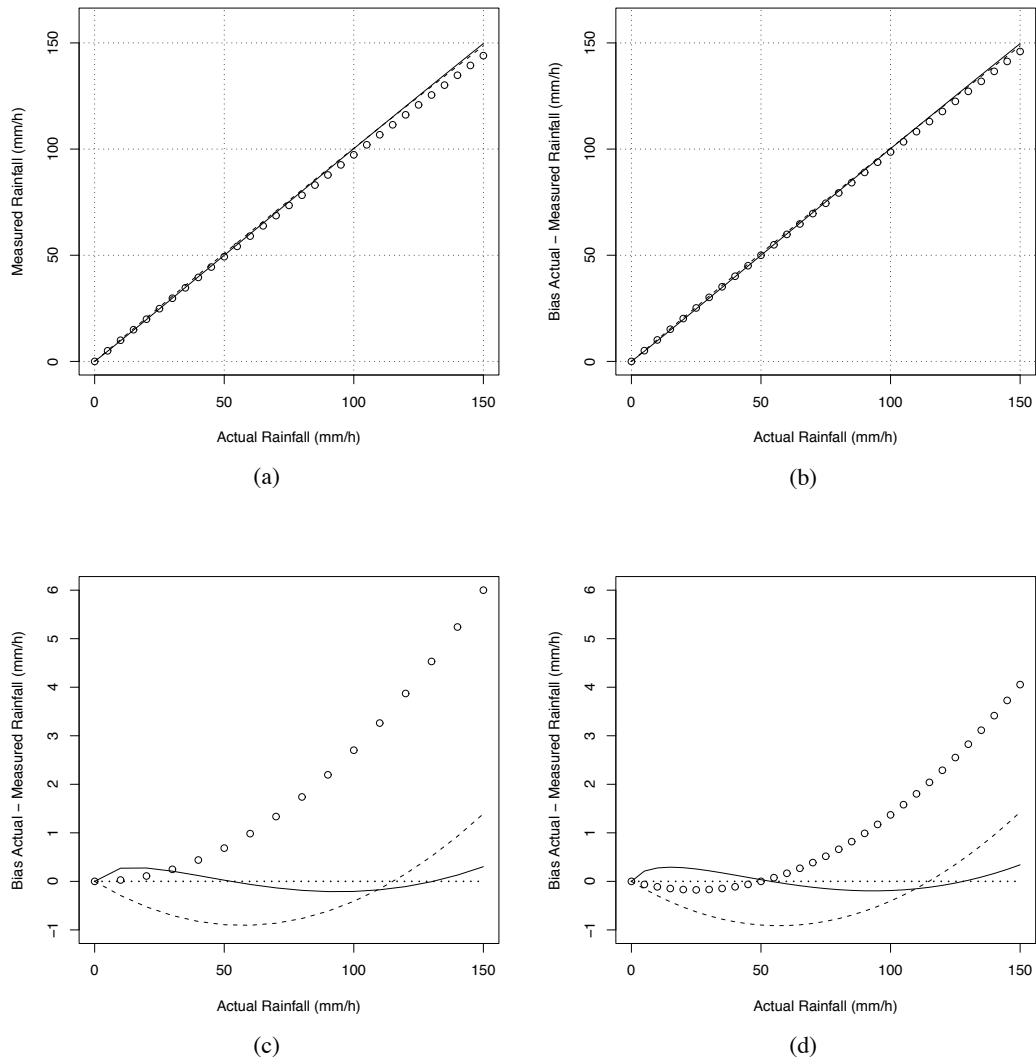


Figure 6: Illustration of tipping bucket gage calibration and corrections. For two optimal rainfall values $I_{opt} = 0 \text{ mm} \cdot \text{h}^{-1}$ (a and c) and $I_{opt} = 50 \text{ mm} \cdot \text{h}^{-1}$ (b and d) the graphs report raw measured rainfall intensities (dots) and correction formulas (dashed line for linear and solid line for dynamic correction formulas and solid line for dynamic correction) for a set of control rainfall rates (0 to $150 \text{ mm} \cdot \text{h}^{-1}$ with steps of $10 \text{ mm} \cdot \text{h}^{-1}$). The top and bottom graphs give, respectively, the corresponding relationships and the relative differences with respect to the control rainfall.

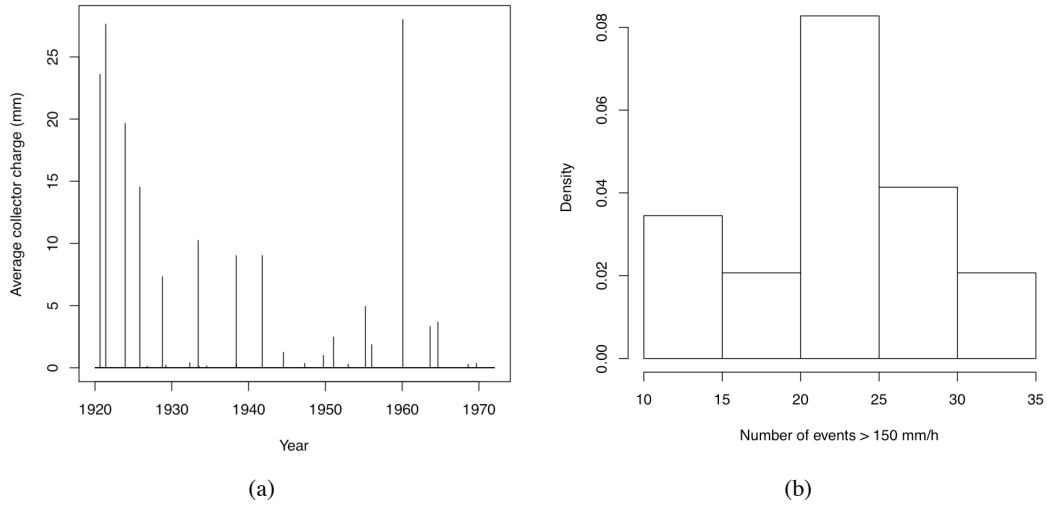


Figure 7: a: Water height in the collector when the rain rate exceeds its evacuation capacity. In this simulation, 25 events over the 52 years featured instantaneous intensities higher than 150 mm h^{-1} causing the collector to enter in charge. b: histogram of the number of events with $I > 150 \text{ mm h}^{-1}$ for the 30 simulations covering 52 years of data.

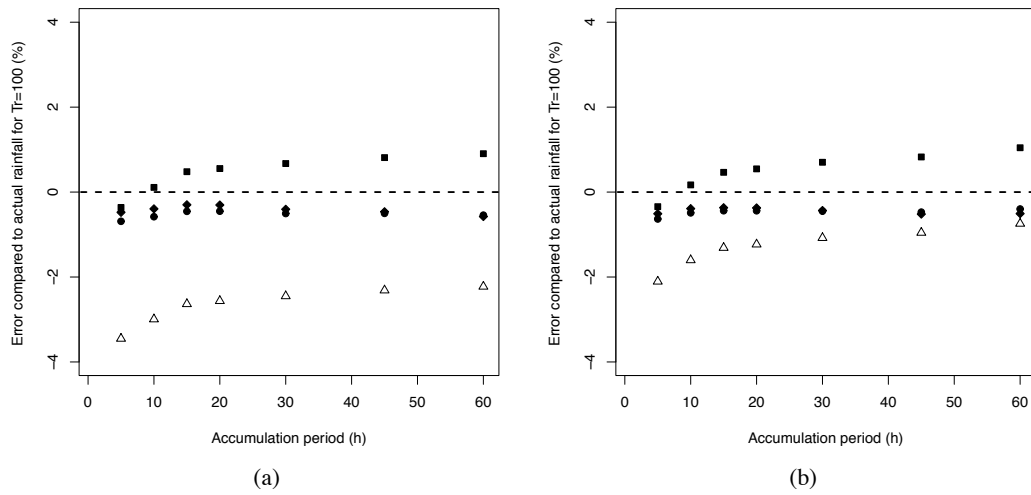


Figure 8: Percent error showing the effect of rainfall correction on the estimation of the 100 year return period rainfall for two optimal rainfall values $I_{opt} = 0 \text{ mm h}^{-1}$ (a) and $I_{opt} = 50 \text{ mm h}^{-1}$ (b). The dashed line represents the actual rainfall used as reference, circles identify the results when applying the physically based correction, the squares when applying linear correction and the diamonds when applying dynamic correction. The triangles represent the raw (uncorrected) measurements.

2.4 Ground measurement network for spatial estimations: limits

The rain gauge network is a collection of point measures that are often utilized for the estimation of the spatial rainfall. The rain gauge network is more or less able to catch the actual behavior of the field, and this capacity basically depend on the ratio between the density of the rain gauge network and correlation structure of the rainfall fields. An infinitely dense rain gauge network is not realizable in practice, and therefore one must rely on a sparse ground measurement network or, where possible, on the radar measured rainfall field, derived from the transformation of the rainfall reflectivity into rainfall intensity. The issues relative to the use of radar fields are described in Section 1.5.2.

Skoien and Blöschl (2006) define the sampling scale triplet of spacing, extent and support to define the spatial dimensions of a monitoring network, examining the random error introduced for the estimation of mean, spatial variance as a function of the various configurations of the sampling triplet. The results indicated that the estimation error and the variable can be of the same order of magnitude.

Intuitively, since the interpolation is a combination of observed values, the rain gauge network must be able to catch the main features of the rainfall field, namely its average, variance, maximum. In the following we use an academic example to verify the capability of different rainfall networks to detect the maximum of rainfall fields.

Several numerical simulation have been performed to estimate the uncertainty in the measure of spatial rainfall by means of a rain gauge network. A series of log-normal fields have been generated¹. The log-normality of spatial rainfall is supported by numerous studies (for example, *Kedem and Chiu* (1987)) and empirical evidences.

Log-normal fields with two different correlation structures (for further details on the correlation structure of spatial variables, go to Section 4) have been generated: in the simulation I, 100 independent log-normal fields characterized by a spherical correlation structure with decorrelation distance of 20 km (approximately the hourly-rainfall correlation distance); in the simulation II, the correlation distance is 100 km (approximately the daily-rainfall correlation distance).

Five different rain gauge networks have been randomly generated, characterized by an increasing density of 81, 100, 144, 196, 324 rain gauges per 10000 km^2 . The simulated rain gauge measurements have been interpolated by appropriate geostatistical techniques (Section 4). The structure of the fields have been reconstructed and the maxima corresponding to different aggregations surfaces (from 1 to 300 km^2) have been computed. Figure 2.5 show an idealized 1-D example of how the field maxima can be underestimated in case of data undersampling.

An example of the rainfield reconstruction for the simulation I is reported in Figure 2.6. We represented the synthetically generated field (Figure 2.6-a), and the reconstruction with networks of increasing densities (Figure 2.6-b:f). Figure 2.7 reports the analogous rainfall field reconstruction for daily rainfall, characterized by approximate decorrelation distance of 100 km (simulation II).

¹using the R package “RandomFields”

We summarize the results of the 100 simulations by representing the average value of the indicator $E[\hat{Z}_{max}/Z_{max}]$ (Figure 2.8). The plot shows that the hourly spatial rainfall can not be efficiently measured unless having an extremely dense rain gauge network (only the network composed by 324 rain gauges provides satisfactory results). For areas lower than 200 km^2 the inference of spatial rainfall is strongly biased. Figure 2.8-b shows that the same measurement networks give more reliable results in measuring spatial rainfall when the rainfall correlation structure is extended up to 100 km.

As we have seen, the rain gauge network does not suffice for the estimation of spatial rainfall excepted for very large surfaces. A slight improvement can be realized by choosing an appropriate rain gauge distribution.

To determine if a regular rain gauge network grid may improve the rain gauge estimation, we have realized a second set of simulations, distributing the rain gauges according to a “random stratified” pattern (Figure 2.9), a pattern resulting from the general tendency to add gauges in order to fill the non gauged sub-regions. It consists in dividing the area in a number of squares equal to the desired number of gauges, and then to randomly locate one rain gauge in one square.

The comparison of Figure 2.8 and 2.10 show that the stratified random sampling has a positive influence in the determination of the spatial rainfall amount, especially when the sampling density is very low. The effect is practically neglectable in the densest networks. Configurations similar to that of OHM-CV survey (density: $1/100 \text{ km}^{-2}$) lead to considerable (higher than 10% in average) underestimations of the maximum spatial rainfall amount for surfaces lower than 200 km^2 for hourly rainfall and 100 km^2 for daily rainfall. We have to remember that, since we have considered the rain gauge as representative of the surface of 1 km^2 , an additional underestimation due to the small-scale rainfall variability could affect the rainfall measurement. According to *Journal and Huijbregts (1978)*, however, this error is practically neglectable.

2.5 Conclusion

The tipping-bucket rain gauge has numerous advantages for the measure of ground rainfall, among them the most important is that each measure is automatic and perfectly repeatable. Nevertheless, the measure of rainfall is submitted to several errors. The error related to the environment and

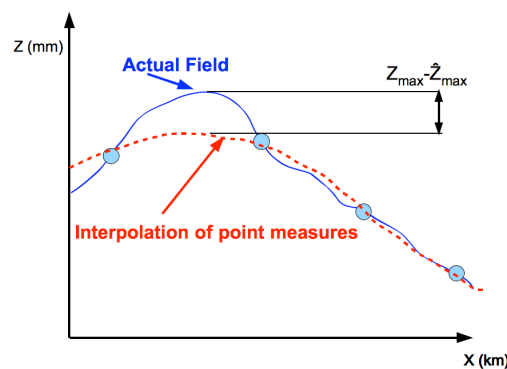


Figure 2.5: Interpolation error in the estimation of the field maximum Z_{max} due to the spatial undersampling.

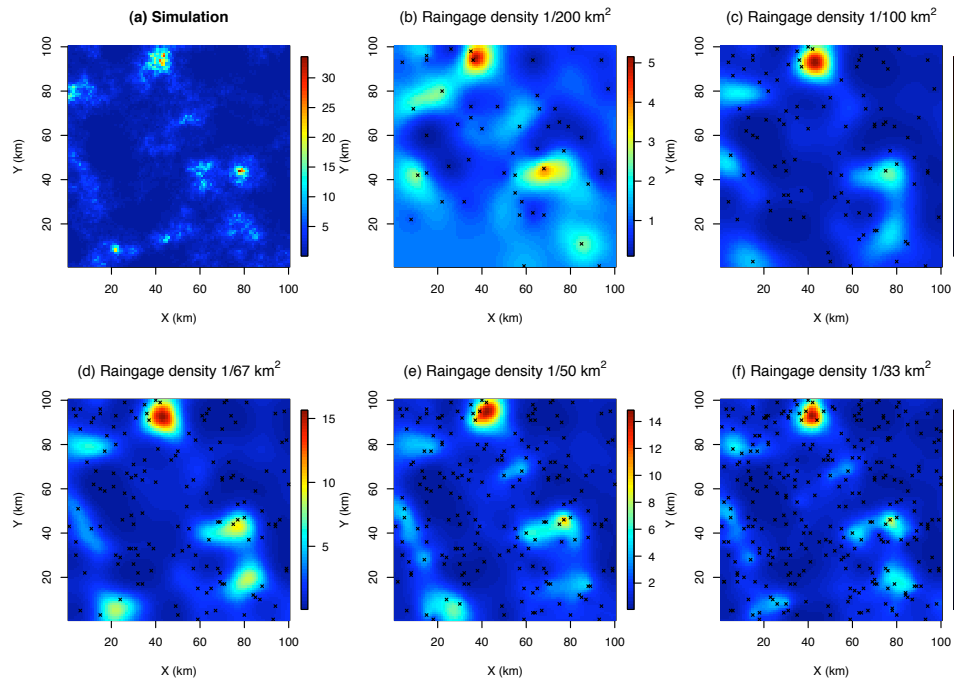


Figure 2.6: Simulated rainfall field and reconstruction by kriging interpolation for different rain gauge densities. The rain gauge repartition is randomly chosen. Decorrelation distance = 20 km (close to that of hourly rainfall)

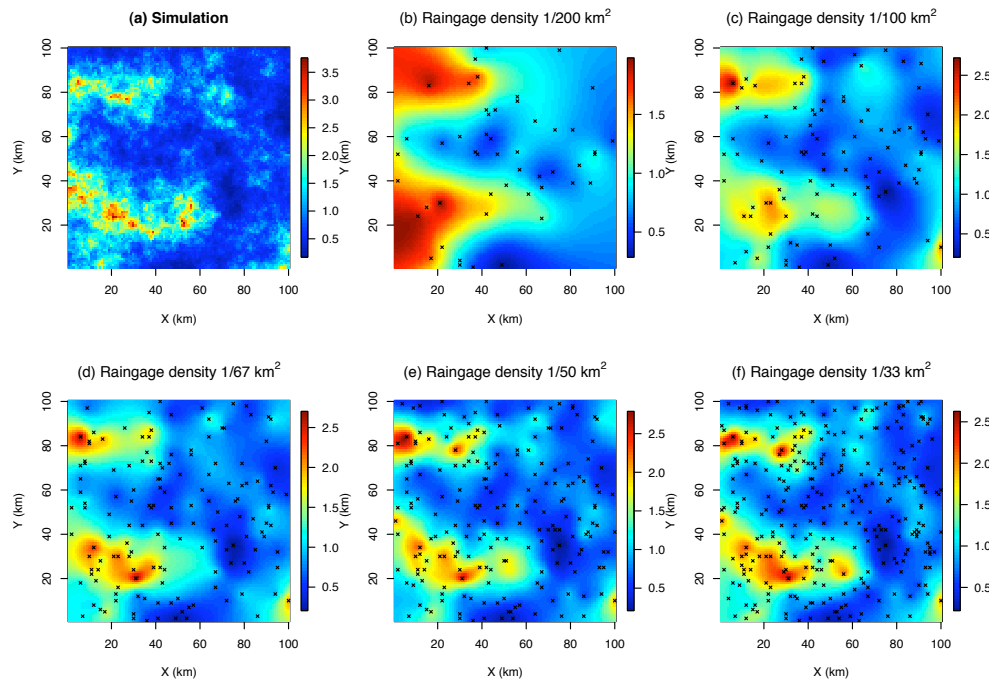


Figure 2.7: Simulated rainfall field and reconstruction by kriging interpolation for different rain gauge densities. The rain gauge repartition is randomly chosen. Decorrelation distance = 100 km (close to that of daily rainfall)

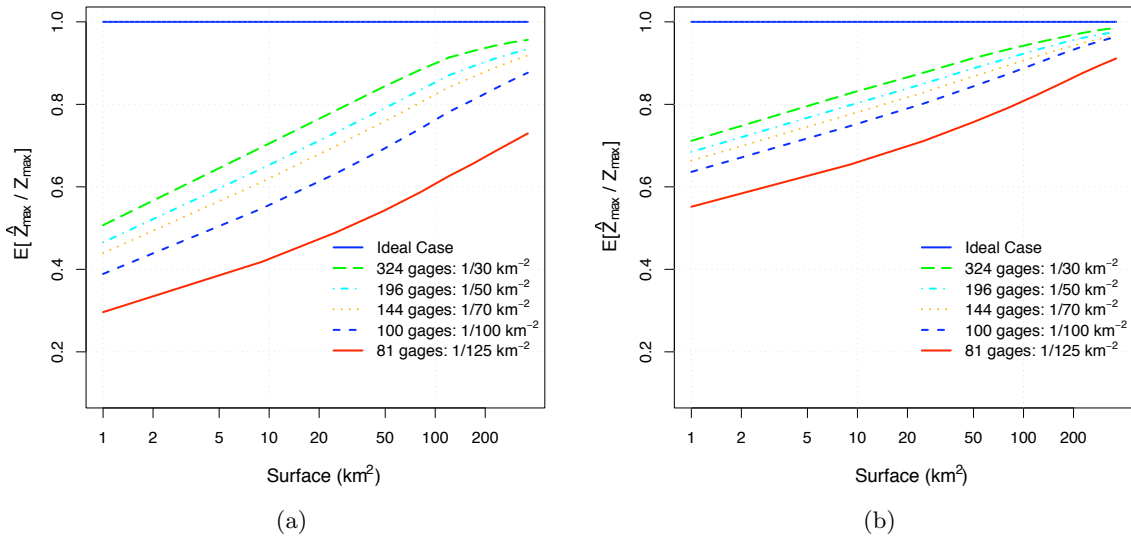


Figure 2.8: Effect of sparseness of the rain gauge network on the spatial rainfall measurements: for 100 simulations, average of the ratio between simulated spatial-rainfall maxima and observed spatial-rainfall maxima. a: Log-normal fields with spherical variogram (range 20 km). b: Log-normal fields with spherical variogram (range 100 km).

local effects are not discussed here. Supposing that the rain gauge is correctly located and properly maintained, mechanical errors remain. Due to the intrinsic structure of the device and to resolution issues, the measured rainfall can be under/overestimated.

The analysis of series at a scale close to the resolution of the measure leads to an underestimation of the magnitude of the extremes. At certain scales, the analysis of extreme observations is affected by systematic errors due to the intrinsic properties of the measuring device, such as the collector storage effect on the instantaneous rainfall intensity measurement in case of instantaneous rainfall amounts higher than $100\text{-}200 \text{ mm h}^{-1}$, as observed for the Cévennes rain gauge network.

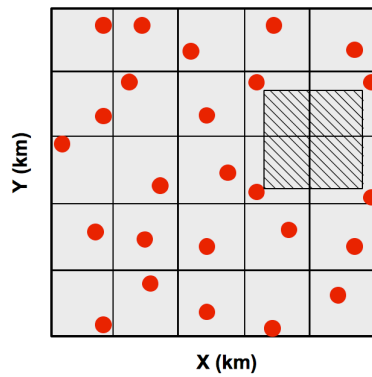


Figure 2.9: Random Stratified sampling pattern. Each pluviometer is located into a square. The hatched area identifies the maximum ungaged area.

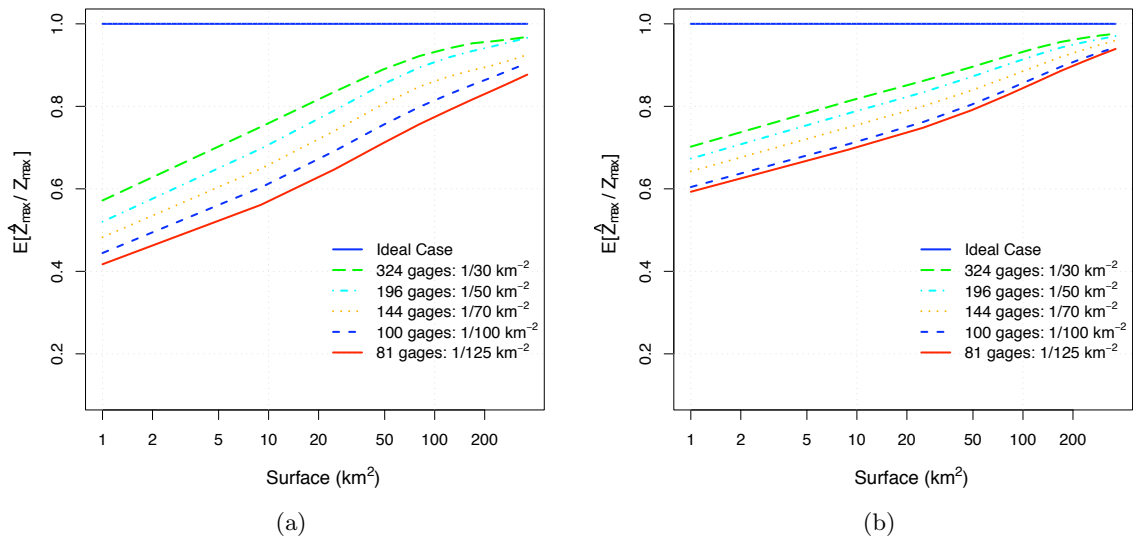


Figure 2.10: Effect of stratified random sampling on the spatial rainfall measurements: for 100 simulations, average of the ratio between simulated spatial-rainfall maxima and observed spatial-rainfall maxima. a: Log-normal fields with spherical variogram (range 20 km). b: Log-normal fields with spherical variogram (range 100 km).

The extreme measurements through tipping-bucket device are affected to systematic errors in the order of 5 % at the 5-min resolution, lower at the hourly resolution. Due to the presence of a single optimum rainfall value for which the rainfall is correctly measured, a correction is needed to avoid incongruences between the annual or monthly rainfall depth measured by tipping-bucket and by graduate cylinder. Only the data correction through appropriate theoretical formulas can provide acceptable results in the total rainfall depth measurement. In case the correction algorithm is not implemented and the tipping-bucket is expected to work as a totalizing rain gauge as well, to minimize the total error the optimum should be fixed at a value corresponding to the average of positive rainfall intensity values.

Besides the point-rainfall errors, the rain gauge network density is often too coarse to correctly detect the rainfall structure. The consequence is that even the most refined interpolation technique to obtain an estimation of the spatial rainfall amount will result in strong underestimations of the rainfall maxima.

Part II

Theoretical background

Résumé

Les chapitres qui suivent sont dédiés à la description du cadre théorique de la thèse. L'objectif du travail de thèse étant la caractérisation de la pluie extrême et de ses propriétés d'invariance d'échelle dans l'espace et dans le temps, le cadre théorique est composé des sujets suivants:

- **Statistique des extrêmes:** comment estimer la fréquence d'occurrence ou l'intensité de variables qui sont rarement observées? Combien est l'estimation robuste?
- **Géostatistique:** comment peut-on définir la structure spatiale (2D) d'un champ de pluie?
- **Invariance d'échelle:** comment peut-on estimer la distribution d'une variable dans des échelles différentes de l'échelle d'observation?
- **Extrêmes de pluie spatiale :** comment coupler les statistique de la pluie spatiale avec le comportement des extrêmes?

Introduction

The following chapters are devoted to the theoretical background of the thesis. The aim of the thesis is the characterization of extreme rainfall and its scaling in space and time. In this thesis, the theoretical framework lies on:

- **Statistics of extremes:** how to estimate the intensity or the frequency of extreme observations? How robust are these estimations?
- **Geostatistics:** how to define the 2D spatial structure of rainfall fields?
- **Scaling:** how to estimate probability density functions at scales different from the observed ones?
- **Spatial Rainfall Extremes:** how to generalize the extreme value theory for the study of spatial rainfall?

Statistics of extreme point-rainfall intensities

Résumé

Dans cette section, nous décrivons la **théorie des extrêmes** et les méthodes pour modéliser les précipitations extrêmes ponctuelles. L'objectif de l'analyse des valeurs extrêmes est de déduire le comportement des extrêmes à partir d'un échantillon limité de données. La conception d'ouvrages, par exemple, exige que les structures (ponts, barrages) puissent résister aux phénomènes extrêmes (inondations) se produisant tous les 200, 500, 1000 ans. Malheureusement, les données disponibles (pluie ou débit de la série) ne sont pas suffisamment longues pour estimer ces quantités directement par extraction du relatif quantile de l'échantillon, la plus longue série ayant 50-100 ans de données.

Grâce à l'analyse des valeurs extrêmes, en supposant que tous les extrêmes soient indépendantes et tirées de la même distribution, il est possible d'extrapoler le modèle des extrêmes pour prédire la valeur maximale que la variable peut vraisemblablement atteindre une fois tous les 200, 500 ou 1000 ans.

L'analyse des valeurs extrêmes est soumise à une série de limitations: i) si la longueur des données est limitée, l'extrapolation produit très grandes incertitudes (quantifiées au travers du concept d'intervalle de confiance), ii) l'indépendance des extrêmes doit être assurée; iii) toutes les observations doivent être tirées de la même distribution; iv) les séries doivent être stationnaires. Enfin, il faut considérer que, malgré les lois de valeurs extrêmes soient valables en théorie, les vraies observations sont soumises à problèmes d'échantillonnage et peuvent être bien loin du comportement idéal.

Après une brève introduction sur les concepts de **période de retour** et quelques définitions, les trois méthodes pour l'analyse des valeurs extrêmes généralement adoptées seront décrites dans cette section:

- l'analyse des **maxima par blocs**;
- l'analyse des **excès par rapport à un seuil**;
- la méthode du **processus ponctuel de Poisson**.

Enfin, des considérations à propos de la **stationnarité** des séries temporelles seront tirées.

3.1 Introduction

In this section we will describe the **theory of extremes** and the methods to model point-rainfall extremes. The objective of the extreme value analysis is to infer the behavior of extremes based on a limited sample of data. The engineering design, for example, requires the structure (a bridge or a dam) to resist to events (floods) occurring every 200, 500, 1000 years. Unfortunately, the available data (rainfall or flow series) are not sufficiently long to estimate the extremes behavior directly on the sample since the longest series features 50-100 years of data.

The extreme analysis consists in the appropriate selection of a sample of independent and identically-distributed (i.i.d.) sample of extremes from the original distribution, with the aim to extrapolate the behavior of a series of extremes for predicting the value that the variable can presumably attain once every 200, 500 or 1000 years.

The extreme value theory is submitted to a series of limitations: if the sample size is limited, the extrapolation will produce very large uncertainties (quantified through the confidence interval concept); the independence of the extremes has to be ensured; all the observations must be drawn from the same distribution. In addition, one has to consider that, despite the extreme value laws are valid in theory, real-life data may show large deviations.

After a brief introduction about the concepts of **return period** and some definitions, the three methods for the extreme value analysis will be described in this section:

- the **block maxima** analysis;
- the **Peaks Over Threshold** analysis;
- the **Point Process** Analysis.

Some considerations about the stationarity of time series will follow.

3.2 Frequency and return period: definitions

Extreme events are, by definition, rare. We can quantify their likelihood of occurrence in terms of frequency, but in hydrology it is usually preferable to work in terms of return period, defined as “The average number of years that last between the reference event and another event of the same magnitude or higher” .

We quantify the likelihood of occurrence $F(x)$ or equivalently the return period $T_R = 1/(1 - F(x))$. Hydrologist are familiar with the latter concept: the classic extreme analysis, based on the extraction of annual maxima of series, expresses the occurrence of rainfall events in terms of return period.

Table 3.1 shows typical values of return periods and their correspondent frequency level.

Table 3.1: Equivalence between cumulative distribution function $F(x)$ of annual maxima and return period T_R for some significant values of the return period T_R .

$F(x)$	0.9	0.95	0.98	0.99	0.995	0.998	0.999
T_R (yrs)	10	20	50	100	200	500	1000

3.3 Extreme Value Theory

The Central Limit Theorem (CLT) defines the conditions under which the mean of a sufficiently large number of independent random variables will be approximately normally distributed.

Let X_1, X_2, \dots, X_N be a set of N independent random variables and each X_i have an arbitrary probability distribution $P(x_1, \dots, x_N)$ with mean μ_i and a finite variance σ_i^2 .

The variable

$$X = \frac{1}{N} \sum_{i=1}^N x_i \quad (3.1)$$

is normally distributed with $\mu_X = \mu_x$ and $\sigma_X = \sigma_x/\sqrt{n}$.

If CLT gives the asymptotic law of the mean, the extreme value theorem (*Gnedenko*, 1943) gives similar results for the maximum.

Let X_1, X_2, \dots, X_N be a sequence of N independent and identically-distributed random variables with common distribution F , let $M_n = \max\{X_1, \dots, X_n\}$. We can define a sequence of normalization parameters (a_n, b_n) such that each $a_n > 0$ and

$$\lim_{n \rightarrow \infty} P\left(\frac{M_n - b_n}{a_n} \leq y\right) = F^n(a_n y + b_n) \rightarrow G(y) \quad (3.2)$$

If G is a non degenerate distribution function, it belongs to the Generalized Extreme Value distribution class (GEV), defined as:

$$\begin{cases} G(y) = \exp(-(1 + \frac{\xi}{\sigma}(y - \mu))_+^{-1/\xi}) & \text{for } \xi \neq 0 \\ G(y) = \exp(-\exp(-\frac{y-\mu}{\sigma})) & \text{for } \xi = 0 \end{cases} \quad (3.3)$$

where ξ is the extreme value index (also known as the shape parameter) and $h_+ = \max(h, 0)$.

The theorem states that the maximum of a sample of i.i.d. random variables after proper renormalization converges in distribution to one of 3 possible attraction domains.

The ξ value identifies the attraction domain for extremes:

- $\xi < 0$: Weibull attraction domain;
- $\xi = 0$: Gumbel attraction domain;
- $\xi > 0$: Fréchet attraction domain;

In Equation 3.2 we dealt with sample maxima. An expression similar to Equation 3.2 can be derived for the values exceeding a threshold. *Pickands* (1975) showed that if X is a random variable for which Equation 3.2 holds, then:

$$P[X \leq y | X > u_p] \rightarrow H(y) \quad (3.4)$$

if the threshold u_p approaches the endpoint μ_{end} , $H(y)$ is the Generalized Pareto Distribution (GPD), defined as:

$$\begin{cases} H(y) = 1 - (1 + \frac{\xi_p}{\beta_p}(y - u_p))_+^{-1/\xi_p} & \text{for } \xi_p \neq 0 \\ H(y) = 1 - \exp(\frac{y-u_p}{\beta_p}) & \text{for } \xi_p = 0 \end{cases} \quad (3.5)$$

where $H(y)$ is the cumulative distribution function, $h_+ = \max(h, 0)$, u_p is the position parameter (the imposed threshold), β_p is the scale parameter and ξ_p is the shape parameter.

These two asymptotic results motivate the modeling of block-maxima with GEV distribution and the peaks over threshold with GPD distribution.

It is worthy to notice that the three attraction domains for extreme do not depend on the extraction protocol, therefore they are defined based on the ξ value both for GEV and for GPD.

In Figure 3.1 an intuitive scheme shows the relationships between parent distributions and attraction domains for the Gumbel and Fréchet attraction domain.

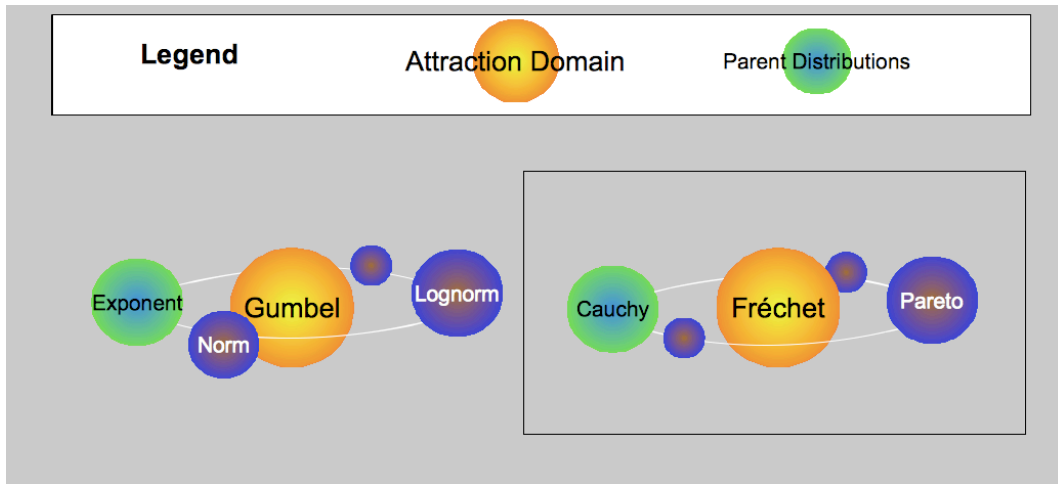


Figure 3.1: The attraction domains of interest in rainfall: Fréchet and Gumbel are the natural attractors of the extremes extracted from the parent distribution. Intuitive scheme inspired to *Mestre* (2008).

3.4 The role of independence in Extreme value analysis

The extreme value analysis requires the sample to be composed by independent and identically distributed (i.i.d.) data. This is the main factor limiting the sample set size. Working with synthetic data, it is easy to produce a large number of i.i.d. observations, but this is not the case when dealing with real data.

Let us analyze the i.i.d. hypothesis in detail:

- Independence. Whatever the extreme value model adopted (see section 3.5 and 3.6), the derivation of the probability density function of extreme random variables assumes the independence of their realizations. A practical method to assess independence between observations is to analyze the auto-correlation function (ACF)¹. The Auto-Correlation function of a random variable $X(t)$ of mean μ and variance σ^2 is defined, for the time lag τ , as:

¹This can be done using the function `{acf}` in the `{base}` R package.

$$ACF(\tau) = \frac{E[(X_t - \mu)(X_{t+\tau} - \mu)]}{\sigma^2} \quad (3.6)$$

where the symbol $E[]$ denotes expected value. The Auto-Correlation Function for the daily series obtained by aggregation of the hourly series at the rain gauge of Montpellier (1920-1972) is reported in Figure 3.2. The independence among observations is ensured for events separated by time lags such that ACF is close to 0 (approximately 100 hours in this case). Since the ACF function (Equation 3.6) is based on the sample mean μ and variance σ^2 , one must take care in examining the auto-correlation function at large time lags (e.g. months, years): the stationarity of data is affected by the seasonal behavior or by climatic trends, and therefore the mean μ as well as the variance σ^2 of the signal variate with the time t .

- The second hypothesis to correctly model extremes is that realizations must be drawn from the same parent distribution. This implies that extremes should be selected among the observations belonging to the same weather type. This condition has several consequences: i) the annual maxima could reasonably be considered as identically-distributed in regions where rainfall events are generated by one prevailing synoptic condition; ii) in regions where two or more weather types could be responsible of extreme events, this condition could be achieved taking into account extremes originated by the same weather type; taking more than one maxima per year could result in a mix of realizations drawn from different weather types that lead to spurious results.

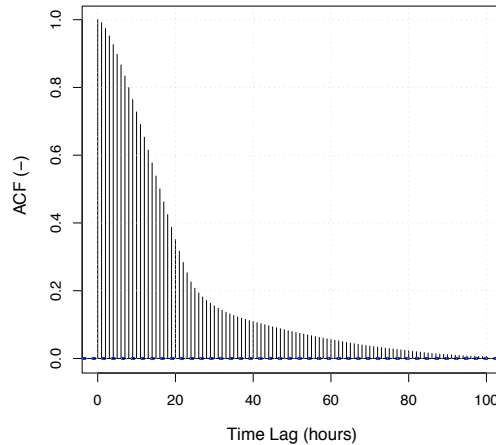


Figure 3.2: Auto-correlation function of the 24-h aggregated series of Montpellier.

3.5 Block-maxima analysis

The block-maxima analysis is the most known approach to model point-rainfall extremes. It consists in extracting the maximum value within a window of fixed size, one year for example (Figure 3.3).

This approach corresponds to the practical application of the extreme value theorem (Equation 3.2), stating that a sample of i.i.d. maxima is distributed as a GEV (Equation 3.3).

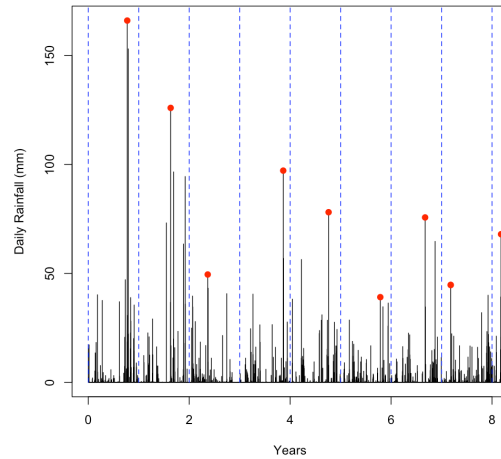


Figure 3.3: Selection of annual maxima (dots). The dashed vertical lines define the blocks. One maximum per block is selected.

The value of ξ determines the distribution limits:

$$\begin{cases} -\infty < x \leq \mu - \frac{\sigma}{\xi} & \text{for } \xi < 0 \\ \mu - \frac{\sigma}{\xi} \leq x < +\infty & \text{for } \xi > 0 \\ -\infty \leq x < +\infty & \text{for } \xi = 0 \end{cases} \quad (3.7)$$

The three configurations of Equation 3.7 are illustrated. For $\xi < 0$ ($\xi = -0.2$ in this case) the random variable is bounded above by a limiting value (Weibull). For $\xi = 0$ the random variable can take any positive value. In the case where $\xi > 0$ (Fréchet density function), the random variable has a lower bound. If μ and σ are equal, a positive ξ (Fréchet) gives higher probabilities of extreme values compared to the Gumbel ($\xi = 0$) case.

The block maxima analysis² is illustrated here on the 52-year long series collected at the station of Montpellier Bel-Air. In order to reduce the sampling resolution effect (Section 2.2), the hourly rainfall rates have been aggregated using a moving-average scan to produce a series of daily rainfall rates.

The method is applied as follows:

- for each year of measures, the absolute 24-hour maximum is retained, obtaining a sample of 52 maxima.
- a first guess of the GEV parameters is done through the application of the method of moments, knowing that

²The package `{ismev}` of R, with the functions “`gev.fit`” and “`gev.diag`”, allows to perform a Maximum Likelihood Estimation of the GEV parameters and to plot the diagnostic graphs.

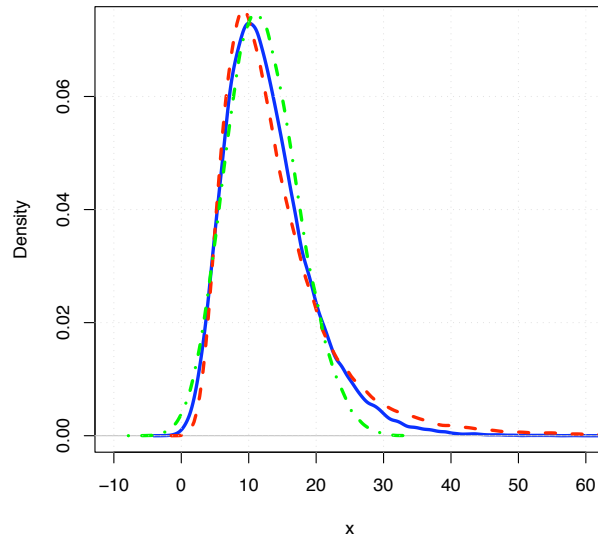


Figure 3.4: Density function of samples belonging to the three attraction domains of GEV-distribution. The three samples have $\mu = 10$ and $\sigma = 5$. The solid blue line identifies a Gumbel sample ($\xi = 0$), the dashed red line a Fréchet sample ($\xi = 0.2$) and the dash-dotted green line a Weibull sample ($\xi = -0.2$).

$$\begin{cases} \mu = E[x] + \frac{\sigma}{\xi} - \frac{\sigma}{\xi} \Gamma(1 - \xi) \\ \sigma^2 = \frac{s[x]^2 \xi^2}{(\Gamma(1-2\xi) - \Gamma(1-\xi)^2)} \end{cases} \quad (3.8)$$

imposing $\xi = 0$. The Gamma function is defined as $\Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t} dt$.

- the fitting is refined by application of the Maximum Likelihood Estimation, taking as initial values the first guess obtained by the moments method.
- the MLE method allows to determine the estimated values as well as the standard error of the estimation.

The estimated parameters are $\hat{\mu}_m = 69.80 \pm 3.8$, $\hat{\sigma}_m = 24.15 \pm 3.10$, $\hat{\xi}_m = 0.205 \pm 0.127$.

Through a series of graphs (Figure 3.5) the results of the fitting can be visually examined.

- The probability plot compares empirical CDF and modeled CDF. The ideal condition is that all the points lie on the bisector (solid line). This plot indicates that the data distribution follows a GEV.
- The quantile plot (or q-q plot) compares observed and modeled values corresponding to the same quantile. This plot helps in verify if the large values have been correctly modeled or if some outliers are present. In this case, it seems that the common as well as the extreme values are well represented.
- If the previous two plots provide satisfactory results, we can rely on the results shown by the return level plot. The critical aspect in the analysis of the Montpellier series is that 52 maxima are

not a sufficient number for a reliable estimation of the shape parameter. The shape parameter ξ influence on the estimation of the extreme behavior increases as the return period to be estimated increases. In the plot, it determines whether the black line will be straight ($\xi = 0$), convex ($\xi > 0$) or concave ($\xi < 0$). In this case, since few observations exceed the 10-years return period, the 95 % confidence levels (blue line) of the estimation are far from the average estimation (black line), indicating a longer sample is needed for a reliable estimation. The confidence interval (vertical distance between the two blue lines) is extremely large at 100 years, with a span of over one magnitude order. It is therefore difficult to define the return level corresponding to return periods higher than 50 years, taking such model and with this sample size.

- The density plot gives an idea of the distribution of the data: the observations appear to be concentrated around the distribution mode; few values can be really considered as extremes. An histogram shows if the empirical data fits well with the model (represented through a density line).

In this example, the shape parameter is positive, indicating Fréchet (hyperbolic tailed) maxima. Nevertheless, the ξ estimation is very sensitive to the outliers. Longer samples would be needed to obtain a reliable estimation of ξ . This is a delicate point, since ξ is the more significant parameter in the determination of the return levels in the extrapolation range (i.e. for T_R higher than the length of the series). The need of larger samples is one of the motivation for the use of Peaks-Over-Threshold method (Section 3.6).

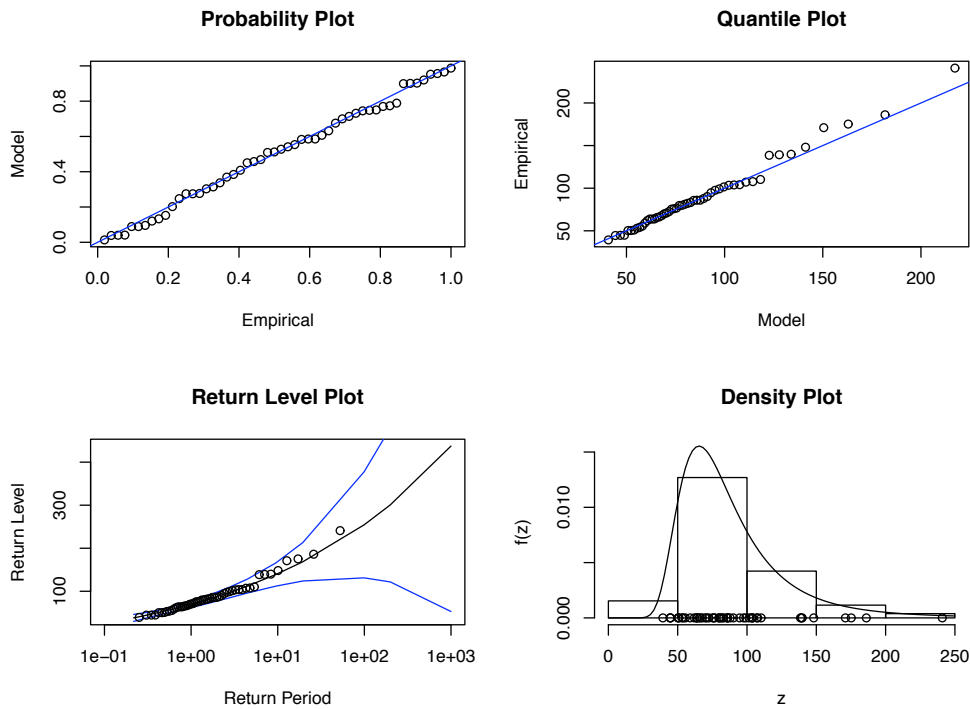


Figure 3.5: GEV summary plot for the 24-h series of Montpellier in the period 1920-1972. In clockwise order from the top left: probability plot, q-q plot, return level plot, density plot.

3.6 Analysis of exceedances: POT

The exceedances analysis, also referred to as Peaks-Over-Threshold (POT), is a method to characterize the probability distribution of extremes by selecting a number of observations exceeding a given threshold. The method consists in selecting the appropriate sample for the application of Equation 3.4 and 3.5. POT can be preferred to the block maxima analysis because it allows to select a higher number of observations.

As one can intuitively expect, the selection method (block-maxima or POT) does not modify the intrinsic behavior of the data. For this reason, the three attraction domains described for the block-maxima analysis illustrated in Figure 3.1 are still valid in the POT framework. The Gumbel attraction domain for $\xi_p = 0$ refers to exponentially-tailed exceedances; the Fréchet attraction domain models exceedances with hyperbolic (power-law) tails; the Weibull attraction domain models exceedances characterized by a higher bound.

Block-maxima and POT approaches have in common not only the presence of three attraction domains: the ξ parameter is exactly the same for the two methods. The two remaining parameters are related among each others by direct relations. The substantial equivalence of the two approaches is at the base of the point process approach (Section 3.8).

3.6.1 GPD fitting on synthetic series

Imagine to generate a random normal series. In this example, we generate 10^5 i.i.d. samples with $\mu = 2$ and $\sigma = 1$. The distribution of the variable is shown by the histogram in Figure 3.6.

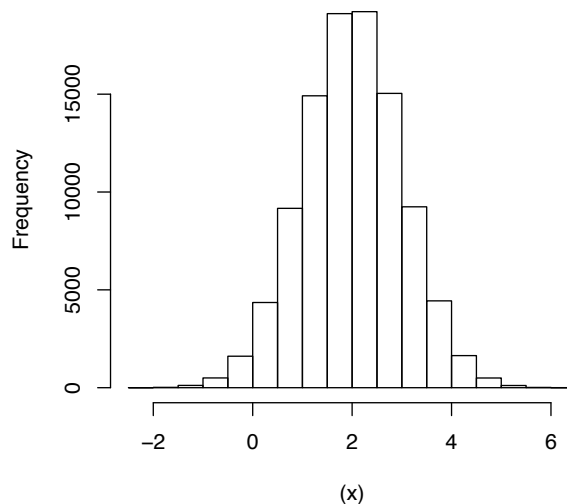


Figure 3.6: Histogram of the synthetic normal series ($n = 10^5$) with $\mu=2$ and $\sigma = 1$.

According to the theory, the observations exceeding a threshold should be GPD distributed. In addition, since the sample is normal, we expect the sample set to belong to the Gumbel attraction domain, with $\xi_p \sim 0$ (see Figure 3.1).

The first operation is the threshold selection. In synthetic series, only numerical criteria are used for the threshold choice. The “mean residual life plot” (Davison and Smith, 1990; Coles, 2001) is the graphical representation of the locus of points

$$\left(u, \frac{1}{n_u} \sum_{i=1}^{n_u} (x_i - u)\right), \quad (3.9)$$

where u is the varying threshold and n_u the number of exceedances above u . If the exceedances are distributed according to a GPD, the mean residual life plot should be approximately linear in $u > u_0$ (Coles, 2001). The Mean Residual life plot for the synthetic series is shown in Figure 3.7-a in the range 4-6, indicating a reliable threshold (linearity of MRL plot) in the range 4.5-5.5.

The second threshold-selection procedure (Coles, 2001) is the estimate of the model at a range of thresholds. Above a level u_0 at which the asymptotic hypothesis of GPD are acceptable, the estimation of the shape parameter (Figure 3.7-b upper plot) should be constant and the estimation of the scale parameter should be linear in u^3 . Poorly speaking, the threshold should be sufficiently high to incorporate only extremes and sufficiently low to have a sufficient number of values such to limit the estimation variance.

The graphs in Figure 3.7-b show the maximum likelihood estimates and confidence intervals of the shape and scale parameters over the threshold range, chosen in the range 2-6. In the case of Figure 3.7-b the graphs suggest to select the threshold within the range 4-4.8, where both the hypotheses (ξ_p constant and σ_p linear) are satisfied. A good compromise between the results derived from the two methods can be the value 4.8.

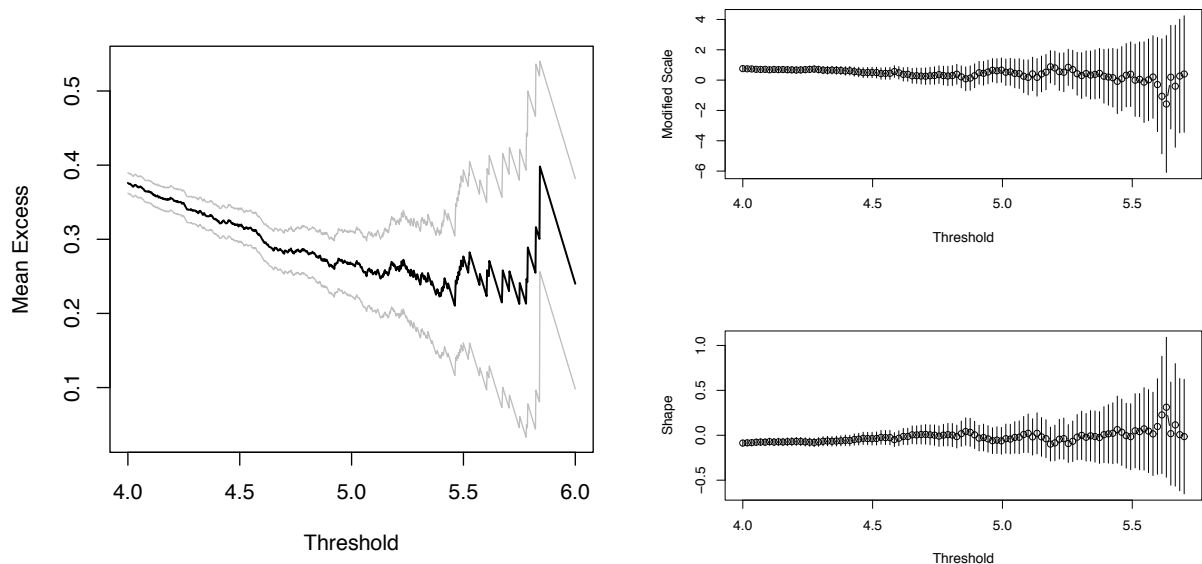


Figure 3.7: a: Mean residual life plot for the synthetic normal series of 100000 realizations with $\mu=2$ and $\sigma = 1$. The grey lines indicate the confidence interval of the estimation. b: Fitting of GPD model over a range of thresholds. The dots are the estimation and the vertical line identifies the confidence interval of the estimation.

³The MLE estimation by interactive threshold selection can be performed by using the function `gpd.fitrange` in the package `{ismev}`

Once selected the threshold, the GPD fitting can be performed through a MLE estimate⁴. The fitting results in a scale parameter $\beta_p = 0.284 \pm 0.034$ and in a shape parameter $\xi_p = -0.060 \pm 0.084$.

Similarly to Section 3.5, it is possible to draw graphs indicating the quality of fitting and the behavior of extremes (Figure 3.8). In the present case, the probability plot and the quantile plot well behave, the empirical observations perfectly lie on the bisector. Supposing that our sample was composed of hourly observations, we can draw the return level plot. It shows a slight concavity of the model ($\xi = -0.06$). This has no consequences for small return periods, but it can affect the estimation for large return periods. Having generated a Gaussian sample, we expected a null shape parameter. Finally, the density plot shows how the sample is apparently well fitted to the model.

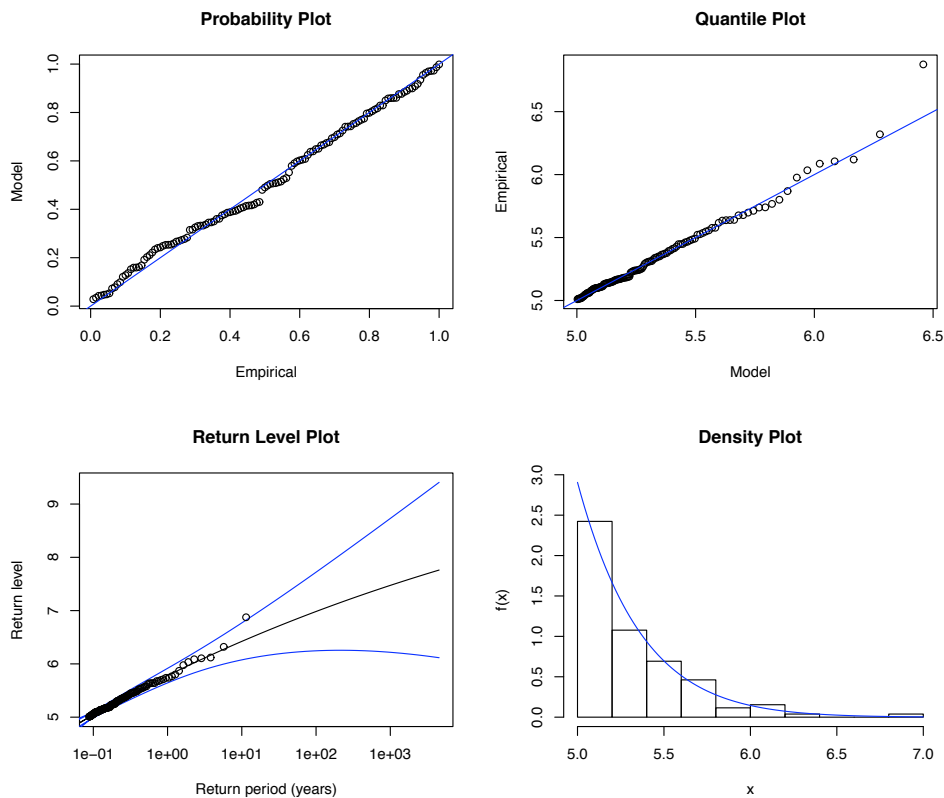


Figure 3.8: Result of the fitting of a GPD on the exceedances ($x \geq u_d = 5$) of Gaussian sample of $n=100000$, $\mu = 1$ and $\sigma = 1$.

The POT analysis on a Gaussian sample is easy because the sample tends to have exponentially-tailed extremes (maxima will be Gumbel distributed and exceedances will be negative exponentially distributed). More awkward is the case of positively skewed samples (e.g. log-normal), where the convergence of the tails to exponential is extremely slow and can lead to spurious values of ξ_p , even with very large samples.

⁴For example, the function `{gpd.fit}` of the `{ismev}` package do this kind of estimation. `gpd.diag` allows to draw the diagnostic plots.

3.6.2 GPD fitting on real series

In order to compare the block maxima and POT method on a practical application, the POT method is hereafter illustrated using the rainfall series of Montpellier (52 years) to estimate the extreme daily rainfall intensities or frequencies. Suppose that we want to estimate the extremes behavior for the accumulation duration of 24 h. We can build a 24-h observations database by means of a moving-average scan.

The main difference with the analysis on a synthetic series is the dependence among observations. Many of the extreme observations appear clustered; this is in contrast with the implicit hypothesis of independence between observations.

Figure 3.2 shows that for the analyzed series, two observations are practically uncorrelated for time lags higher than 100 hours. We can use this information for the POT analysis⁵: only the independent observations separated by a lag higher than the decorrelation time will be retained.

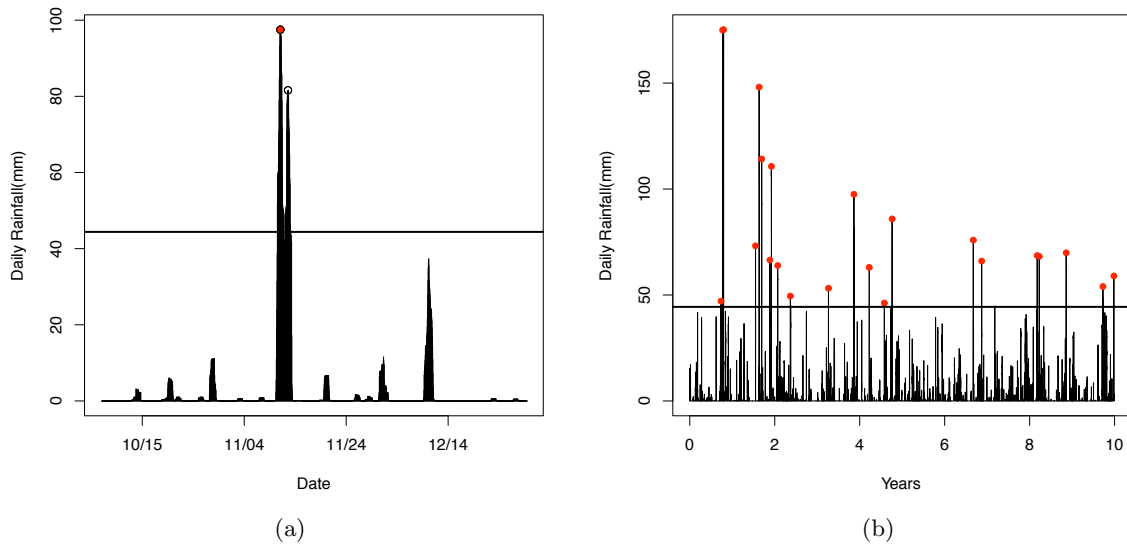


Figure 3.9: Selection of observations exceeding a selected threshold of 44.4 mm/day (horizontal line) corresponding to an average of 3 events per year. a: Particular of the de-clustering process for removing the dependent peaks: only the solid red observation is retained. b: Selection of data above a threshold and separated by a lag of 100 h.

The threshold choice could be the result of a numerical analysis, similarly to Section 3.6.1. Figure 3.10 reports the mean residual life plot for the station of Montpellier for thresholds higher than 1 mm. The grey lines represent the 95 % confidence level for the estimate.

According to Figure 3.10, it exists a range of u_p (30-75 mm) where the MRL plot is linear. The threshold should be detected in this range.

⁵The Peaks-Over-Threshold analysis here shown is performed by using the package POT (Ribatet, 2007); the preliminary de-clustering is performed by means of the function {clust} in the same package.

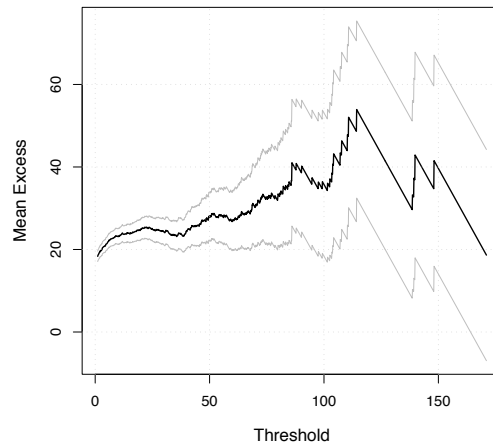


Figure 3.10: Mean Residual Life Plot of the i.i.d. exceedances for the accumulation duration of 24 h - Rain Gauge series of Montpellier.

Again, we can fit the GPD over a range of thresholds (Figure 3.11) in order to determine the value for which the estimation bias and variance reach the better compromise. In the range $44\text{-}65 \text{ mm } d^{-1}$, the parameters u_p and β_p are relatively stable and exhibit low estimation variance.

The Mean Residual Life plot and the fitting of the GPD over a range of thresholds provide information about the range in which the threshold should be selected. This threshold corresponds to an average of 2-4 events per year. A good compromise can be to choose the threshold as the value corresponding to 3 events per year ($\hat{u}_p = 44.4 \text{ mm } \text{day}^{-1}$).

Once fixed the threshold, the maximum likelihood estimator for GPD can be used⁶. The estimation gives $\hat{\beta}_p = 24.79 \pm 3.42$ and $\hat{\xi}_p = 0.118 \pm 0.102$.

The results can be evaluated by means of 4 different graphs, similar to those reported in Figure 3.5: probability plot, q-q plot, return level plot, density plot. See 3.5 for a detailed description of each plot.

3.7 Comparison Block-Maxima - POT

In the previous chapters, we described the two widespread approaches for the modeling of extreme values. The two methods (Block-Maxima and Peaks-Over-Threshold) are the practical application of the extreme value theorem for modeling the maxima and the exceedances-over-threshold of a sample, respectively. The extreme value theorems state that a series of i.i.d. maxima of a random sample are GEV-distributed, while a series of exceedances over a fixed threshold is GPD-distributed. The two approaches are intimately related: (Salvadori and De Michele, 2001) show that the GEV and the GPD parameters can be related through the following equalities:

⁶function fitgpd{POT}

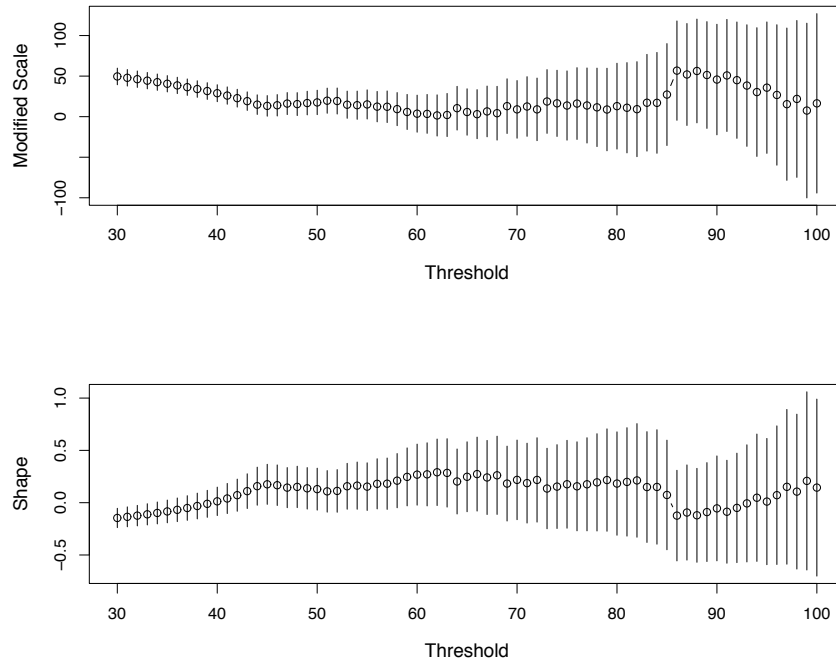


Figure 3.11: GPD fit of the parameters σ_p and ξ_p as a function of the threshold u_p .

$$\begin{cases} \mu = u_p - \frac{\beta_p}{\xi} (1 - n^\xi) \\ \sigma = \beta_p n^\xi \\ \xi = \xi_p \end{cases} \quad (3.10)$$

where n is the average number of events over a block (i.e. a year) selected in the Peaks-Over-Threshold method. If one value is taken in average for each block ($n = 1$), we have $\mu = u_p$, $\sigma = \beta_p$ and $\xi = \xi_p$. However, practical tests on synthetic series show that the transformation actually provide different results respect to the direct estimation. The two estimation methods do not give perfectly coincident results either in case of extremely long series.

In order to compare the two extreme modeling methods, we examine the results obtained at the rain gauge station of Montpellier (featuring 52 years of hourly records) at the time scale of 24 h.

Table 3.2 shows a summary of the obtained results from which we can draw interesting conclusions: since the number of samples is consistently smaller in the Block-Maxima approach, we expect the estimation variance to be higher, as for the ξ parameter. A single outlier can sensibly modify the estimation of GEV parameters, while in GPD its influence is less marked. The deviation of the GPD ξ_p parameter towards 0 can be a consequence of gathering regular observations into the extreme sample. The return level for $T_R = 100$ yrs shows a good agreement between the two methods. The POT return level is lower than the Block-Maxima result, presumably because of the insertion of regular values into the sample, leading to higher biases in the estimation of the extreme behaviour. However, the 3-parameter GEV fitting with a sample of only 52 realization is unsafe, leading to higher

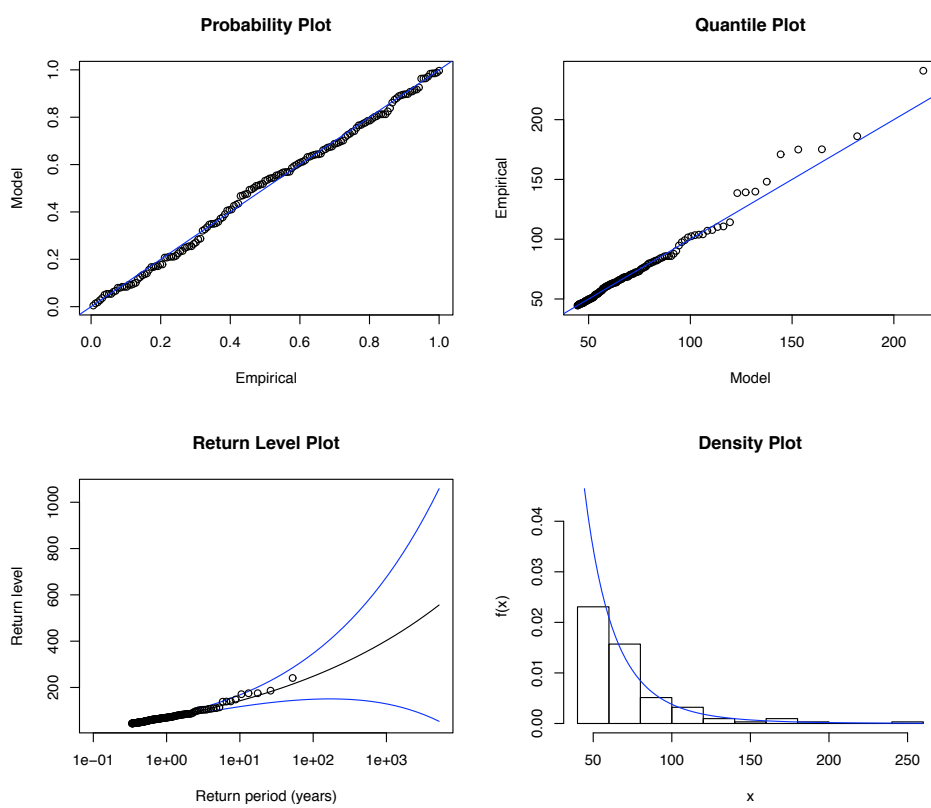


Figure 3.12: Results of the POT analysis on the rain gauge series of Montpellier for $D=24h$, taking an average of 3 events per year.

estimation variance. For this reason, the use of Peaks-Over-Threshold method should be preferred dealing with limited samples.

In general terms, we can detect the main advantages/drawbacks of block-maxima versus the Peaks-Over-Threshold method for the determination of the behavior of extremes:

- theoretical framework. The two approaches are both theoretically correct and intimately related: direct relation exists between the GEV and GPD parameters.
- infinite sample size. At the limit of infinite sample size, the two methods are totally equivalent giving equal results in terms of return levels;
- finite sample size. Block-maxima approach is limited to one sample per block, excluding possible extremes occurring in the same block; in some cases the sample size of the sample maxima does not allow a reliable estimation;
- reliability of the estimation: practical tests demonstrated that the minimum sample size for having reliable estimation of the GEV parameters is 100;
- arbitrariness. Block-maxima approach is objective and perfectly repeatable while POT approach is submitted to the choice of the threshold by the operator (even though helped by indicators such as the mean residual life plot or by physical reasoning);
- dependence among data. The block-maxima approach implicitly removes dependent observations; Peaks-Over-Threshold approach can be performed in real series only after de-clustering of data.
- appropriate data use. With block-maxima approach most of the information is discarded; Peaks-Over-Threshold approaches allows a better use of the information contained in the database increasing the sample size;
- bias. Taking one sample per year, block-maxima samples is likely to contain only extremes; POT samples may contain regular values and then the estimation may be biased towards the regular values;
- variance. The size of samples is such that the block-maxima approach result in very high variance estimation (unless to fix a priori one parameter, e.g. $\xi = 0$ or to establish relations between variables, μ and σ for instance); the POT estimation variance is generally lower.

Method	Distrib.	Sample Size	Position (mm)	Scale (mm)	Shape	$T_R = 100$
Block-Maxima	GEV	52	$\mu=69.8\pm 3.8$	$\sigma=24.5 \pm 3.10$	$\xi=0.205\pm 0.127$	257 mm
Peaks-Over-Threshold	GPD	156	$u_p= 44.4$ (fixed)	$\beta_p=24.79 \pm 3.42$	$\xi_p=0.118\pm 0.102$	247 mm

Table 3.2: Comparison block-maxima and Peaks-Over-Threshold methods for the modeling of rainfall extremes at the station of Montpellier Bel-Air, duration 24 h. The data is relative to the period 1920-1972, where no significant climatic change evidences have been detected.

3.8 Point process: an unified framework for extreme analysis

In the previous sections we have seen that the Block-Maxima method directly provides results as a function of the return period and therefore it is easy to be managed; POT method gives more robust estimations because it is possible to select a larger number of observations. Since the two approaches have more than one common property, the interest of the scientific community was to find an unified approach.

The point process (PP) theory (*Coles, 2001*) is a recent and elegant formulation that provides an interpretation of extreme value theory unifying the classic models. Some basic concepts were early introduced by *Pickands (1975)*. The models described in Section 3.5 and 3.6 can be derived from particular cases of the point process theory. We consider noteworthy to report the outline of the point process theory as a necessary step for an unified theory of extremes.

As highlighted by *Coles (2001)*, the main advantage of using a point process model is to obtain more robust samples than the block-maxima approach, keeping the easy formulation and parametrization of GEV in which the probability can be directly related to the return period through the relation $P = 1 - 1/T_R$.

The derivation of the point-process theory is based on the consideration that the number of events exceeding a sufficiently high threshold u is Poisson distributed and their intensity is given by the extreme value theory.

In practice, if block-maxima approach models the sample composed by the maxima of each block, and POT models the exceedances over a given threshold, PP models the number of exceedances over a threshold within a block (Figure 3.13), knowing that the intensity can be modeled with GPD and the number of exceedances through a Poisson process.

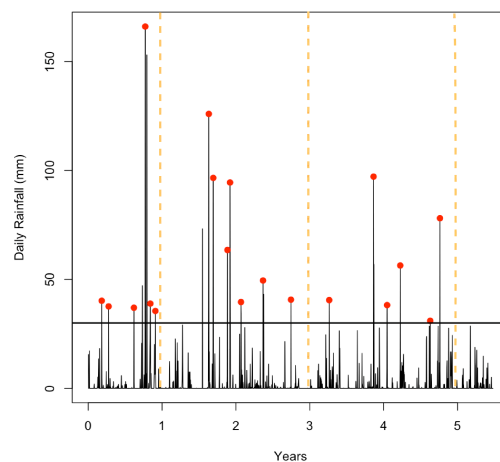


Figure 3.13: Scheme of the point-process theory. The number of observations (red dots) that exceed a given threshold (horizontal black line) within a block (yellow dashed line) is Poisson-distributed.

Coles (2001) gives the mathematical derivation of the PP model and the maximum likelihood function.

3.9 Stationarity of the rainfall series

The increasing concentration of atmospheric carbon dioxide in the last decades has almost certainly led to changes in the global mean temperature. If there are incontestable evidences of the temperature increase after 1970, the effects or climate change on the precipitation is less clear.

As shown in many papers (*Palmer and Räisänen, 2002; Zhang et al., 2008; Naidu et al., 2009; Allamano et al., 2009; Costa and Soares, 2009*), the detection of trends in rainfall or river flow series is a tough work due to the rainfall variability and series availability. Rarely, such analyses have highlighted a significant trend in the rainfall series.

Having a database covering about 15 years of hourly rainfall data (1993-2008), we are unable to detect a drift in the behavior of heavy rainfall.

For longer series, such as the reference series of Montpellier, we can apply methods to estimate the behavior of extremes for two disjoint sub-periods. If a trend is detected, the extreme value theory cannot be applied anymore unless expressing the drift of parameters with time.

To verify that no significant trend is present on the rainfall series, we analyze the rainfall excesses through a POT analysis. To have a robust sample without including common observations, we fix the number of excesses in 4 per year. We fix the shape parameter equal to zero, due to the poor number of i.i.d. observations (considering that we just want to check the stationarity, we may fix this constraint).

We compute the GPD parameters for 11-year moving windows, as suggested by *Naidu et al. (2009)* (Figure 3.14).

In the period 1920-1972, no significant trends are observed.

In any case, the estimation of large return levels can be done only after verification of the stationarity hypotheses. The MLE estimation methods allow to modify the expression in order to take into account possible trends in the behavior of extremes due to climatic change.

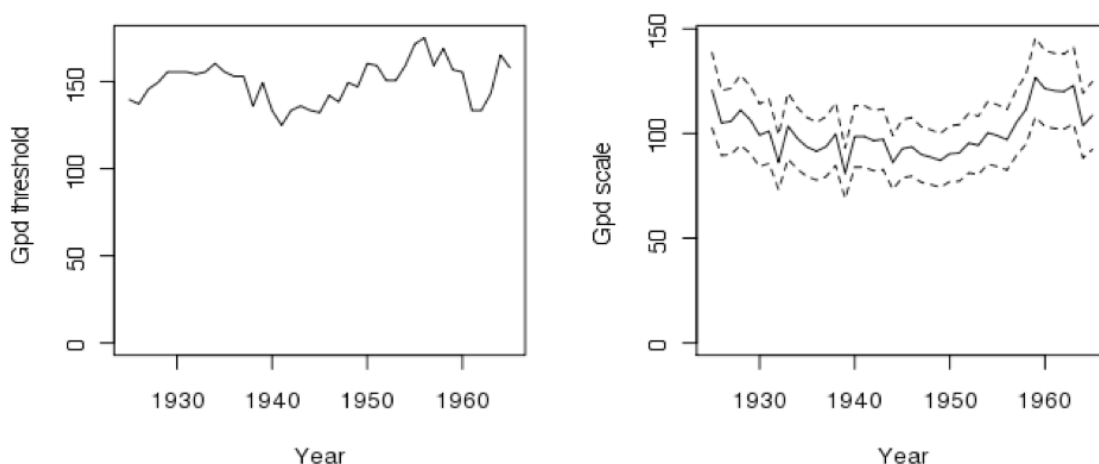


Figure 3.14: Temporal evolution of the threshold (a) and of the scale (b) parameters of the GPD from a POT analysis of the Montpellier rainfall series. The parameters are computed over 11-year sliding periods. An average of 4 excesses per year are selected.

3.10 Conclusion

In this chapter, we described the theory and the practical methods to model the extremes of a process. The modeling of extremes is submitted to a series of hypothesis to be respected, such as the independence of observations and the identical distribution of the parent process.

The commonly used methods have a common origin and lead to the same results in terms of estimation of the return levels, even if they involve different sample-extraction methods and a different law for the extremes. Each model has some advantages and drawbacks and we gave insights on the method to be preferred as a function of the data size and data dependence. Applying the main methods for the modeling of the daily extremes⁷ of a rainfall series covering 52 years of data, we found similar results both in terms of parameter ξ and of return level for $T_R = 100$.

In the final part we provide some elements concerning the stationarity of rainfall series. The classic extreme analysis can't be applied in non-stationary series. The extreme value parameters, in this case, must be expressed as a function of the time. Due to the evidences of thermal increase after 1970, a modification of the weather types and a consequent modification of the extreme precipitation events is expected. Nevertheless, the precipitation series, especially at the hourly resolution, are not long enough to establish whether a precipitation modification has occurred in the last decades due to climatic change.

⁷in Appendix C we provide a tutorial on how to perform a generic extreme analysis on synthetic or real data with R.

1 Geostatistics

Résumé

Cette section est consacrée à la description des techniques d'analyse spatiale et des techniques d'interpolation connues sous le nom “géostatistique”. La géostatistique est largement utilisée dans les géosciences pour caractériser la structure spatiale des variables aléatoires ou des champs aléatoires tels que les dépôts minéraux, l'hauteur piézométrique d'une nappe phréatique, les précipitations sur un bassin, la température et la pression dans l'atmosphère, la concentration de polluants. La variabilité de ces phénomènes empêche de les décrire avec des fonctions mathématiques simples, telles que la régression linéaire.

*Le chapitre est organisé de la façon suivante: dans le premier paragraphe, nous définissons les propriétés que un champ aléatoire doit présenter afin de pouvoir utiliser la géostatistique pour son analyse spatiale. Après, nous présentons brièvement la théorie de l'**analyse variographique** montrant les applications 1D et 2D, y compris des cas particuliers de champs avec dérive, champs anisotropes, présence de co-variables. Enfin, nous présentons l'outil d'interpolation géostatistique “**krigeage**”, qui exploite les informations contenues dans le variogramme pour effectuer la meilleure estimation linéaire de la valeur d'un champ à des endroits non instrumentés*

4.1 Introduction

This section is devoted to the review of the spatial analysis and spatial interpolation techniques known as **geostatistics**. Geostatistics is widely used in geosciences to characterize either the time or space structure of random variables or random fields such as grade of mineral deposit, depth of piezometric heights, rainfall depth, temperature and pressure in the atmosphere, pollutants concentration.

The chapter is organized as follows: in the first paragraph, we define the properties that a field must exhibit in order to use geostatistics for its spatial analysis. After, we briefly present the **variogram theory** showing 2D applications including special cases of drifted fields, anisotropic fields, multi-field analysis and presence of co-variables. Finally, we present the geostatistical interpolation tool known as **kriging**, that exploits the information contained in the variogram to perform the best linear estimation of the value of a field at ungedged locations.

4.2 Stationarity of random functions

The structural analysis of a random field is differently conducted depending on the kind of stationarity that the field exhibits. We discuss in the following the concepts of strict stationarity, first and second-order stationarity, intrinsic hypothesis and ergodic hypothesis.

A random function $Z(X_1, \dots, X_h)$ is strict-sense stationary (*Chiles and Delfiner, 1999*) where its finite-dimensional distributions are invariant under any arbitrary translation:

$$P[Z(x_1) < z_1, \dots, Z(x_k) < z_k] = P[Z(x_1 + h) < z_1, \dots, Z(x_k + h) < z_k] \quad (4.1)$$

poorly speaking, a random function is strict-sense stationarity when its probability distribution does not vary with the location of the x_n points. This equality is valid for the whole statistical distribution, and therefore it concerns all the quantiles and moments. In many cases, instead to the whole distribution, we may be interested to a limited number of moments. Limiting the analysis to the first two moments of the distribution, the field is second-order stationary (wide-sense stationary) if the mean m is constant and the covariance C only depends on the separation h :

$$\begin{cases} E[Z(x)] = m \\ E[(Z(x) - m)(Z(x + h) - m)] = C(h) \end{cases} \quad (4.2)$$

The structural analysis can be conducted even if the mean of the process Z is not a constant but linearly increasing/decreasing with x . In this case the stationarity criteria are applied to the increments $Y_h(x) = Z(x + h) - Z(x)$. $Z(x)$ is called an “intrinsic random function”, and its statistical moments are defined as:

$$\begin{cases} E[Z(x)] = \langle a, h \rangle \\ Var[(Z(x + h) - Z(x))] = 2\gamma(h) \end{cases} \quad (4.3)$$

where $\langle a, h \rangle$ is the linear drift of the intrinsic random function and $\gamma(h)$ is the variogram function (Section 4.3).

The ergodic property is another insightful concept in geostatistics. The realization ω of the random function Z at the generic point x is noted $Z(x, \omega)$. In case of ergodicity (of the mean), one can infer the mean of a random function Z even if only one realization at the point x_i is known. Poorly speaking, the time average of the process $Z(x, \omega)$ function along the trajectories is related to its spatial average.

A stationary random function $Z(x, \omega)$ is ergodic in the mean if the spatial average of $Z(x, \omega)$ over a domain $V \subset R^n$ converges to the expected value $m = E[Z(x, \omega)]$ when V tends to infinity:

$$\lim_{V \rightarrow \infty} \frac{1}{|V|} \int_V Z(x, \omega) dx = m \quad (4.4)$$

4.3 Theoretical variogram

Let us consider a (wide-sense) stationary random function $Z(x)$ characterized by its mean m and its covariance function $C(h)$:

$$\begin{cases} m = E[Z(x)] \\ C(h) = E[(Z(x) - m)(Z(x+h) - m)] \end{cases} \quad (4.5)$$

The covariance function describes the spatial correlation of the field as a function of the separation distance h . It can be computed only in case the mean is stationary (i.e. stationary random functions). If the mean is not constant (as in intrinsic random functions), the variogram γ can be a diagnostic of the spatial structure of Z :

$$\gamma(h) = \frac{1}{2} \text{Var}[Z(x+h) - z(x)] \quad (4.6)$$

Both the covariance and the theoretical variogram are even functions ($\gamma(h) = \gamma(-h)$ and $C(h) = C(-h)$). The covariance function as well as the variogram are definite positive. If Z is a stationary random function, the variogram and covariance are strictly related:

$$\gamma(h) = C(0) - C(h) \quad (4.7)$$

where $C(0)$ is the theoretical variance of the process measured at 2 points with $h \rightarrow 0$.

The variogram gives indications about the correlation structure of a variable in space. In totally uncorrelated fields (e.g. white noise), the correlation is independent of the spatial lag, therefore the variogram will be a constant. A constant field will have null variogram whatever the lag.

4.4 Sample variogram

The theoretical variogram (*Matheron, 1965*) defines the correlation of the process at two points separated by a lag h . Working with real data, the variogram of the process can be estimated more or less accurately depending on the sampling features.

In one-dimensional case, the sample variogram of Z can be computed for each couple of points x_n and x_{n+1} :

$$\gamma(h_{n,n+1}) = [z(x_n) - z(x_{n+1})]^2 \quad (4.8)$$

where $h_{n,n+1}$ is the spatial lag between the two points (in 1-D corresponds to $|x_n - x_{n+1}|$). An example of application to real data is reported in Figure 4.1. The sample is composed by 50 equally-spaced measures of a Gaussian process (Figure 4.1-a).

Representing all the obtained values of γ respect to the distance h , we obtain the “variogram cloud” (Figure 4.1-b).

The variogram cloud presents difficult interpretation: the sample variability is high and it is difficult to find a regular behavior of the points. In order to include a sufficient number of points in the variogram computation, the sample variogram is usually calculated for distance classes with center h , as:

$$\hat{\gamma}(h) = \frac{1}{2N_h} \sum_{|x_i - x_j| \simeq h} [z(x_i) - z(x_j)]^2 \quad (4.9)$$

Averaging the sample variogram by regularly spaced classes of distances, an efficient description of the spatial structure of the process can be obtained (Figure 4.1-c).

This definition can be generalized to data in \mathfrak{R}^D with $D=1,2,3$; the sample variogram at each points is computed as in Equation 4.8; the 1-D coordinate x is substituted by the coordinate vector \mathbf{x} .

Let us now take an example of a 2D Gaussian spatial field defined over a grid of size $100 \times 100 \text{ km}^2$, with decorrelation distance equal to 50 km (Figure 4.2-a)¹. The field has unitary variance. The sample variogram of the field is computed for classes of distances whose centers are separated by 10 km (Figure 4.2-b)².

From Figure 4.2-b we can extract elements about the meaning of sample variogram. The empirical variogram, in this case, shows $\lim_{h \rightarrow 0} \gamma(h) = 0$, meaning that points that are close in distance are well-correlated. This is not always the case with empirical variograms. $\gamma(0) > 0$ may indicate that either the sample variogram has missed the small-scale variability (under-sampling) of the field or the point estimation is submitted to a significant sampling uncertainty. In this case, the empirical variogram is said to have a “nugget” (the term is related to the first field of application of variograms, mining engineering). A white noise is referred to as “pure nugget process”. If the empirical variogram reaches an asymptotic value, as in Figure 4.2-b, this is called “sill” and corresponds to the variance of the field. The distance for which the asymptotic value is reached is referred to as “range”, i.e. the decorrelation distance of the field. Two points further than the variogram range are likely to be completely uncorrelated.

In some cases, the sample variogram does not reach a finite asymptotic value: this can be due to the limited window size (in this case, in addition, the variance of the field will not correspond to the maximum value of the variogram), or to the presence of a drift (the field has not constant mean, Equation 4.2). In the latter case, a de-trended variogram can remove this effect. In other cases, an unlimited variogram could indicate the presence of long-range correlation (Section 5.9.1).

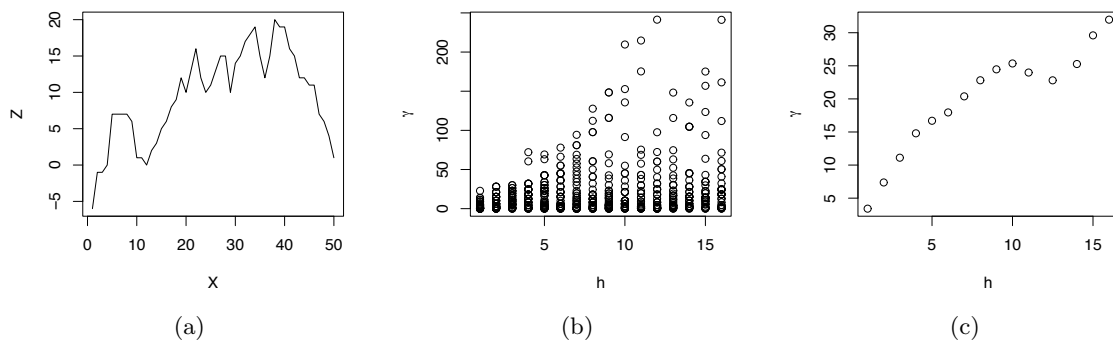


Figure 4.1: a: Simulation of a Gaussian process in one dimension, measured at 50 equally spaced points. b: Variogram cloud obtained representing the variogram of each couple as a function of the couple distance h . c: Sample variogram obtained averaging the variogram for a finite number (15) of distance classes.

¹The fields are simulated in R with the GaussRF function in the package RandomFields with the following code: `x=1:100; y=1:100; FIELD=GaussRF(x,y,grid=TRUE,model="spherical",param=c(0,1,0,50))`

²The function EmpiricalVariogram in the R package RandomFields is used: `EmpiricalVariogram(data=FIELD,x=1:100,y=1:100,grid=TRUE,bin=seq(0,100,10))`

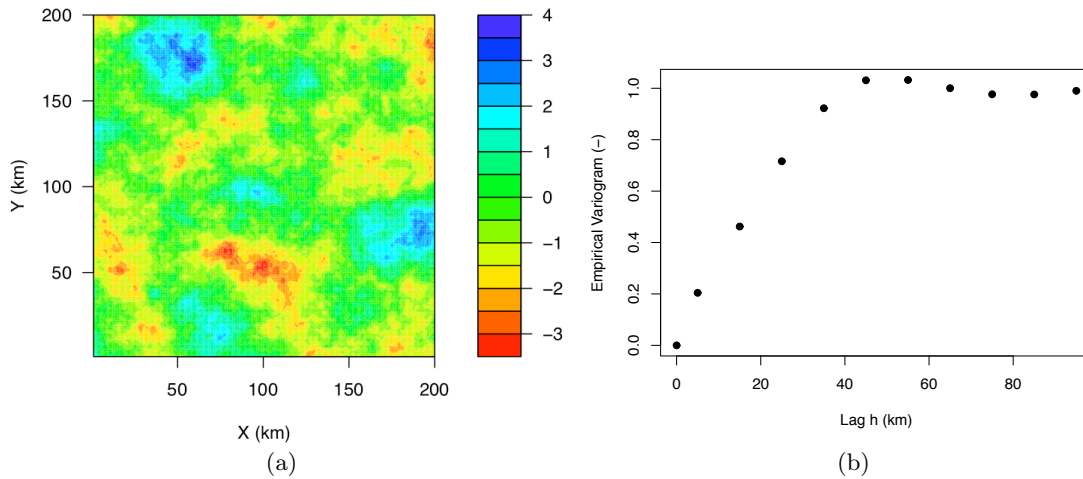


Figure 4.2: a: Generation of a random field with decorrelation distance of 50 km over a 200x200 km grid. b: Empirical variogram of the field for lags up to 100 km, for classes of width 10 km.

4.4.1 Anisotropic variogram

In spatial data, we often have to face with elongated structures, where the correlation distance is likely to be higher along one particular orientation axis. In these situations, it may be useful to compute the variogram by taking into account the orientation of the considered couple with respect to the North, and to regroup the observations by angle classes. The resulting variograms are not only function of $|h|$ but also of the orientation of the \mathbf{h} vector, and they are said to be anisotropic.

Evidences of elongated structures are common studying rainfall. This is the case, for example, of the orographic band highlighted in *Miniscloux et al.* (2001) that contributes up to 40% to the total rainfall amount recorded in mountainous regions (*Godart, 2009*).

We show here an example of anisotropic instantaneous rainfall field. The radar scan of 08 September 2002 at 22:45 UTC shows an elongated structure towards NNE, i.e. 30° with respect to the North (Figure 4.3-a). We compute the anisotropic variogram for 4 angles: 30° , 75° , 120° , 165° (Figure 4.3-b), showing that, for 30° , the sill is far to be reached while, for 120° , an asymptotic value is reached for distances in the range 20-25 km.

4.4.2 Variogram of an intrinsic random function

Previously, we have seen that the unknown mean is a limit for the determination of the covariance function (e.g. Equation 4.2), and that in these cases variogram shall be used. However, the use of variograms may lead to several issues: when the field is the sum of a second-order stationary random field but the mean depends on the spatial coordinates, the variogram does not provide a reliable description of the structure of the field (*Matheron (1972)* describes this problem). *Goovaerts (1997)* demonstrated that removing the field drift we can then perform simple variogram on the residuals. This technique is called detrended variogram.

A simple example of the utility of detrended variograms is shown in Figure 4.4. A mono-dimensional variable has a clear drift (Figure 4.4-a), i.e. the average of the process depends on the x

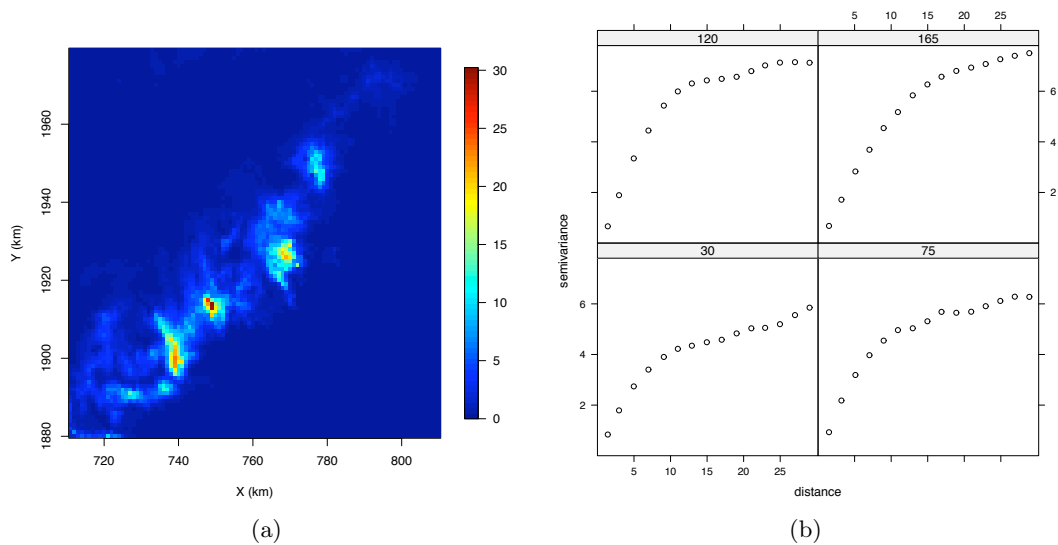


Figure 4.3: a: Instantaneous radar scan recorded on 08 September 2002 at 22:45 UTC. b: histogram of the logarithm of the field. b: Anisotropic variogram for 30, 75, 120 and 165°.

coordinate. This behavior causes that the points separated by large distances have large variograms; in particular, the variogram takes a parabolic behavior (Figure 4.4-b solid line) because the variogram is proportional to $[Z(x) - Z(x + h)]^2$. To remove this artifact, a detrended variogram is computed. The operation corresponds to the variogram computation on the field obtained by removing the drift. As we can see, the de-trended variogram (red dashed line) reaches a sill differently from the simple variogram.

In Figure 4.5-a, a 2D drifted field is reported. The presence of a trend is difficultly detected when the field is represented as an image. Nevertheless, the empirical variogram (Figure 4.5-a, solid circles) shows a parabolic increase. De-trending the variogram as a function of the coordinates X and Y solves this issue, depurating the drift effect and showing that actually an asymptotic value of the variogram is reached.

4.4.3 Indicator variogram

The rainfall spatial structure can be dependent on the intensity level. Varying the intensity threshold the spatial conformations or anisotropy conditions may vary leading to different variogram models. *Journel* (1983) and *Goovaerts* (1994, 1997) introduced the concept of indicator variograms to deal with this property of natural processes. *Barancourt et al.* (1992) discussed about the possibility to deal with rainfall fields separating the rain-no rain intermittency from the rainfall variability. They demonstrated that this separation is possible, and these two processes could be defined by two different variograms. The “Indicator variogram” method uses the intensity of a process as the discriminant factor for computing different spatial analyses. In Figure 4.6-c, the composition of two different fields for rain-no rain intermittency (a) and positive rainfall (b) is shown. Figure 4.6-d reports the indicator variogram for the intermittency field and for the positive rainfall field.

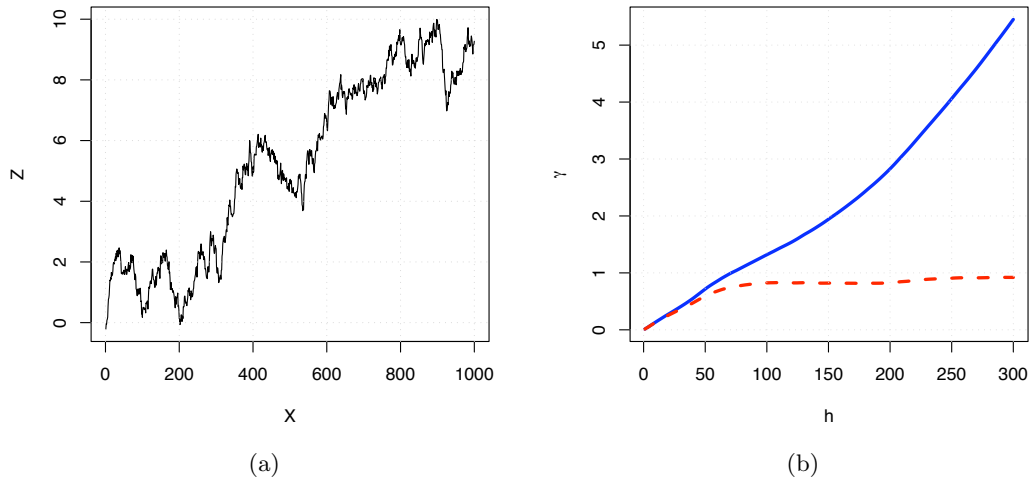


Figure 4.4: Effects of a drift on the variogram computation. a: 1-D process stochastic process (correlation distance = 100) with a drift. b: Simple and de-trended variograms of the field.

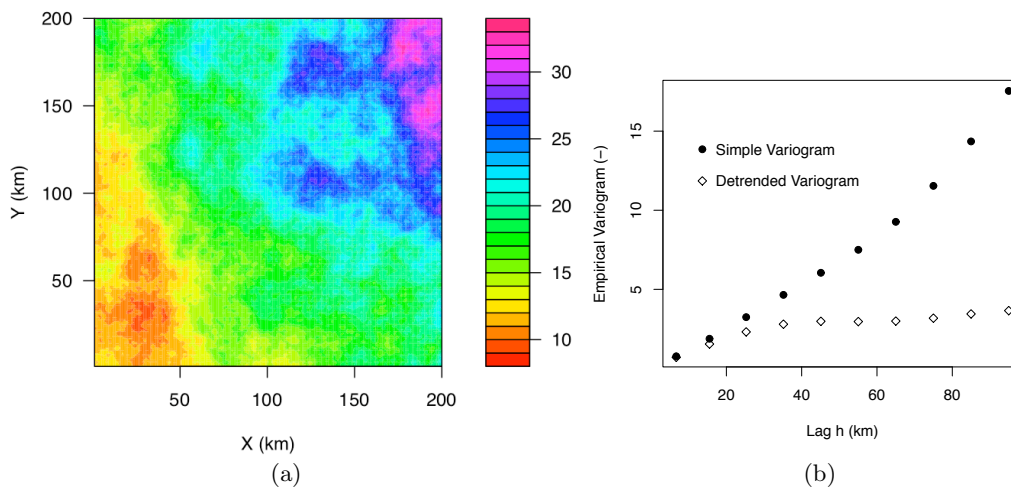


Figure 4.5: a: Gaussian random field with trend (dependent on X and Y axis) over a 200x200 km grid. b: Empirical variogram of the field for lags up to 100 km, for classes of width 10 km; black dots= simple variogram; diamonds=de-trended variogram.

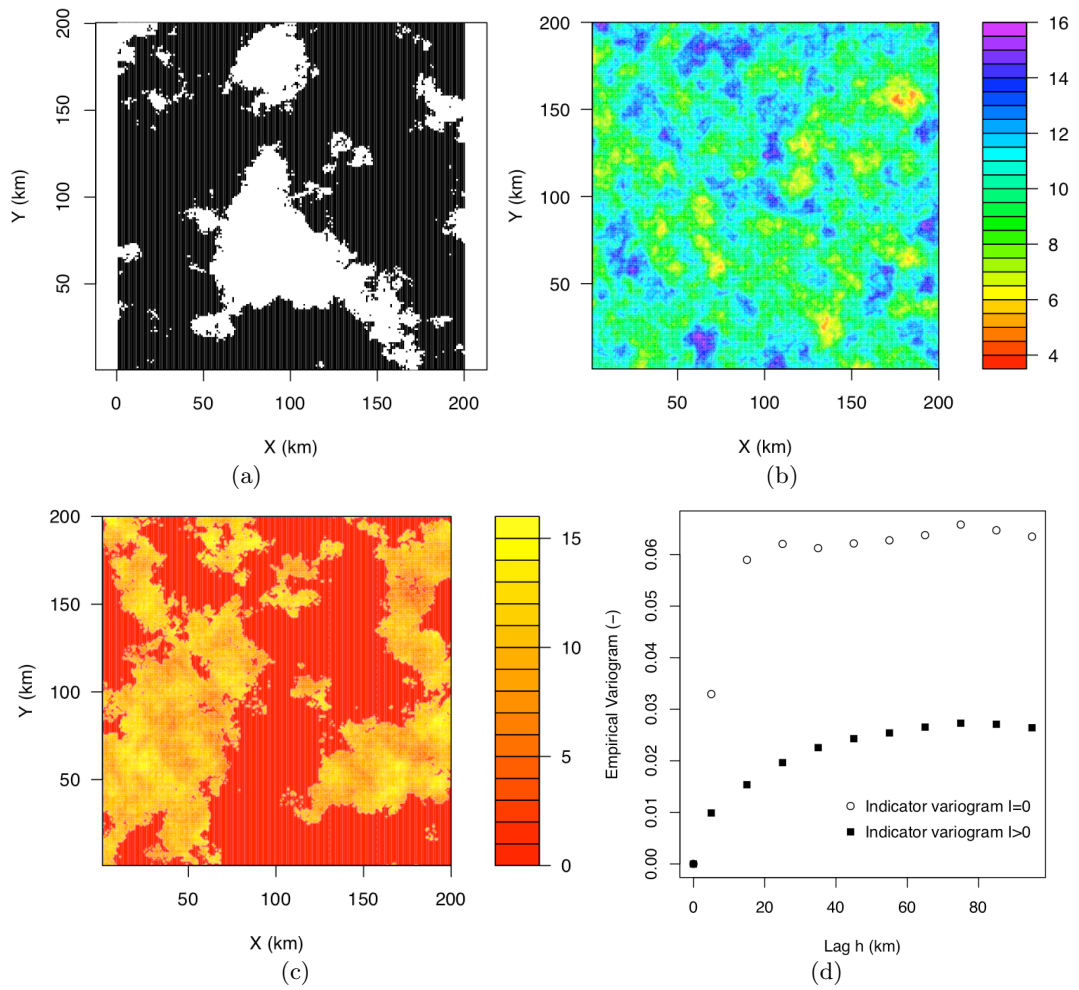


Figure 4.6: a: Rain-no rain intermittent field (decorrelation range=50 km). b: Positive rainfall field (decorrelation range = 20 km). c: Composite rainfall field obtained by product of the two previous fields. d: Indicator variograms: rain-no rain variogram (squares), positive rainfall variogram (circles).

The indicator variogram can be generalized to a number of intensity levels: generally, the higher the intensity of the field, the smaller the correlation range (and the larger the singularity degree).

4.4.4 Climatological variogram

Originally, variograms were conceived for the use in mining engineering. In this context, the observations constitute a single-realization field of the random variable. The extension of geostatistical methods to atmospheric multi-fields data requires an adaptation: the determination of the spatial structure of highly intermittent random variables such as the storm rainfall through the spatial analysis of single fields can lead to spurious results. This motivates the earlier studies concerning the definition of an average variogram (*Delhomme and Delfiner, 1973*). *Lebel and Bastin* (1985) introduced the concept of scaled climatological variogram with the intention to diagnose the spatial structure of K realizations of a random function Z . Its implementation consists in dividing each realization by its standard deviation. The climatological variogram is then determined averaging the resulting variograms for each distance class h :

$$\hat{\gamma}(h) = \frac{1}{2KN_h} \sum_{x_i-x_j \sim h} \sum_{k=1:K} \frac{[z_k(x_i) - z_k(x_j)]^2}{s_k} \quad (4.10)$$

where $N(h)$ is the number of points within the distance class h , K is the number of events, s_k the sample standard deviation of the field.

In this way, each field has unit variance and therefore has the same weight in determining the resulting variogram, which has been referred to as ‘‘climatological’’. The climatological variogram has been extensively used (*Bastin et al., 1984; Lebel and Bastin, 1985; Lebel and Laborde, 1988*) for determining the spatial structure of monthly rainfall maxima for limited surfaces (up to 400 km²) providing also a framework for the frequency-evaluation of extreme spatial-rainfall events.

Figure 4.7 reports a practical example of the normalization required to evaluate a climatological variogram based on three realizations of a random function with the same spatial structure but different range of intensity and therefore different variances of the field. The three fields in Figure 4.7-a,b,c have different ranges of intensity. The variogram computation (Figure 4.7-d) gives insights on the structure of the field but the magnitude of the variogram is extremely different in the three cases. By normalizing each difference as reported in Equation 4.10 by the observed standard deviation of the field, the variogram is defined in the range 0-1³ (Figure 4.7-d). Averaging the three normalized fields in Figure 4.7-e we may obtain indications about the average spatial structure of the random variable.

4.4.5 Variogram models

Sample variograms give a representation of the correlation between point separated by a lag h . The sample variogram is provided at a discrete number of points, corresponding to the center of classes. To give a schematized picture of the spatial structure of the fields, it is frequent to fit the empirical

³The sill=1 is reached when the correlation range is $r \ll d$ where d is the size of the domain. In the other case, due to the under-estimation of the variance in case of correlated data, the variogram sill can be higher than 1 (*Delclaux and Thawin, 1993*).

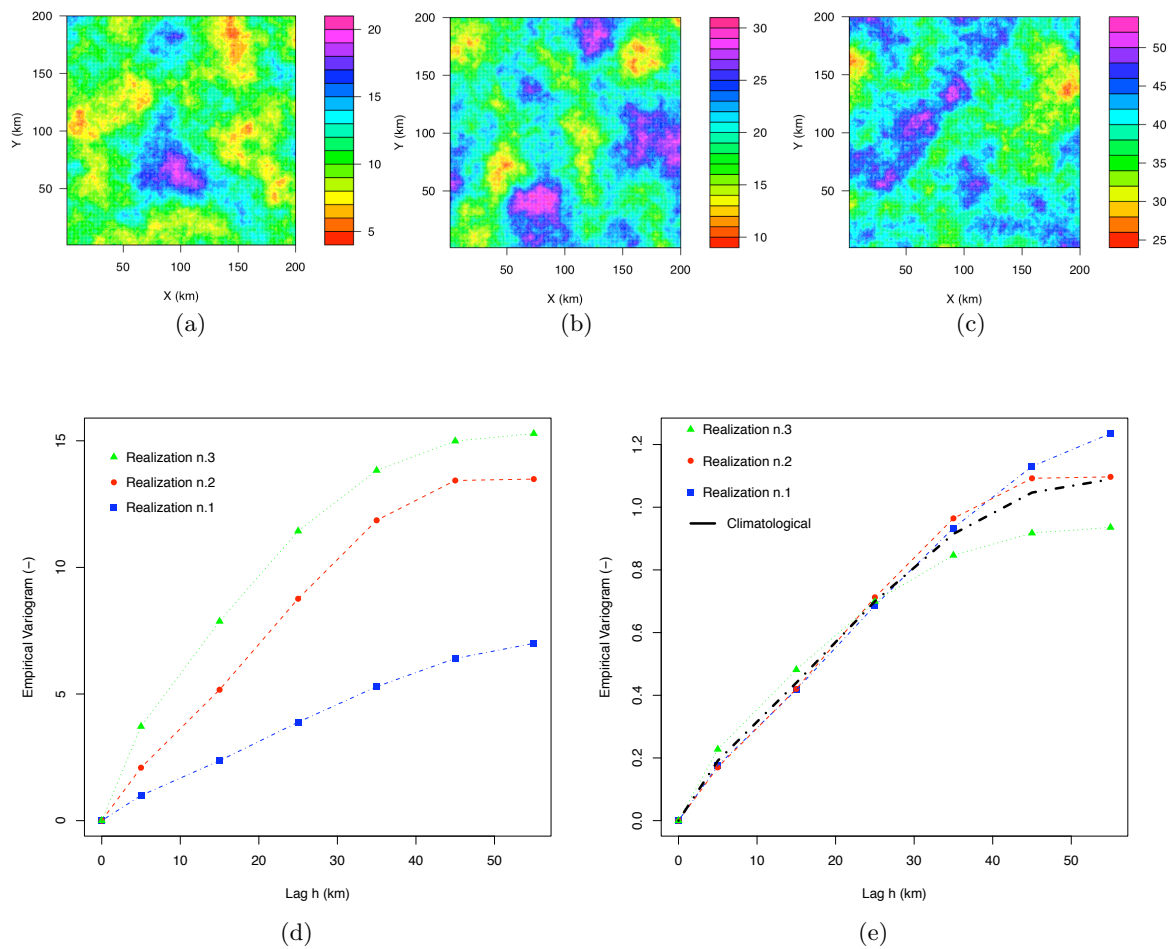


Figure 4.7: Contour-plot of three random fields characterized by the same spatial structure but different intensity range (a,b,c). d: Empirical variogram of the three fields. e: Empirical variogram of the fields normalized by their standard deviation and climatological variogram (black line).

variograms to variogram models. These purely statistical models can be used to replace the physical models when the latter are too complex to be efficiently represented (very often in atmospheric sciences). In addition, the use of variogram models is mandatory for the geostatistical interpolation, being the main input of the kriging interpolation.

Many variogram models are proposed in literature; *Chiles and Delfiner* (1999) report the most used. Generally, variogram models can be chosen according to two criteria: i) physical significance; ii) ease of use and parameter readability. In Figure 4.8 the 4 models we schematically plot the 4 models detailed in the following.

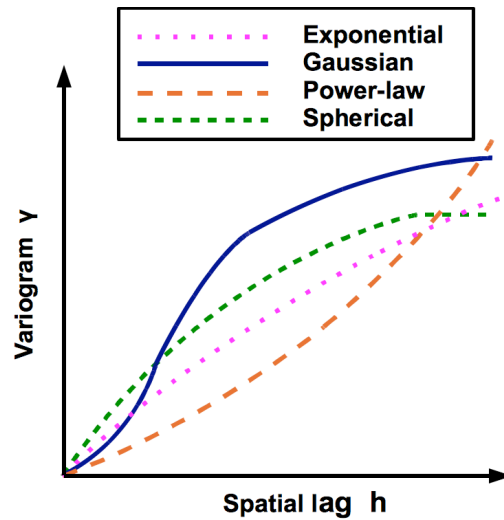


Figure 4.8: Schematic example of exponential, gaussian, power-law and spherical variogram model.

The exponential model issues from an analytical derivation of the covariance function of a continuous-time Markov processes possessing the property of conditional independence between the past and the future when the present is known.

$$\gamma(h) = n + s(1 - (\exp(-h/a))) \quad (4.11)$$

where a expresses the lag for which the variogram reaches the 0.63 of the sill. The model has range $r \rightarrow \infty$, and at about $3a$ we can define the practical range, as the lag for which the 95% of the sill is reached (*Chiles and Delfiner*, 1999).

Two other physically-based models merit to be mentioned. the Gaussian model

$$\gamma(h) = n + s(1 - (\exp(-h^2/a^2))) \quad (4.12)$$

which is associated with an infinitely differentiable stationary random function, and the power-law model

$$\gamma(h) = n + s(x^{\alpha_p}) \quad (4.13)$$

which is of particular interest since it satisfies the property of self-similarity. For $\alpha_p = 1$ the linear variogram is obtained.

Although the spherical variogram model is not the result of a particular kind of process, many natural phenomena show empirical evidence of this correlation structure. It is one of the most used variogram models, mainly because of the fact that the three parameters are directly readable from the sample variogram plot. The spherical variogram can be expressed as:

$$\begin{cases} \gamma(h) = n + s(1 - (1 - \frac{3r}{2h} + \frac{r^3}{2h^3})) & \text{for } h \leq r \\ \gamma(h) = n + s & \text{for } h \geq r \end{cases} \quad (4.14)$$

where h is the lag, and r is the lag for which the variogram reaches its asymptotic value, otherwise called range. The range r corresponds to the decorrelation distance of the process. The parameter n is the so-called nugget (its properties and practical interest have been discussed in Section 4.8), corresponding to $\gamma(0)$ and s is the sill, the difference between the variogram for $h > a$ (the average variance of the field) and the nugget n . Figure 4.9 shows how the three parameters can be visually detected.

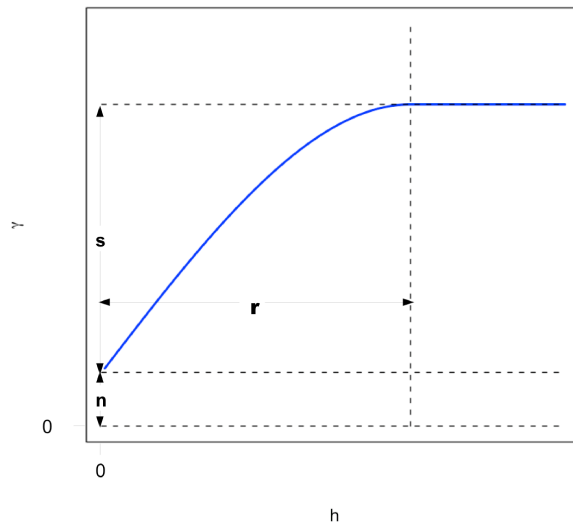


Figure 4.9: Spherical variogram model and graphical parameters determination. In the plot, the nugget n , the sill s and the range r of the variogram, can be graphically determined.

The nugget n indicates that the observations for lags $h \rightarrow 0$ are not completely correlated. This behavior has two possible origins: i) the process has a small-scale variability not caught by the measurement network due to its sparseness; ii) the point measurements are affected by uncertainty. In practice, rarely the nugget is found equal to 0. The practical implication of $n \neq 0$ for the interpolation process is that the interpolation function will not pass exactly through the observed points.

4.5 Interpolation of point data

In this section the most known methods for the interpolation of spatial data are briefly presented, focusing on their advantages and drawbacks.

The interpolation methods can be divided in two categories (*Arnaud and Emery, 2000*):

- deterministic methods (barycentric methods, space-partitioning methods and splines)
- stochastic methods (regression, local regression and kriging).

The barycentric methods infer the value at unknown points based on a weighted sum of the value at close points. The most known example is the inverse distance interpolation for which the estimation at a point s_0 takes the form:

$$\hat{z}(s_0) = \sum_{i \in V(s_0)} \frac{|s_i - s_0|^{-d}}{\sum_{i \in V(s_0)} |s_i - s_0|^{-d}} z(s_i) \quad (4.15)$$

where $d > 0$ is the inverse distance power. For $d = 2$ the weight associated to $z(s_i)$ decreases with the square of the distance from s_0 .

The space-partitioning methods infer the value at the unknown point as the value at a point supposed to have similar features. The nearest neighbour interpolation assigns to the unknown point the value of the closest known observation. The Thiessen polygons determine, for each known point, an area of influence; each point is considered the best estimator for its area of influence.

The splines are expressions of the intrinsic tendency of nature to minimize the energy of a system. The idea behind splines is to assimilate the field to the behavior of a thin metal plate forced to pass through a series of control points. In one dimensional fields, the cubic spline consists in fitting a 3rd order polynomial in each interval between two points, with the following constraints: i) the function must pass for the two points ii) the function must be continuous and derivable iii) the first derivative must be continuous iv) at the boundaries, the second derivative has to be null. The 2D case is the generalization of the 1D cubic spline, in which the number of parameters needed to interpolate a set of K parameters is $2(K + 3)$ (*Hastie and Tibshirani, 1990*). The main drawback of splines is that they have very poor skills in the inference of the field properties outside the boundaries: since the null second derivative is the only boundary condition, the field can exhibit drifts and the values could be affected by the bending in proximity to the last point of measure (in signal processing, this effect is referred to as “Gibbs phenomenon”).

4.5.1 Kriging

Kriging (*Krige, 1951; Matheron, 1962; Cressie, 1993; Chiles and Delfiner, 1999*) is an unbiased geostatistical interpolation method. It is based on the statistical modeling of the spatial correlation structure of a field.

The main difference between kriging and the other interpolation methods is that kriging is based on a statistical model of a process rather than on an interpolation function (*Chiles and Delfiner, 1999*). Similarly to the Inverse Distance Weighting, it consists in inferring the value of a random variable at a point based on a weighted sum of neighbours. Kriging is the best linear interpolator and it is unbiased: at each point, the estimation is made by means of an error minimization. By construction, kriging is designed to yield the error committed by the interpolation at each point.

Suppose to have a multi-field realization of a random field $Z(x, \omega)$ in space-time (Figure 4.10). We know the field values at some points x_1, \dots, x_N and we want to estimate, based on the information collected at all the realizations ω , the values at the point x_o for which we do not have estimates.

The weights λ_i associated with the point x_o are chosen with respect the following conditions:

- the estimator is linear: $\hat{Z}(x_o, \omega) = \sum_{i=1}^N \lambda_i(x_o)Z(x_i, \omega)$;
- the estimator must be, in an expected value sense, the best estimator;
- the estimator is not biased: $\hat{Z}(x_o, \omega)$ must correspond (in average) to $Z(x_o, \omega)$. For each realization, the expected value of the error $E[\epsilon(x_o, \omega)]$ is null. This corresponds to: $E[Z(x_o, \omega)] = E[\hat{Z}(x_o, \omega)] = E[\sum_{i=1}^N \lambda_i(x_o)Z(x_i, \omega)]$;
- The error variance

$$\sigma_{\epsilon(x_o, \omega)}^2 = E[(Z(x_o, \omega) - \hat{Z}(x_o, \omega))^2] \quad (4.16)$$

must be minimized.

Mainly, three types of kriging are available, depending on the properties of the mean of the process:

- Simple kriging: the mean of the process $\mu(s) = m$ is a known constant;
- Ordinary kriging: the mean of the process $\mu(s) = \mu$ is an unknown constant;
- Universal kriging: the mean of the process depends on the position s through a linear combination $\mu = \sum_{j=0}^p f_j(s)\beta_j$.

In the next sections, some additional details for each of these methods are provided.

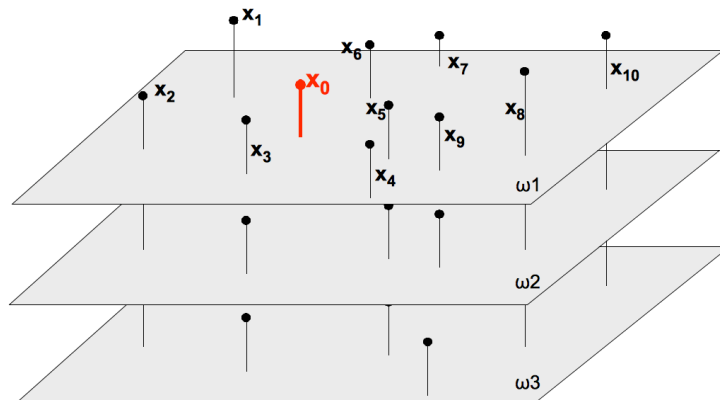


Figure 4.10: Example of 3 realizations of the random process $Z(x, \omega)$.

4.5.2 Simple kriging

The simplest kriging application is when the field is stationary (i.e. the mean does not depend on the location) and the mean is known. Since the mean is constant and known, simple kriging (*Matheron, 1970*) can be equally expressed in terms of covariance or variogram. Since the covariance function can be evaluated only when the mean of the process is known (and stationary), simple kriging is rarely suitable for practical problems.

In simple kriging the mean is known. It is possible to subtract the mean to the field in order to obtain, at each point, a distribution with null mean: $\mu(x) = 0$.

In this case, the non bias condition is trivial since $E[Z(x_0, \omega)]$ and $E[Z(x_i, \omega)]$ are equal to zero.

Let us deal with the minimization of the error variance. The error variance can be written $\sigma_{\epsilon(x_0, \omega)}^2 = \text{Var}[\epsilon(x_0, \omega)] = E[(Z(x_0, \omega) - \hat{Z}(x_0, \omega))^2]$; developing the right hand side of the equation we obtain:

$$\sigma_{\epsilon(x_0, \omega)}^2 = E[Z(x_0, \omega)^2 - 2 \sum_{i=1}^N \lambda_i(x_0) Z(x_i, \omega) Z(x_0, \omega) + \sum_{i=1}^N \sum_{j=1}^N \lambda_i \lambda_j Z(x_i, \omega) Z(x_j, \omega)] \quad (4.17)$$

where the first member is equal to $E[Z(x_0)]^2$ that is the variance of the process; the remaining terms can be expressed in terms of the covariance:

$$\sigma_{\epsilon(x_0, \omega)}^2 = \text{Var}[Z(x_0, \omega)] - 2 \sum_{i=1}^N \lambda_i(x_0) \text{Cov}[Z(x_i), Z(x_0)] + \sum_{i=1}^N \sum_{j=1}^N \lambda_i \lambda_j \text{Cov}[Z(x_i), Z(x_j)] \quad (4.18)$$

The minimization of the error variance corresponds to impose

$$\left\{ \begin{array}{l} \frac{\partial \sigma_{\epsilon(x_0)}^2}{\partial \lambda^k} = 0 \end{array} \right. \quad (4.19)$$

It is demonstrated (*Obled, 2007*) that the development of this partial derivative system leads to the following equation:

$$\left\{ \begin{array}{l} \frac{\partial \sigma_{\epsilon(x_0)}^2}{\partial \lambda^k} = -2 \text{Cov}[Z(x_k) Z(x_0)] + 2 \sum_{j=1}^N \lambda_j \text{Cov}[Z(x_k) Z(x_j)] = 0 \end{array} \right. \quad (4.20)$$

For each k we can thus write the kriging system in matrix form:

$$\begin{pmatrix} C_{x_1, x_1} & C_{x_1, x_2} & \cdots & C_{x_1, x_N} \\ \vdots & \vdots & \ddots & \vdots \\ C_{x_k, x_1} & C_{x_k, x_2} & \cdots & C_{x_k, x_N} \\ \vdots & \vdots & \ddots & \vdots \\ C_{x_N, x_1} & C_{x_N, x_2} & \cdots & C_{x_N, x_N} \end{pmatrix} \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_k \\ \vdots \\ \lambda_N \end{pmatrix} = \begin{pmatrix} C_{x_1, x_0} \\ \vdots \\ C_{x_k, x_0} \\ \vdots \\ C_{x_N, x_0} \end{pmatrix} \quad (4.21)$$

This example was the simplest since the mean is known. In ordinary kriging, the optimization is conditioned to the fact that the weights sum is unit and therefore requires the use of Lagrange multipliers, with increased complexity.

4.5.3 Ordinary kriging

With a small increase in complexity, the kriging interpolation can be generalized for cases where the mean is stationary but unknown. This interpolation is called “ordinary kriging” (Matheron, 1970). Instead of the covariance function, the variogram function (Section 4.3) expresses the spatial correlation as a function of the distance lag, h . In this case, the random variable $Z(x)$ is required to be stationary of second order.

The non-bias condition is expressed by fixing the weight sum to the unit: $\sum_{i=1}^N \lambda_i(x_0) = 1$. The error variance

$$\sigma_{\epsilon(x_0, \omega)}^2 = -\gamma(x_0, x_0) - \sum_{i=1}^N \sum_{j=1}^N \lambda_i \lambda_j \gamma(x_i, x_j) + 2 \sum_{i=1}^N \gamma(x_i, x_0) \quad (4.22)$$

must be minimized. In this case, the mean is unknown and therefore we can not directly express the covariance term. We add to the system the condition $\sum_{i=1}^N \lambda_i - 1 = 0$. The system to be minimized is then subject to a constraint and the method of Lagrange multipliers must be used. This method is a necessary condition for the optimization in constrained problems, and allows for finding the maxima and minima of the constrained function.

The system becomes

$$\begin{cases} \frac{\partial[\sigma_{\epsilon(x_0)} - 2\nu(\sum_{i=1}^N \lambda_i - 1)]}{\partial \lambda_k} = 0 \\ \frac{\partial[\sigma_{\epsilon(x_0)} - 2\nu(\sum_{i=1}^N \lambda_i - 1)]}{\partial \nu} = 0 \end{cases} \quad (4.23)$$

that, in matrix form, can be expressed as

$$\begin{pmatrix} \gamma_{x_1, x_1} & \cdots & \gamma_{x_1, x_N} & 1 \\ \vdots & \ddots & \vdots & \vdots \\ \gamma_{x_N, x_1} & \cdots & \gamma_{x_N, x_N} & 1 \\ 1 & 1 & \cdots & 0 \end{pmatrix} \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_N \\ \nu \end{pmatrix} = \begin{pmatrix} \gamma_{x_1, x_0} \\ \vdots \\ \gamma_{x_N, x_0} \\ 1 \end{pmatrix} \quad (4.24)$$

An example of the kriging interpolation method is shown in Figure 4.11. We prefer to show how kriging works in a mono-dimensional case. Suppose to have a series of daily rainfall observations, along a line, where stations are separated by 5 km. We aim to interpolate the points in order to obtain a continuous surface.

The first operation is to evaluate the variogram of the daily rainfall amounts for each couple and to average the obtained values for class of distances. Then, a variogram model is fitted to the sample variogram. In Figure 4.11-a, a Gaussian variogram seems to be adapted to describe the spatial structure of the field. An automatic fitting procedure has been applied. Kriging is then applied to estimate the value of the field at intermediate points (Figure 4.11-b). In this case, a point every 2 km is estimated. The adopted variogram model has nugget $n \neq 0$. It means that either the measures

are affected by a systematic error or the process exhibits small-scale variability that a coarse network is not capable to detect. We now check the influence of a correct nugget estimation. Figure 4.11-b shows the interpolated rainfall intensity for two kriging schemes. The blue curve corresponds to a kriging without nugget while a positive nugget is introduced in the kriging scheme leading to the red curve. Imposing a null nugget, the interpolated intensity exactly matches the known data points but displays a rough curve when close data points have close intensities. When kriging includes a nugget, the curve of the interpolated intensity is smoother but does not exactly match the data.

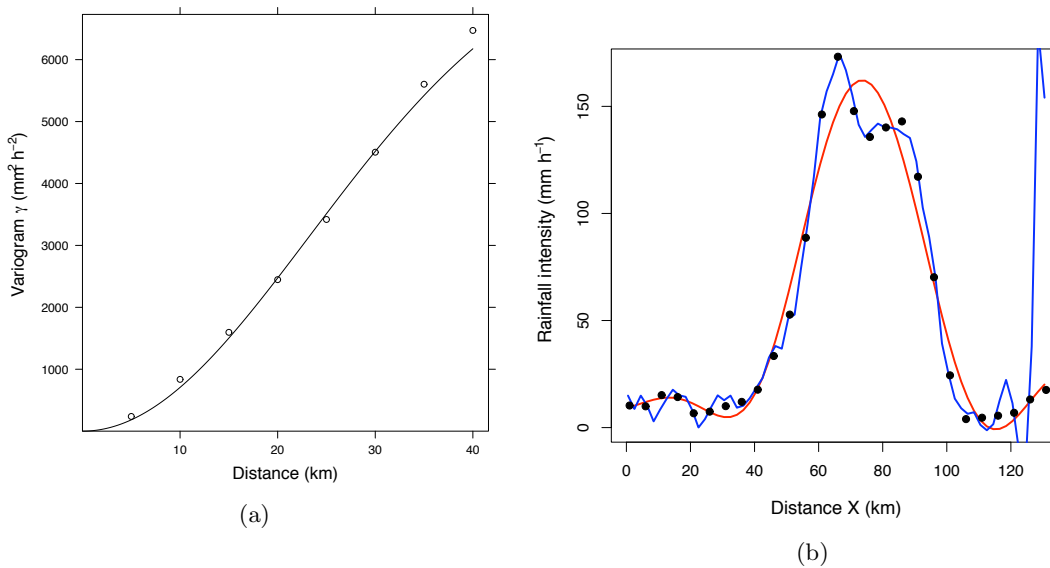


Figure 4.11: a: Sample variogram for the example rainfall field. b: Observed rainfall field (an observation every 5 km), Gaussian variogram model (red line), Gaussian model with imposed zero nugget (blue line).

In Appendix D a code for the 1D kriging is proposed.

4.5.4 Universal kriging

The variogram formulation has been modified for dealing with particular situations where the mean of the field changes as a function of the spatial coordinates. When the stationarity of the mean (upon which ordinary and simple kriging are based) is not verified, the “Universal Kriging” method, also known as kriging with an external drift, can be used. Kriging with an external drift (*Goovaerts, 1997; Wackernagel, 1998*) consists in associating to the primary variable Z , known at few locations, an auxiliary variable Y , available everywhere in the spatial domain. For the application of the method the second-order stationarity must be verified. The use of universal kriging is of particular interest in mining engineering or in hydrogeology. In this latter field, for example, it is common to have a drift in the piezometric height, and the sample variogram show the behavior illustrated in Figure 4.5. In atmospheric sciences, the use of universal kriging is limited to low-variability fields, such as monthly or annual rainfall. In high-variability fields (such as instantaneous rainfall), there is high probability of introducing an erroneous drift that can lead to anomalous results.

An example of kriging with an external drift is the use of the elevation as an auxiliary variable for mapping the annual rainfall in the Cévennes-Vivarais region 4.12. The assumption is that the rainfall

depth at the monthly or annual accumulation time is locally correlated with the elevation (*Gottardi, 2009*). Even if an overall trend rainfall depth versus elevation can not be established in the region, the method uses the local information to detect local linear trends of the rainfall depth with the elevation. *Goovaerts (2000)* showed that kriging with an external drift is the best method for taking into account the local information provided by elevation.

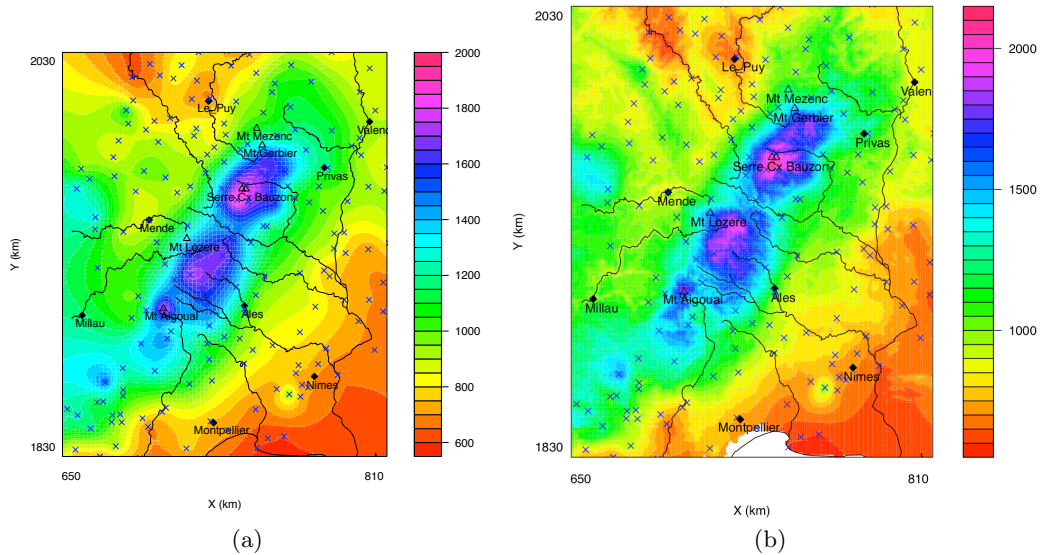


Figure 4.12: a: Annual rainfall-depth map (in mm) using ordinary kriging. b: Annual rainfall-depth map using elevation as external drift.

4.6 Conclusion

This section has been devoted to a basic description of the structural analysis and spatial interpolation methods known as geostatistics. In the first paragraph a short description of the main properties of random fields has been reported. We discussed about stationarity, intrinsic hypothesis and ergodic hypothesis. These properties of the fields will intervene in the choice of the structural tool needed for their characterization and of the best interpolation method for the estimation of the field value at ungauged locations.

We introduced the use of variograms as a descriptor of spatial fields. Variogram is more flexible than covariance to describe the properties of random fields, in the sense that variogram just need stationarity of the mean of the field (while for computing covariance the mean must be known). The use of variograms can also be extended to fields that do not obey to the second order stationarity, i.e. fields with a linear drift of the mean. Detrended variograms can be used for these purposes. Indicator variograms have been proposed for deal with fields presenting different correlation structures as a function of the intensity range (e.g. rain-no rain intermittency and rainfall variability). The need of atmospheric sciences to detect a multi-field behavior of a random variable and the impossibility to draw significant information from a single storm justify the introduction of climatological variograms.

The empirical variograms evaluated from data can be suitably fitted with variogram models expressing with a continuous function the correlation structure of the fields. The variogram model can be seen as a purely statistical modeling of the random variable at different scales. In addition, it is the basic ingredient for the kriging interpolation, a stochastic interpolation method widely used for inferring the value of a random variable at unaged locations. The advantages of kriging are that i) it is the best linear interpolator and ii) it provides the uncertainty of the estimation at each point. Similarly to variograms, different kriging techniques exist for taking into account of the properties of the analyzed fields: known stationary mean, unknown stationary mean, linear drift of the mean, presence of an auxiliary variable or of a co-variable.

Scale invariance and self-similarity



Résumé

La définition “**analyse multi-échelle**” peut faire référence à deux analyses différentes. Généralement ce terme indique que l’étude est conduite de façon indépendante pour plusieurs échelles d’intérêt. Dans notre cas, l’analyse multi-échelle consiste à définir les relations entre les échelles, avec l’objectif de déterminer les lois qui expriment la “**relation d’échelle**”. Un grand nombre de processus naturels possède des propriétés invariantes-d’échelle, notamment leur propriétés statistiques sont auto-similaires à toutes échelles.

L’invariance d’échelle à été détectée dans certains objets géométriques ainsi que dans une multitude de processus naturels, de la géophysique à l’économie, de la biologie à la géographie. La complexité et l’absence d’une formulation mathématique rigoureuse ont limité la diffusion de l’analyse de l’invariance d’échelle, malgré les nombreuses évidences physiques de son existence.

Le chapitre est consacré à la description de l’état de l’art sur l’étude de l’invariance d’échelle de la pluie. Dans une première partie, nous introduisons la théorie de l’invariance d’échelle. On présente en suite les principales approches pour analyser l’invariance d’échelle des processus géophysiques et le cadre mathématique qui amène aux différentes formulations.

En conclusion, nous présentons une description des principaux travaux dans la modélisation de l’invariance d’échelle dans trois contextes différentes: **invariance d’échelle spatiale**, **invariance d’échelle temporelle**, **invariance d’échelle dans l’espace espace-temps**.

5.1 Introduction

The definition “**multi-scale analysis**” may refer to two different analyses. Generally this term signifies that the studies are conducted, separately, at more than one scale of interest. In our case, the term refers to the studies that define the relationships between the scales, with the aim to define the law governing the scale transition (the so-called “**scaling relation**”). Many processes exhibit scale-invariant properties, i.e. their statistical properties are similar whatever the scale.

Scale-invariance has been detected in the geometrical properties of abstract entities as well as in the statistical properties of a multitude of processes, from geophysics to economics, from biology to



Figure 5.1: Construction of the first six iterations of a cantor set. The segments have been depicted as bars of fixed width.

geography. The general complexity and the lack of a rigorous mathematical formulation has limited the acceptance and the massive usage of scale-invariance, despite the physical evidences.

This chapter is devoted to the description of the state of the art regarding the scale-invariance of rainfall. In the first part, a brief introduction on the scale-invariant theory is given. The two main approaches to verify the scale-invariance of geophysical processes are reviewed, with a description of the various types of scaling that have been found in natural processes. We give some details on the mathematical framework leading to the different formulations.

In conclusion, we present a description of the main works in the modeling of scale invariance in three different context: **spatial scaling**, **temporal scaling**, **space-time scaling**.

5.2 Scaling in geometry: fractals

According to the definition of *Mandelbrot* (1967), a fractal is "a rough or fragmented geometric shape that can be split into parts, each of which is (at least approximately) a reduced-size copy of the whole".

Many decades before the definition of the "fractal" concept, objects with unusual properties (namely the fractional dimension) were studied by Cantor, Koch, Peano, Sierpinski and Levy (*Mandelbrot*, 1982). The Cantor ternary set (1894, Figure 5.1) could be a good example to understand these unusual properties: it is created by repeatedly deleting the open middle thirds of a set of line segments. If the segment line is defined in the interval $(0, 1)$, the first step is the deletion of the middle third $(1/3, 2/3)$. The second step is the removal of the middle third of each of the two remaining segments, $(1/9, 2/9)$ and $(7/9, 8/9)$. The process is iterated for an infinite number of steps k .

It is well known that points have geometrical dimension 0, lines have unit geometrical dimension, and plane have geometrical dimension 2. Fractals are characterized by the fact to present fractional dimension: the Cantor set, for example, is generated by fragmentation of a segment, of dimension $D = 1$, that in turn is composed by an infinite number of points ($D=0$, for this reason the Cantor set is also referred to as "Cantor dust").

The fractal dimension of the set is defined as:

$$D = \lim_{\epsilon \rightarrow 0} \frac{\log N(\epsilon)}{\log \frac{1}{\epsilon}} \quad (5.1)$$

where $N(\epsilon)$ is the number of self-similar structures of size ϵ needed to cover the whole structure. In the case of Cantor dust,

$$D = \lim_{k \rightarrow +\infty} \frac{\log 2^k}{\log \frac{1}{1/3^k}} = \lim_{k \rightarrow +\infty} \frac{\log 2^k}{\log 3^k} = 0.6309... \quad (5.2)$$

where k is the step number.

The fractal dimension is an important property of geometrical objects but it does not univocally define a geometrical object; sets with the same fractal dimension but different appearance can be built.

5.3 Fractals in nature

Since the second half of XX century, numerous natural and anthropogenic objects have been discovered to have fractal properties. Starting from the second half of XX century, analyzing economic as well as natural observations, unexpected properties of real series related to “fractality” have been discovered.

One of the first findings concerns the measure of the objects length. “How long is the Coast of Britain”, contained in *Mandelbrot* (1967) clearly expresses the concept of fractal objects. Let us imagine to measure the length of a coastline drawn on a 2D surface, between two limits. Its length is at least equal to the distance measured along the straight line connecting its beginning and its end. If we take supports of smaller size, it is possible to notice that the coastline length increases with the dimension of the support. The approximate length of objects with the dimension of the support varies following a straight line in log-log plot.

As an application, let us determine the length of a river included in the study region: the Allier river in the Auvergne department shown in Figure 5.2-d. With rulers of decreasing length, we measure the length of the river on a map (Figure 5.2-a,b,c). The ruler length corresponds to the values of 32, 16, 8, 4, 2 and 1 km. The measures are reported in Figure 5.2-e. It is worthy to notice not only that the length of the river increases with the inverse of the ruler length, but also that the relation is approximately linear in double logarithmic plots. To highlight the log-log linear relationship, the regression line ($R^2 = 0.94$) is drawn. The log-log slope of the regression line in Figure 5.2-e represents the Hurst exponent H (also known as the fractal exponent). Therefore, the genetic processes of the Allier riverbed acted building a seemingly fractal object.

From this example it is clear that the properties of a variable change with respect to the resolution of analysis and that some of these properties change in a regular and predictable way. *Mandelbrot* (1982) performed the same analysis for a number of geographical variables, finding different Hurst exponents (Figure 5.3). In the next sections we will move from geographical to geophysical variables. For most of the geophysical variables we will see that a single exponent (the fractal dimension) is not enough to describe the scaling features of a variable and a continuous spectrum of exponents should be adopted. The main scope of the scaling analysis is to determine this spectrum of exponents (also referred to as “singularity spectrum”) of a geophysical variable.

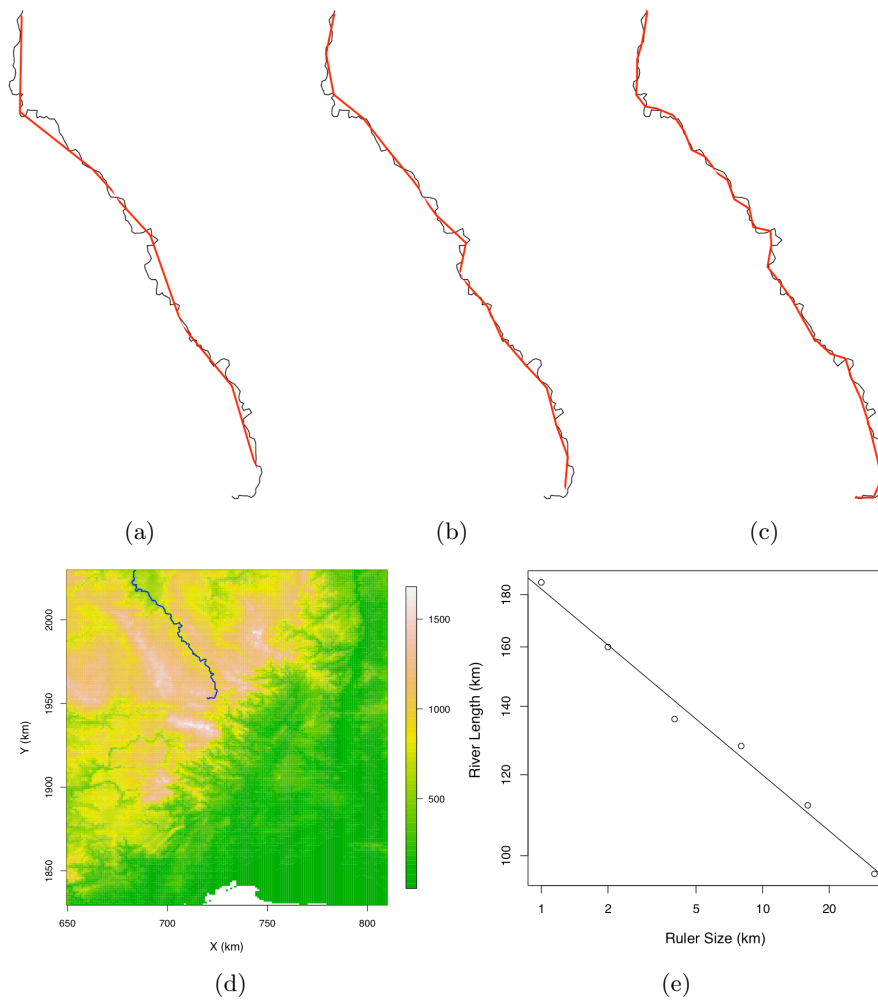


Figure 5.2: Allier River, effect of the ruler size on the measure of the river length. On panel a,b,c and d the gray line shows the Allier river course as given by 1 km digital terrain model. The red segments are juxtaposed rulers of lengths: a: 16 km; b: 8 km; c: 4 km. d: Localisation of the Allier River within the study region. e: Log-Log Plot of the estimated river length as a function of the ruler scale and regression line.

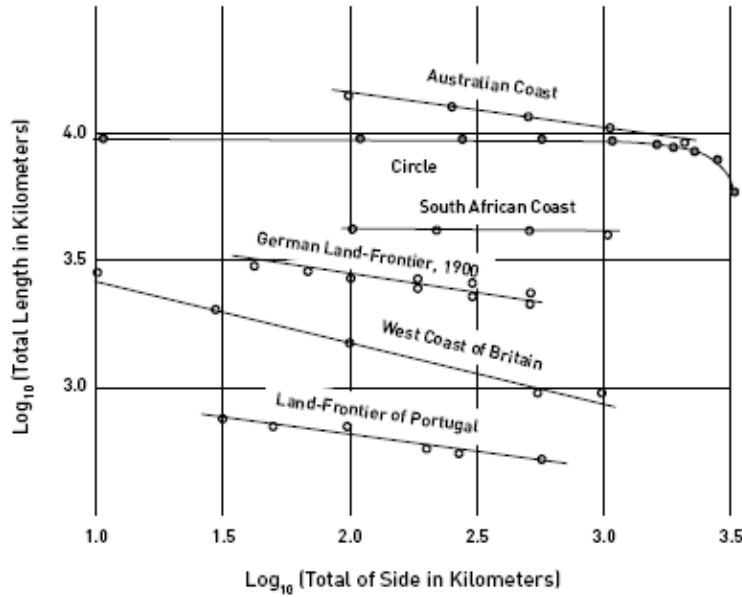


Figure 5.3: Fractal analysis on the coastline length in several regions of the world, from *Mandelbrot* (1982). The slope in log-log axes gives the fractal dimension of the set.

5.4 The origin of scaling in nature: turbulence

The first evidences of scaling in natural processes are relative to the study of turbulence. In turbulence, the most of kinetic energy is contained in the large scale structures and the energy flux cascades down from the large scale structures to smaller ones, creating smaller and smaller structures which produce a hierarchy of eddies. By means of a dimensional analysis, *Kolmogorov* (1941) demonstrated that the energy spectrum function can be defined as

$$E(k) = C\epsilon_d^{2/3}k^{5/3} \quad (5.3)$$

where k is the wave number, ϵ_d the energy dissipation rate, C is a constant.

The power-law function expressed in Equation 5.3 implies that all the scales equally contribute to define the variability of the eddies; no characteristic scale is present and turbulence is statistically self-similar at different scales. Many experiments (*Frisch*, 1995) demonstrated the validity of Equation 5.3.

We can depict the contribution of each scale to the variability of the process through the energy spectrum. The energy spectrum describes how the energy of a signal $f(t)$ is distributed with the frequency ω . For a continuous signal, we can define the spectral density $\Phi(\omega)$ as:

$$\Phi(\omega) = \left| \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t)e^{-i\omega t} dt \right|^2 = \frac{F(\omega)F^*(\omega)}{2\pi} \quad (5.4)$$

where $F(\omega)$ is the continuous Fourier transform and $F^*(\omega)$ is its complex conjugate¹.

¹In cases when the function to study is not square integrable (i.e. $\int_{-\infty}^{+\infty} |f(x)|^2 dx \leq +\infty$), the Fourier transform can not be computed. Thanks to the Wiener-Khinchine theorem it is possible to compute the spectral density of a

The energy spectrum can also be interpreted as the distribution of the variance at different scales, and therefore indicates the contribution of each scale to the variability of the process. A flat spectrum means that the variability is not distributed across the spectrum, but concentrated at one scale: an example is given by the white noise (Figure 5.4-a). A signal with no characteristic scale is linear as a function of the frequency in double logarithmic plot (like in Figure 5.4-b, for frequencies corresponding to the range 2 hours-1 week).

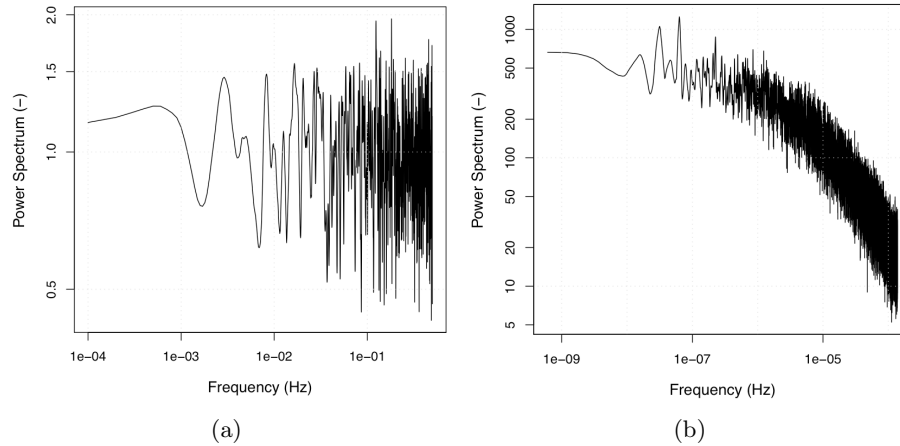


Figure 5.4: a: Power spectrum of a white noise. b: Power spectrum of the rainfall series of Montpellier (time resolution=1h). In both cases, a smoothing has been performed in order to reduce the variability of the spectrum.

As reported in *Schertzer and Lovejoy (1987)*, the initial purpose of the scale-invariance studies was the improvement of the numerical modeling of atmospheric processes. In particular, they argue that the model sub-grid parametrizations could be different from one scale to another, making questionable the physical coherence of the variables. In contrast, many fundamental equations are invariant under scale transformations. The Navier-Stokes equations, for instance (*Frisch and Parisi, 1985*):

$$\begin{cases} \partial_t u + u \nabla u = -\nabla p + \nu \nabla^2 u \\ \nabla v = 0 + w \end{cases} \quad (5.6)$$

are formally invariant under the group of transformations

$$r \rightarrow \lambda r, \quad v \rightarrow \lambda^h v, \quad t \rightarrow \lambda^{1-h} t, \quad \text{for } \lambda > 0 \quad (5.7)$$

where w is a constant including boundary and initial conditions, ν is the fluid viscosity, p the pressure and u the turbulent velocity.

signal. The theorem states that, if and only if the signal is a wide-sense stationary process (all the moments are finite and stationary), the spectral density can be computed by Fourier transform of the Auto-Correlation Function (ACF, Equation 3.6):

$$\Phi(\omega) = \int_{-\infty}^{+\infty} ACF(\tau) \exp^{-i2\pi\omega\tau} d\tau \quad (5.5)$$

Similar scale-invariant relations could be used to improve the description of numerous physical phenomena, in the range of validity of scale-invariance. At least two methods for the assessment of the scaling have been proposed, providing complementary information. They are described in the next section.

5.5 Generalization of the fractal concept: Multifractals

The simple scaling hypothesis is the result of additive processes and it is rarely verified in nature. Strong nonlinearities have been detected in natural phenomena; these deviations were predicted by *Frisch and Parisi* (1985) and stem from multiplicative processes (*Schertzer and Lovejoy*, 1985). In simple scaling, a single scaling exponent suffices to describe the behavior of the statistical moments at different scales. In contrast, multiple scaling requires multiple exponents (e.g. mean and variance of the process scale differently) and it is therefore more general. In terms of probability distribution, we typically observe that the fractal dimension decreases as the threshold is increased. Such processes are referred to as “multifractals” (*Frisch and Parisi*, 1985).

In the previous sections, we have seen that the scale invariance of a random variable can be qualitatively detected through the spectral density analysis. If the spectral density is linear in double logarithmic plots, the process has not characteristic scale. The absence of a characteristic scale is the first requirement of scale-invariance.

However, the power density slope can not provide sufficient information about the kind of scale-invariance. Two concurrent methods have been developed to quantitatively describe the scale-invariance of a process: the Generalized Structure Function $\zeta(q)$ (*Frisch and Parisi*, 1985; *Harris et al.*, 2001) and the Moment Scaling Function $\tau(q)$ (or a related function $K(q)$) (*Schertzer and Lovejoy*, 1987; *Gupta and Waymire*, 1990). The two approaches are related but have significant differences: $\zeta(q)$ is computed by evaluating differences of the field at increasing distances; $\tau(q)$ deals with the subsequent averaging of the field over scales λ .

5.5.1 Scaling of the Generalized Structure Function ζ

The self-similarity concept is strictly related to the findings of *Kolmogorov* (1941) regarding the power-law relation in Equation 5.3. In their pioneering work regarding the fully developed turbulence, *Frisch and Parisi* (1985) have proposed a description of the longitudinal velocity signal $u(x)$. They found that the local scaling behavior of the velocity increment $\delta u(x_0, l)$ around x_0 , for $l \rightarrow 0$, can be fully characterized by the expression

$$\delta u(x_0 + l) = [u(x_0 + l) - u(x_0)] \sim l^{h(x_0)} \quad (5.8)$$

where $h(x_0)$ is called local singularity exponent.

The singularity spectrum $D(h)$ of the field is then defined as the function that gives, for a fixed h , the fractal dimension of the set of points x for which the exponent $h(x)$ is equal to h . *Frisch and Parisi* (1985) proposed an experimental method for estimating the $D(h)$ spectrum based on the

scaling exponent $\zeta(p)$ derived from the computation of the expected value of the p^{th} power of the turbulent velocity u fluctuations as a function of the displacement l (i.e. the scale):

$$E[|u(x+l) - u(x)|^p] \sim l^{\zeta(p)} \quad (5.9)$$

$\zeta(p)$ is the scaling function of the Generalized Structure; when this power-law relationship is verified the field is said to be self-similar. $\zeta(1)$ corresponds to the Hurst exponent.

The relation between $\zeta(p)$ and $D(h)$ is given by the Legendre transform

$$D(h) = \min_p(ph - \zeta(p) + 1) \quad (5.10)$$

Equation 5.9 has important consequences:

- the Generalized Structure Function $E[|u(x+l)u(x)|^p]$ has a power-law relationship with the separation distance l , expressing the scale-invariance;
- the existence of the $\zeta(p)$ function, non linear with l , models the deviation from simple scaling;
- the scale-invariance could be determined in non-stationary fields, but they must have stationary increments;

In few cases, the function $\zeta(p)$ is linear with p . This condition regards, for instance, Brownian motion.

5.5.2 Moment scaling analysis

The moment scaling analysis significantly differs from Generalized Structure Function analysis in that, instead of looking at the variability of the fine-scale field variability at different separation lags l , it models its ensemble moments as a function of the scale.

Simple scaling is the simplest kind of scaling. The mathematical description of the simple scaling is provided by *Gupta and Waymire* (1990): an arbitrary random field Y in \mathfrak{R}^d is simple scaling if, for each λ , there is a scale function C_λ such that the following equality holds for any arbitrary set of points $\mathbf{x}_1, \dots, \mathbf{x}_n$:

$$P[C_\lambda^{-1}Y_{\lambda_1}(\mathbf{x}_1) < y_1, \dots, C_\lambda^{-1}Y_{\lambda_1}(\mathbf{x}_n) < y_n] = P[Y_{\lambda_2}(\mathbf{x}_1) < y_1, \dots, Y_{\lambda_2}(\mathbf{x}_n) < y_n] \quad (5.11)$$

where λ is an index that is function of the ratio $\frac{\lambda_1}{\lambda_2}$.

We can express the same equality in probability distribution in the contracted form as:

$$C_\lambda^{-1}Y_{\lambda_1}(\mathbf{x}) \stackrel{d}{=} Y_{\lambda_2}(\mathbf{x}) \quad (5.12)$$

The only acceptable form of the scaling function C_λ is $C_\lambda = \lambda^c$ (*Gupta and Waymire*, 1990; *Sornette*, 2004).

Therefore, simple scaling of the distributions of the same random variable Y at two scales λ_1 and λ_2 can be expressed as:

$$Y_{\lambda_1} \stackrel{d}{=} \lambda^{-c} Y_{\lambda_2} \quad (5.13)$$

where c is a parameter referred to as co-dimension function and $\lambda = \frac{\lambda_1}{\lambda_2}$ is the scale ratio.

Gupta and Waymire (1990) demonstrate that Equation 5.13 involves a similar relation in terms of the ensemble moments; this property is known as “wide-sense simple scaling”:

$$E[Y_{\lambda_1}^q] = \lambda^{-\tau} E[Y_{\lambda_2}^q] \quad (5.14)$$

where q is the moment order and λ is the scale ratio; τ is called the “moment scaling function”. In simple scaling, it is linear with q . *Gupta and Waymire* (1990) specify that the transition from “strict” to “wide” scaling is verified only if the statistical moments of the distribution exist. The inverse relation is always true, i.e. “wide sense” implies “strict sense” scaling. The relation in Equation 5.14 is a power-law relationship: the scaling of ensemble moments occurs when the ensemble moments exhibit linearity in log-log axes as a function of the scale λ .

The parameters c and τ of the relations 5.13 and 5.14 are inter-dependent and define the type of scaling:

- when the function c in Equation 5.13 is a constant and τ in Equation 5.14 is linear with the moment order q , the relation is called “simple scaling”: this phenomenon is rarely observed in geophysics.
- more frequently c depends on the intensity and it is function of the “singularity strenght” γ_s ; the moment scaling function and the codimension c are related through the relation $\gamma_s(q) = \frac{dK(q)}{dq}$; the moment scaling function $\tau(q)$ is not linear with the moment order q . In this case, we have “multiple scaling”.

In Figure 5.5 we show the differences between simple and multiple scaling of the distributions. In simple scaling (Figure 5.5-a), the probability distribution undergoes a simple contraction/expansion, i.e. it is multiplied by a constant scaling factor; in multiple scaling the probability distribution shape may change because each quantile and moment scale in a different manner.

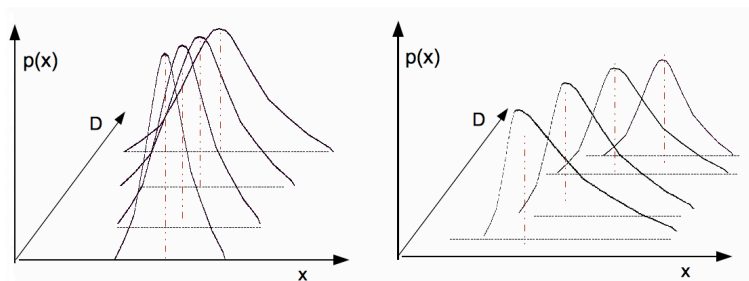


Figure 5.5: Scheme of simple (left) and multiple (right) scaling of probability distribution functions. In simple scaling, a single factor is needed to transform the pdf at different scales (shape conservation); in multiple scaling an infinite number of factors are needed, modifying the pdf shape.

5.5.2.1 Applicability of the moment scaling analysis: Statistical self-affinity

Schertzer and Lovejoy (1987) state that the use of moment-scaling analysis when the process is not conservative leads to spurious results. A process is conservative if, in the Fourier space, the power spectrum slope (in log-log) of a n -dimensional process is steeper than n . When the process is not conservative, the moment scaling may not be verified on the field itself and a differentiation is required to detect scale-invariance (*Schertzer and Lovejoy*, 1987). The differentiation can be of integer or fractional order. Processes that exhibit scale-invariance after fractional differentiation are called “self-affine”. Through the convolution of the original field Y with a specific filter (a log-log straight line in the Fourier space), it is possible to obtain a conservative field ϕ , apt for the moment scaling analysis.

If β is the spectral slope of the process Y , the fractional derivation order required to obtain a conserved process is (*Tessier et al.*, 1993):

$$H = \frac{\beta - 1 + K(2)}{2} \quad (5.15)$$

H is the Hurst exponent, also computed as $\zeta(q)$ in Equation 5.9 for $q = 1$. To fractionally derive a process, one may work in the frequency domain. If $F(\omega)$ is the Fourier transform of the function $f(x)$, the Fourier transform of the H^{th} order derivative, $f^{(H)}(x)$ is equal to $(i\omega)^H F(\omega)$. When H is negative, the transformation is a fractional integration, i.e. the inverse process required to return to the non-conservative process from the conservative one. An example of fractional differentiation, for the 30-min wind-speed station of Luzern, Switzerland, is shown in Figure 5.6. The power spectrum of the series is reported in Figure 5.6-a, showing a power-law slope of 1.18. The fractional derivation yields a field with unit slope (Figure 5.6-b); this field can be used to estimate the moment scaling function. We must point out that the choice of the arbitrary unit in the y axes does not affect the slope estimation, since log-log plots are used.

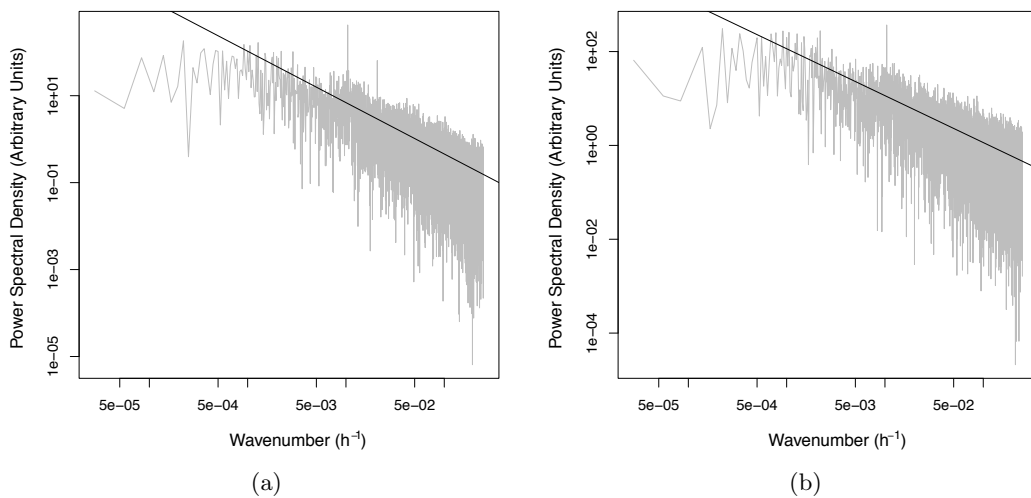


Figure 5.6: Spectral densities (obtained by Equation 5.4) of 30-min wind-speed series of Luzern, Switzerland. a: Spectral density plot of the original time series, $\beta = 1.18$. b: Spectral density of the fractionally differentiated flux, $\beta = 1$.

A simpler approach to construct a conservative field from a non-conservative one consists in taking the field fluctuations. In this case, attention must be devoted to the finer resolution scale of the analysis: due to sampling errors, deviations from scaling have been detected (*de Montera et al.*, 2008).

5.5.2.2 Moment scaling function of conservative processes

In this section, the analysed random variable Y is conservative. We detail the computation of its moment scaling function. The general scaling relation of ensemble moments known as Mandelbrot-Kahane-Peyriere (MKP, *Mandelbrot (1974); Kahane and Peyriere (1976)*) is defined as:

$$\tau(q) = 1 - q + \log_b E[Y^q] \quad (5.16)$$

where q is the moment order and b the ratio between two successive scales of analysis (in the simplest case, $b = 2$).

The main requirement is to have long enough series ($N \rightarrow +\infty$) to make the sample moments $\frac{1}{N} \sum_{i=1}^N I_\lambda^q$ converge to the ensemble moments $E[I^q]$ to which the scale-invariance applies. The scale-invariance is verified when the moments lie on a straight line in log-log axes as a function of the scale ratio λ .

Gupta and Waymire (1990) define a series of probability distributions that may fit with scale-invariance. *Gupta and Waymire (1993)* derived the expression of $\tau(q)$ for different kinds of cascades: for log-normal distributed fields (firstly introduced by *Kolmogorov (1962)* in the framework of the statistical theory of turbulence), we have:

$$\tau(q) = \frac{\sigma^2}{2 \log(b)} (q^2 - q) \quad (5.17)$$

With just one parameter (σ in Equation 5.17), one can model the infinite hierarchy of singularities proper of log-normal multi-scaling processes in a continuum of scales. *Gupta and Waymire (1993)* also demonstrate the possibility of using Levy-stable processes (Equation 5.55) to model scale-invariance in case of fat tailed distributions; Levy-stable distribution is used in the Universal Multifractal Model proposed by *Schertzer and Lovejoy (1992)*.

5.5.2.3 Universal Multifractal Model

Schertzer and Lovejoy (1992) have developed the model referred to as “Universal Multifractal model”, a generalization of the log-normal case in Equation 5.17 that accommodates exponential as well as heavy tailed distributions. They assume that the fields are originated by a multiplicative cascade where the weights are distributed following Levy’s stable distribution, the only distribution with the Gaussian to exhibit stability-upon-addition but, differently from the latter, characterized by hyperbolic tails. As Gaussian distribution is a particular case of Levy distribution, with a single model is possible to represent both hyperbolic and exponentially-tailed multiplicative processes.

The great innovation of the model is that it is conceived to be applied over a continuum of scales, through scale-densification; differently from Equation 5.17, the parametrization does not depend on the ratio between the resolution of two adjacent scales b .

The Universal Multi-fractal model is defined through its moment scaling function K :

$$\begin{cases} K(q) = -qH + C_1 \frac{q^\alpha - q}{\alpha - 1} & \text{for } \alpha \neq 1 \\ K(q) = -qH + C_1 q \log(q) & \text{for } \alpha = 1 \end{cases} \quad (5.18)$$

H is the Hurst exponent expressing the non-conservativity of the field ($H=0$ in conservative fields), C_1 is the mean co-dimension of the process (representing the degree of sparseness of the mean) and α is the Levy's stability index.

The expression of $K(q)$ corresponds to $-\tau(q)$ introduced in Section 5.5.2; we find convenient to report the two approaches separately, since i) they have been originated from two studies conducted at the same time; ii) nowadays none of the two notations has prevailed over the other.

Knowing that $\gamma_s(q) = \frac{dK(q)}{dq}$, it is also possible to express the Universal Multifractal Model in terms of the singularity strength γ_s , defined as

$$\gamma_s(q) = \frac{dK(q)}{dq} \quad (5.19)$$

The moment scaling function K and the codimension function c are related through the Legendre and inverse Legendre transform, respectively (*Frisch and Parisi, 1985; Schertzer and Lovejoy, 1987*):

$$K(q) = \max_{\gamma_s} [q\gamma_s - c(\gamma_s)] \quad (5.20)$$

and

$$c(\gamma_s) = \max_q [q\gamma_s - K(q)] \quad (5.21)$$

Figure 5.7 shows the duality between scaling in probability-distribution (through $c(\gamma_s)$) and scaling in statistical-moments (through $K(q)$) for a 1D series, when $H = 0$. We see that while the first moment (average of the process) does not vary with the scale (in fact $K(1) = 0$ in the Universal Multifractal Model), the second moment varies with scale, in accordance with the changes in the distribution variability (variance). A similar pattern concerns the third moment, that is related to the asymmetry of the distribution.

The Universal Multifractal Model can be expressed as a function of the codimension $c(\gamma_s)$, for $0 \leq \alpha \leq 2$ and for $H \neq 0$:

$$\begin{cases} c(\gamma_s + H) = C_1 \left(\frac{\gamma_s}{C_1^{\alpha'}} + \frac{1}{\alpha} \right)^{\alpha'} & \text{for } \alpha \neq 1 \\ c(\gamma_s + H) = C_1 \exp \left(\frac{\gamma_s}{C_1} - 1 \right) & \text{for } \alpha = 1 \end{cases} \quad (5.22)$$

where $\alpha' = \alpha/(\alpha - 1)$. The derivation of the expression above, starting from the moment scaling function, is reported in Appendix F for the case of $\alpha \neq 0$ and $H = 0$.

The relation between moment scaling function $K(q)$ and codimension function $c(\gamma_s)$ is shown in Figure 5.8 for the case $\alpha = 1.75$, $C_1 = 0.15$ and $H = 0$. From these figures it appears clear the meaning of the Legendre transform for relating the moment scaling function K and the codimension c : it corresponds to the maximum distance between the function $q\gamma_s$ (in both cases the dashed line with zero intercept) and the function $K(q)$ or $c(\gamma_s)$, in Figure 5.8-a and b, respectively.

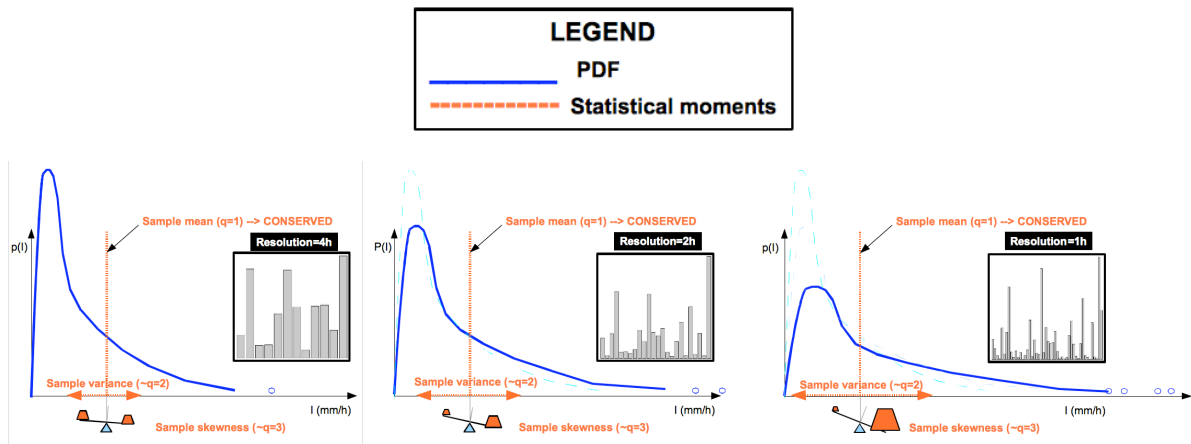


Figure 5.7: Intuitive scheme of the probability distribution change with the scale in multi-scaling process. The blue line identifies the probability distribution, while the orange line refers to the centred statistical moments of the distribution.

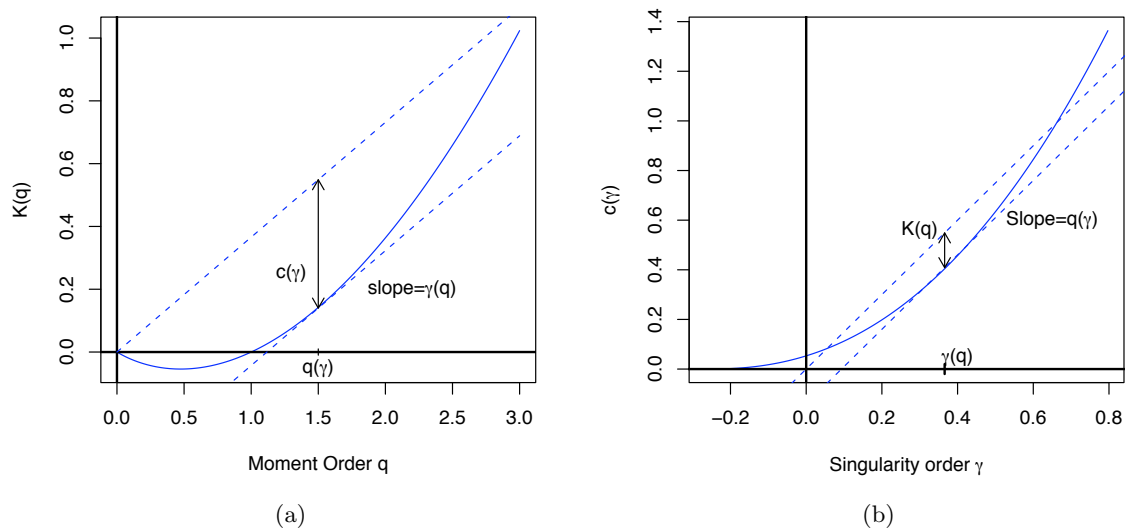


Figure 5.8: a: Moment scaling function $K(q)$ of the Universal Fractal Model, relation between $K(q)$, codimension c and γ_s . b: Codimension function of the Universal Fractal Model and relations between $c(\gamma_s)$, moment order q and moment scaling function $K(q)$. Taken from *Tessier et al. (1993)*.

The expression “Universal Multifractal” refers to the possibility to model multiplicative processes characterized by either heavy or exponential tails. Actually, the approach followed by *Gupta and Waymire* (1993), through application of the MKP expression (Equation 5.16), is more general: any probability distribution can be implemented into the model.

5.6 Towards an unified multifractal formalism

The Generalized Structure Function theory (Section 5.5.1) and the Moment Scaling theory (Section 5.5.2) share several features:

- both originate from the generalization of the scaling of turbulent flows of *Kolmogorov* (1941);
- both are expression of the non-linearity of geophysical processes;
- a Legendre transform relates the GSF scaling function $\zeta(p)$ and the GSF singularity spectrum $D(h)$; the same relation establishes between the moment scaling function $\tau(q)$ and the codimension function $c(\gamma_s)$;

Vainshtein et al. (1994) define the relation between the scaling functions of GSF and of the scaling of the moments. *Menabde et al.* (1997) expresses this relation as:

$$\tau(q) = \zeta(1) \cdot q - \zeta(q)$$

The singularity spectrum $D(h)$, estimated through Legendre transformation in Equation 5.10 as a function of the scaling exponent of GSF, can be equally derived as a function of the moment-scaling function $\tau(q)$, as:

$$D(h) = \min_q [qh - \tau(q)] \quad (5.23)$$

The derivation as a function of $\tau(q)$ is more general than the one in Equation 5.10, since $\tau(q)$ is, in most cases, the exact Legendre transform of the $D(h)$ singularity spectrum (*Muzy et al.*, 1993).

5.7 Multifractal spectrum estimation

In this section we briefly define the methods for estimating the multifractal spectrum of a sample set. We focus on the estimation of the moment-scaling function given by *Schertzer and Lovejoy* (1987) and *Lavallée* (1991), and on an alternative technique based on wavelets proposed by *Muzy et al.* (1993), successfully developed by *Venugopal et al.* (2006a,b).

5.7.1 Moment-based estimation

To analyze the multi-fractal character of a finite sample, two methods have been successively implemented: the Trace Moment and the Double Trace Moment. The moment analysis should be performed

paying attention to the chosen moments: low order moments can be affected by the instrument resolution while, for high order moments, the strongest observation can be predominant², leading to linearity of the moment scaling function $K(q)$ for q higher than a critical moment q_s .

According to *Schertzer and Lovejoy* (1992) and *Tessier et al.* (1993), the critical moment q_s is defined as:

$$q_s = \left(\frac{D + D_s}{C_1} \right)^{1/\alpha} \quad (5.24)$$

where D is the euclidean dimension of the field and D_s the sampling dimension, defined as

$$D_s = \frac{\log(N_s)}{\log(\Lambda)} \quad (5.25)$$

where N_s is the number of elements at the finer resolutions that compose the coarse-resolution field and Λ is the ratio between large scale and small scale.

The moment scaling function $K(q)$ can be computed by means of the Trace Moment (Section 5.7.1.1) and Double Trace Moment (Section 5.7.1.2) methods, as a function of the two multi-fractal parameters α and C_1 .

5.7.1.1 The Trace Moment Method

The Trace Moment method (*Schertzer and Lovejoy*, 1987) consists in evaluating the scale invariance of a process by applying a two-step algorithm. In a first step, one plots the statistical raw moments of various orders on double logarithmic diagram as a function of the aggregation scale (Figure 5.9-a reports a schematic example of the statistical moments of order ranging from 0.5 to 2); in the range of log-log linearity of moments, the moment slope ($K = K(q)$) is calculated by least squares optimization. In the second step the empirical slope $K(q)$ is plotted as a function of the moment order q (Figure 5.9-b), and the Universal Multifractal law

$$K(q) = C_1 \frac{q^\alpha - q}{\alpha - 1} \quad (5.26)$$

is fitted to the empirical data. α is the Levy's stability index and C_1 the mean co-dimension of the process. The main drawback of the method is that the correct fitting of the empirical moment scaling function with a two-parameter function can be affected by the inverse dependence between α and C_1 . A second drawback is the fact that, being a two-step process, the estimation uncertainties can not be directly assessed by maximum likelihood expressions, and numerical estimations, such as jackknife methods, need to be used.

5.7.1.2 The Double Trace Moment Method

The inter-dependence between C_1 and α of the Trace Moment Method limits the accuracy of estimation of the two multifractal parameters. *Lavallée* (1991) propose a new method for the characterization

²This phenomenon is referred to as "hard phase transition" by *Schertzer and Lovejoy* (1992)

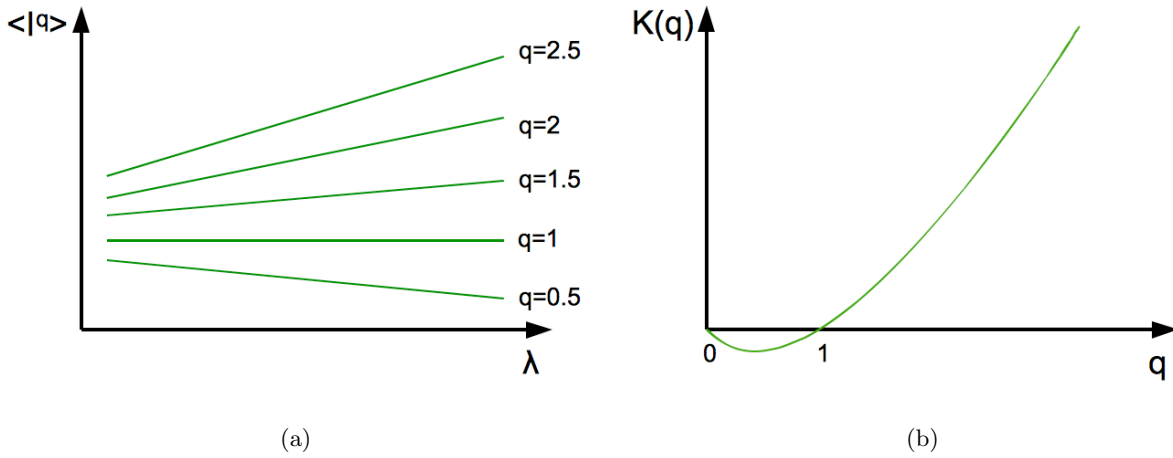


Figure 5.9: Schematic of the Trace Moment estimation method. a: empirical moments as a function of the scale λ . The slope of the straight line for each q corresponds to $K(q)$. b: Moment scaling function K as a function of the moment order q .

of the multifractal behavior of processes. The method has been referred to as “Double Trace Moment” (DTM), because it uses an auxiliary moment η other than q .

The first step is to raise the initial field to a set of auxiliary moments η , obtaining a series of different fields which will be separately analyzed with a two-step algorithm similar to the TM method. For each raised field, the log-log slope of the moments $K(q, \eta)$ is computed as a function of the scale λ (in Figure 5.10-a). The last step is to plot $\log(K(q, \eta))$ empirical moments as a function of η (Figure 5.10-b). For sufficiently small and high η values, the $K(q, \eta)$ function does not depend on η . For η close to the unit, the function has a linear behavior with respect to $\log(\eta)$, whose slope corresponds to α . The moment scaling function $K(q, \eta)$ can be expressed as a function of the moment scaling function $K(q)$ of the Universal Multifractal Method by the following Equation 5.27:

$$K(q, \eta) = \eta^\alpha K(q). \quad (5.27)$$

The C_1 parameter is consequently evaluated by inverting Equation 5.26 for $\eta = 1$. The advantage of the technique is that the Levy’s stability index α is obtained independently from C_1 . Compared to TM method, the parameter estimation is improved by the double log-transformations that yield normally-distributed residuals. The DTM provides a very efficient method for estimating the multifractal parameters. Nevertheless, its implementation is submitted to arbitrary interpretations. Let us examine such points in detail:

- the inflexion point of $K(q, \eta)$, whose slope yields the α value, can appear relatively far from $\eta = 1$; the inflexion point is found by imposing null second derivative of an arbitrary function (e.g. a third order polynomial) that has been fitted to the empirical values of $K(q, \eta)$;
- the choice of the set of η values upon which to perform the fitting is arbitrary;
- the choice of q seems to affect the estimation of α . Given that for $q = 1$ the DTM can not be evaluated, α is usually assessed taking arbitrary values of q higher than the unit and lower

than q_D (for $q \geq q_D$ the moments diverge). For example, *Ladoy et al.* (1993) chose $q = 1.5$ and *Tessier et al.* (1993) $q = 2$.

Despite these arbitrary choices, the DTM method provides the better multifractal estimation, and therefore it should be preferred. Nevertheless, in case of intensive data analysis, such as in the case of regionalization of rainfall data at a station, it is preferable to adopt a completely automatic process working for all stations in the same way, similarly to Trace Moments method.

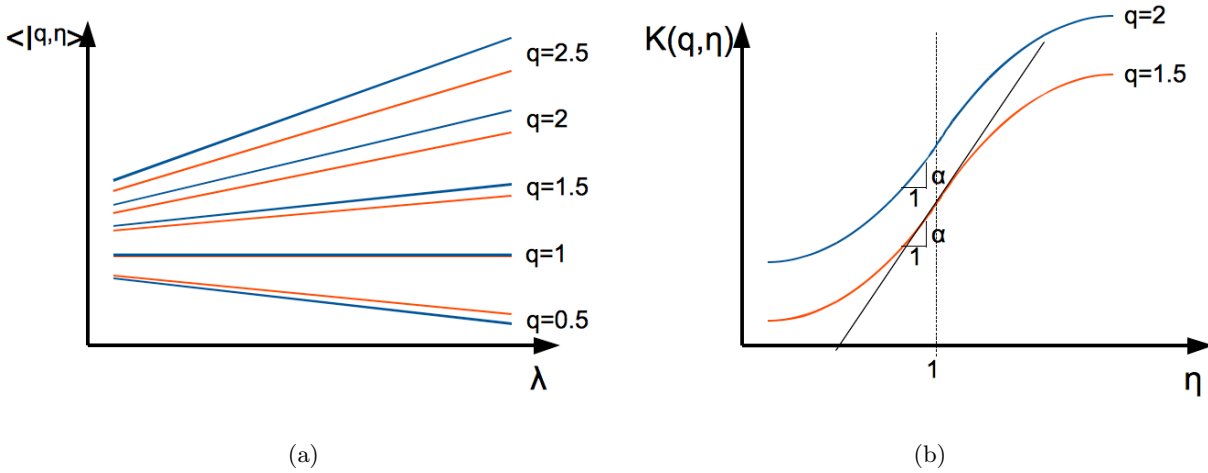


Figure 5.10: Schematic of the Double Trace Moment estimation method. a: Empirical moments as a function of the scale λ , for a given value of the auxiliary moment η . The slope of the straight line for each q corresponds to $K(q)$. b: Moment scaling function K as a function of the moment order q . The blue color identifies the value $\eta = 2$, the orange $\eta = 1.5$. Both the plots are in double logarithmic scale.

5.7.2 Wavelet Estimation

The methods above cited (Equation 5.18, Equation 5.9) allows to obtain, after Legendre transformation, the singularity spectrum of the field. *Venugopal et al.* (2006a) established that traditional techniques can result in spurious estimates of the spectrum of scaling exponents $\zeta(q)$ (Equation (5.9)) and consequently of the singularity spectrum $D(\gamma_s)$ (equal to $D - c(\gamma_s)$ where D is the euclidean dimension of the field, e.g. 1 for time series). An alternate formalism based on wavelets is proposed for the direct estimation of the singularity spectrum. The Wavelet Transform Modulus Maxima coefficients (WTMM, *Muzy et al.* (1993)) may provide a robust estimate of the singularity spectrum $D(\gamma_s)$. A great advantage of this technique compared to the Continuous Wavelet Transform and to the moment analysis, is that the whole range of singularities is explored, including the decaying part of the $D(\gamma_s)$ curve (not possible with the moment analysis, due to the impossibility to explore negative moments).

Particular attention is devoted to the integration order of the signal and its relationships with the wavelet used for the analysis. For instance, *Venugopal et al.* (2006b) used wavelets of increasing order, corresponding to increasing order of the moments; this allows to use integrated signals characterized by various integration orders. The analysis performed using wavelets of increasing orders allows to rely on the independence of the singularity spectrum relatively to the used wavelet.

The process is implemented following four steps:

- the CWT (Continuous Wavelet Transform) is applied to the data series. The wavelet coefficients for each time step and for each scale are provided. The maxima lines, or branches, illustrate the hierarchical singularity structure.
- the WTMM analysis is performed. At each scale of analysis, a research of the relative maxima is conducted. This allows to identify the trees, i.e. the series of maxima which are related at different scales.
- only the wavelet coefficients relative to each branch are retained and the structure function $\zeta(q)$ of these singularities is evaluated, considering that

$$N_\lambda \langle |T_\lambda(x)|^p \rangle \sim \lambda^{\zeta(p)} \quad (5.28)$$

where N_λ is the number of maxima lines at the scale λ and the symbol $\langle \rangle$ denotes the ensemble average of the field.

- through the Legendre transform in Equation 5.20, the $\zeta(p)$ structure function is used to determine the singularity spectrum $D(\gamma_s)$.

The main limit of this technique is the impossibility to estimate the rain-no rain fractal exponent (i.e. the co-dimension of the support); therefore the technique is not suitable for dealing with rain-no rain intermittency.

To illustrate this technique, we analyze a 10-hours event recorded in Ales, South-East of France by a disdrometer. The temporal resolution is 1 min. The time series of the rainfall intensity is shown in Figure 5.11-a. The wavelet spectrum with the main branches is reported in Figure 5.11-b, whereas the consequent evaluation of the moment scaling function ζ and of the singularity spectrum is reported in the Figure 5.11-c.

Comparing the results of the WTMM method with the structure-function techniques for the analysis of rainfall intensities, *Venugopal et al.* (2006b) found that structure function analysis provides spurious results. They state that a proper small-scale analysis of rainfall intensity must be performed through higher order vanishing moments, able to detect all the singularity of the signal. A rainfall intensity series, for instance, cannot be analyzed through a Gaussian wavelet but rather by a Mexican hat wavelet (2nd order derivative of the Gaussian).

This technique has been demonstrated to be reliable for the analysis of 1D signals respect to the GSF or moment-scaling analyses. However, applications of WTMM method are, at the present day, limited to 1D fields and the increase of complexity of the wavelet theory for n-dimensional fields may limit these analyses.

5.8 Scale-invariance of spatial fields

The observation of geophysical variables such as rainfall and cloud cover by means of satellite imagery, often limited in resolution but largely extended in space, made possible the multi-scale analysis of

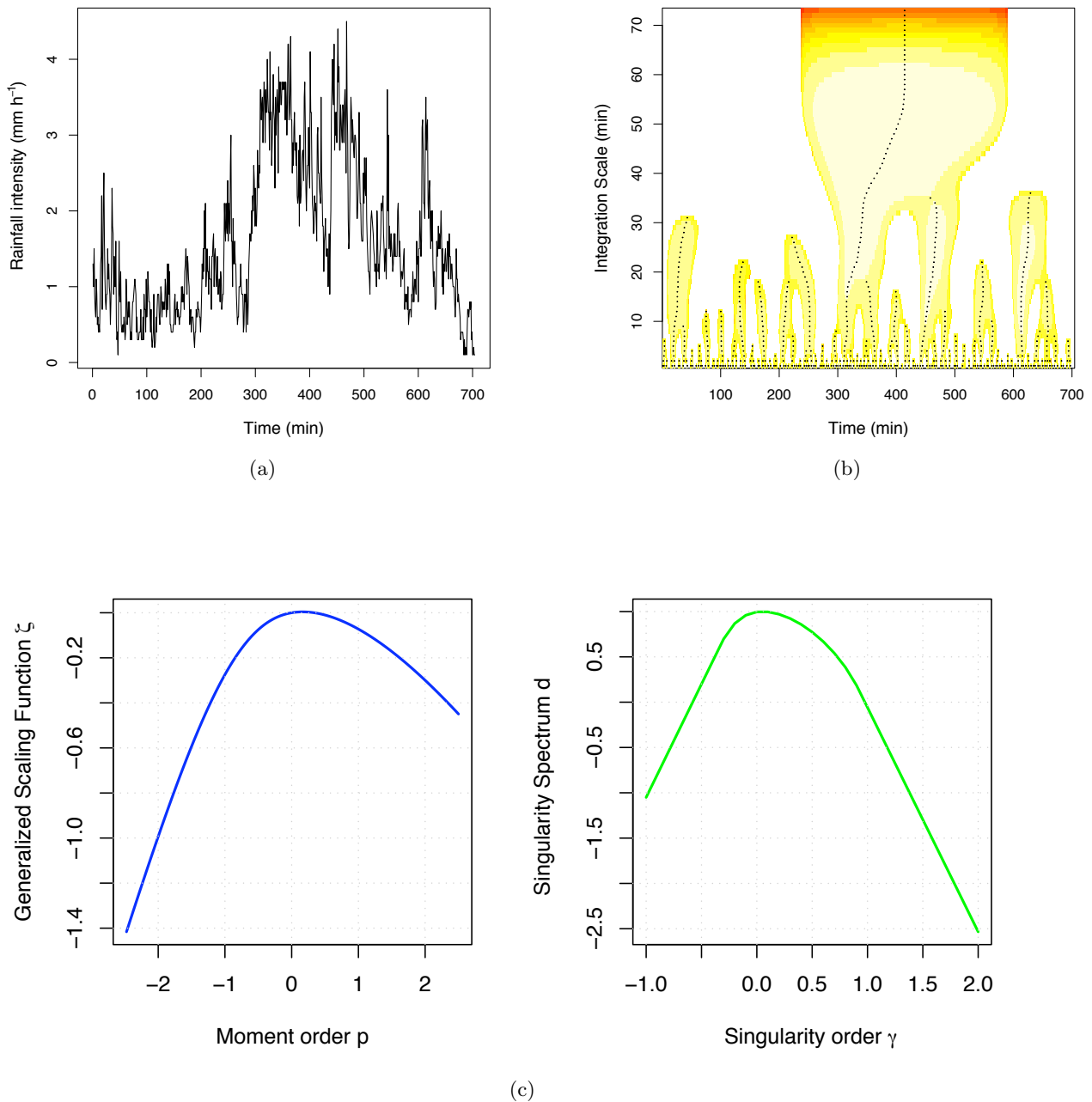


Figure 5.11: Disdrometer scan of an event lasting 700 hours measured at the Alés Disdrometer. a: Time series. b: wavelet spectrum with respect of the time and of the scale. c: Moment scaling function $\zeta(q)$ and singularity spectrum $d(\gamma_s)$. A Mexican hat wavelet has been used.

spatial fields. The gain respect to the older statistical methods for the modeling of rainfall events (see *Waymire and Gupta* (1981) for a detailed review) is to provide a scale-free description of the atmospheric phenomena. In one of the first applications, *Lovejoy* (1982) studied the geometrical relations of rainy areas and cloud covering such as the ratio between area and perimeter of cloud and rainfall fields. However, the main idea of scale-invariance analysis is to apply the scale-invariance concepts introduced for turbulent fluxes by *Frisch and Parisi* (1985) for the large-scale analysis of atmospheric spatial fields.

The earlier studies of the scale-invariance of spatial rainfall have been motivated by the evidence that rainfall can be organized into embedded structures (*Austin and Houze*, 1972; *Gupta and Waymire*, 1979): small-scale clusters of high rainfall intensity are embedded within low-intensity meso-scale areas, in turn embedded within larger area associated with lower intensities up to the synoptic scale (Figure 5.12).

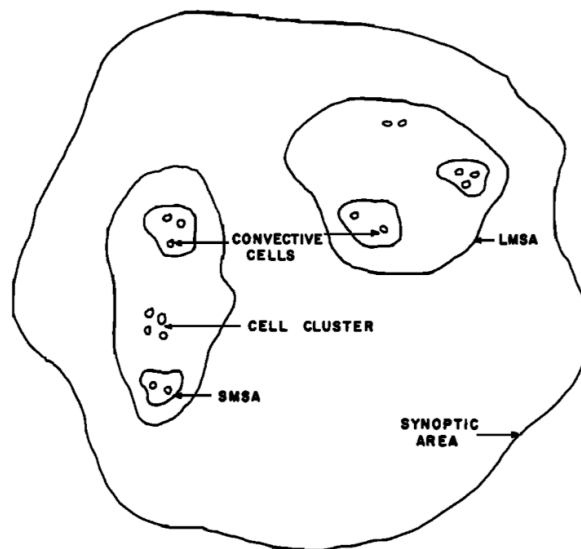


Figure 5.12: Schematic depiction of sub-synoptic rainfall features (from *Gupta and Waymire* (1979))

The spatial scale invariance modeling has been the object of several studies (*Schertzer and Lovejoy*, 1987; *Tessier et al.*, 1993; *Olsson and Niemczynowicz*, 1996; *Olsson et al.*, 1999; *Harris et al.*, 2001). *Schertzer and Lovejoy* (1987) verified the moment scaling of radar reflectivity measured with the volumetric radar of McGill Weather Radar, Canada. Based on these data, they developed empirical techniques such as TM (Section 5.7.1.1) to detect the multiscaling behavior of spatial fields. Using similar techniques, *Tessier et al.* (1993) showed that cloud and rainfall images taken from Landsat, Meteosat and NOAA satellites are scale-invariant, adopting the DTM technique (Section 5.7.1.2). The multifractality of “spatialized” (i.e. interpolated) point rainfall has been studied in *Olsson and Niemczynowicz* (1996), by analyzing a uniform rain gauge network located in Southern Sweden; scale-invariance in spatial scales ranging between 70 and 8000 km^2 has been found. The effect of spatial averaging on the multifractal behavior of daily rainfall has been examined in *Olsson et al.* (1999). *Harris et al.* (2001) applied the GSF, as well as the spectrum analysis and the moment analysis,

to compare the capability of the model forecasts to reproduce the scale-invariant structure of radar-observed fields, which is verified for spatial resolutions in the range 1-15 km.

We have seen in Section 5 that the Fourier analysis may only provide a quantitative assessment of scale-invariance. Showing several advantages, wavelets have been tested for the characterization of the spatial scaling. In spatial fields, the early usages of wavelet analysis are relative to the studies of *Kumar and Foufoula-Georgiou* (1993) and *Perica and Foufoula-Georgiou* (1996). Both studies focused on the characterization of the self-similarity of rainfall fluctuations, by using the simplest wavelet, i.e. the Haar Wavelet, defined as:

$$\psi(t) = \begin{cases} 1, & 0 \leq x < \frac{1}{2} \\ -1, & \frac{1}{2} \leq x < 1 \\ 0, & \text{otherwise} \end{cases} \quad (5.29)$$

In practice, *Perica and Foufoula-Georgiou* (1996) described the “standardized rainfall fluctuations”, defined as $\varepsilon_{m,i} = \frac{X'_{m,i}}{X_m}$, where

$$\begin{cases} X_{m,1} = \frac{1}{4}[(\bar{X}_{m-1}(i,j) + \bar{X}_{m-1}(i,j+1)) - (\bar{X}_{m-1}(i+1,j) + \bar{X}_{m-1}(i+1,j+1))] \\ X_{m,2} = \frac{1}{4}[(\bar{X}_{m-1}(i,j) + \bar{X}_{m-1}(i+1,j)) - (\bar{X}_{m-1}(i,j+1) + \bar{X}_{m-1}(i+1,j+1))] \\ X_{m,3} = \frac{1}{4}[(\bar{X}_{m-1}(i,j) - \bar{X}_{m-1}(i+1,j)) - (\bar{X}_{m-1}(i,j+1) - \bar{X}_{m-1}(i+1,j+1))] \end{cases} \quad (5.30)$$

are the 3 directional fluctuation components and

$$\bar{X}_m = \frac{1}{4}(X_{m-1}(i,j) + X_{m-1}(i,j+1) + X_{m-1}(i+1,j) + X_{m-1}(i+1,j+1)) \quad (5.31)$$

is the local mean of the process. At each scale, the variable ε_m has an approximately gaussian distribution and it is centered at 0, so that the only standard deviation of ε_m allows a complete characterization of the rainfall fluctuations.

This approach has similarities with both the GSF and moment analysis, in that:

- the field is subsequently averaged and the statistics computed on the degraded fields;
- fluctuations are examined, but differently from GSF, only distances equal to the spatial resolution are taken;

Through this kind of analysis, no statistical moments are explored but information on three fluctuation components (“horizontal”, “vertical” and “diagonal”) is detected. In this approach, a field is considered as the superposition of a mean field and of a fluctuation field. It is significant to notice that this wavelet algorithm is commonly used in data compression algorithms.

5.9 Scale-invariance of time series

Up to this section, we examined the spatial scale invariance. The same approaches have been extended to 1D random variables with similar results. As we highlight in Section 5.9.1, scale invariance and

long-memory of the processes are concepts well known since centuries, but the first theoretical analyses corroborated by empirical observations are relatively recent.

5.9.1 Long-range correlation: the first evidence of long-memory systems

Moving from geographical to geophysical variables, *Mandelbrot and Wallis* (1969a) surprisingly found that most of the processes exhibit the so-called long-range dependence. This finding has very important consequences on the perception of natural processes: they conserve the memory of their states for ages or decades. References to the long-memory effect are present in the oral tradition. As reported by *Mandelbrot and Wallis* (1969b), the biblical Joseph Effect expresses the fact that high or low levels in rivers tend to persist “seven fat and seven lean years”. Similar concepts and observations are found, among the others, in meteorology, geophysics, hydrology, physics and economics.

Let us introduce the concept of long-range dependence as defined by *Mandelbrot and Wallis* (1969a). $R(t, s)$ is the maximum difference between two values contained in the interval within t and $t + s$, and it is defined as in Equation 5.32:

$$R(t, s) = \max_{0 \leq u \leq s} [X(t + u) - X(t) - (\frac{u}{s})(X(t + s) - X(t))] + \min_{0 \leq u \leq s} [X(t + u) - X(t) - (\frac{u}{s})(X(t + s) - X(t))] \tag{5.32}$$

$S(t, s)$ is defined as the sample variance of the sub-record from time $t + 1$ to time $t + s$ (Equation 5.33):

$$S^2(t, s) = s^{-1} \sum_{u=t+1}^{t+s} X^2(u) - [s^{-1} \sum_{u=t+1}^{t+s} X(u)]^2. \tag{5.33}$$

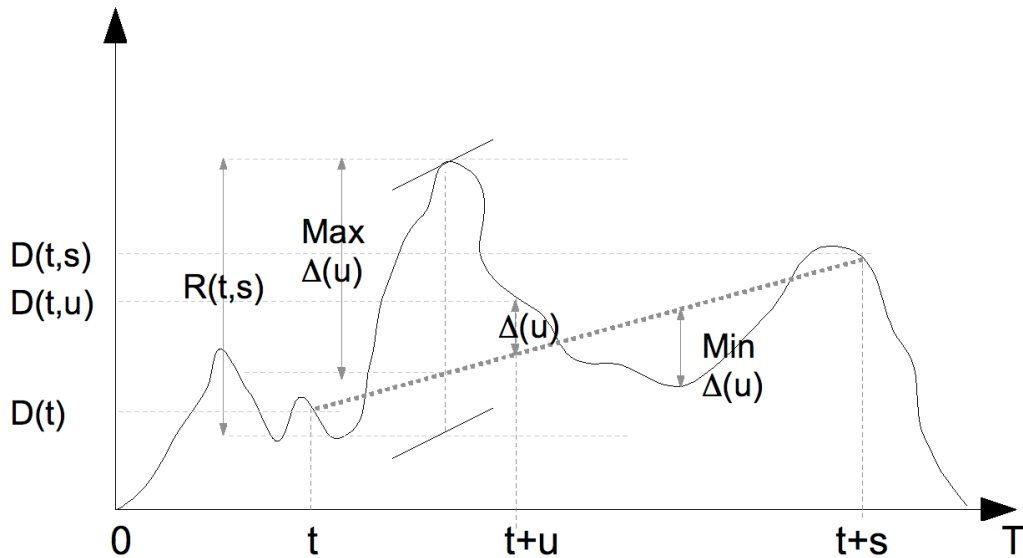


Figure 5.13: Construction of the sample range $R(t,s)$ for a 1D process X reproduced from *Mandelbrot and Wallis* (1969a).

According to *Mandelbrot and Wallis* (1969a,b), the deviation $R(t, s)$, normalized by the variance of the process S^2 , is 0.5 in Brownian motion, while it is higher in natural phenomena. *Mandelbrot*

and Wallis (1969b) stress that an apparent long-range correlation can be given by the seasonality of series, that should be removed before the computation.

5.9.2 Scaling range of rainfall time series

Thanks to an electromechanical disdrometer called “sonic gauge”, *Fabry* (1996) explored the scaling features of high-resolution rainfall intensities. In this study, scales between 0.1 s and 1 hour are explored by comparison with other measurements, such as the vertical-pointing radar of McGill or some daily gauge series. Analyzing the power spectrum (Section 5.4) of the sonic gauge time series, the presence of a scaling regime in the range 10 s-1000 s has been verified. The scaling behavior is confirmed by the vertical-pointing radar analysis, for rainfall as well as solid precipitations (radar scan at height = 3500 m).

In the study, *Fabry* (1996) expressed the temporal scale in terms of spatial scale, by assuming constant advection in the order of 10 m s^{-1} .

Figure 5.14 shows the results obtained by *Fabry* (1996) gathering the power spectrum of sonic gauges with drop-counting and daily gauge devices. The graph shows at least three scaling regimes, from 5 s to 1 h (Turbulence-driven structures), from 2 h to 15 days (baroclinic forcing) and from 15 days to some years (succession of weather systems), where the spectrum is flat.

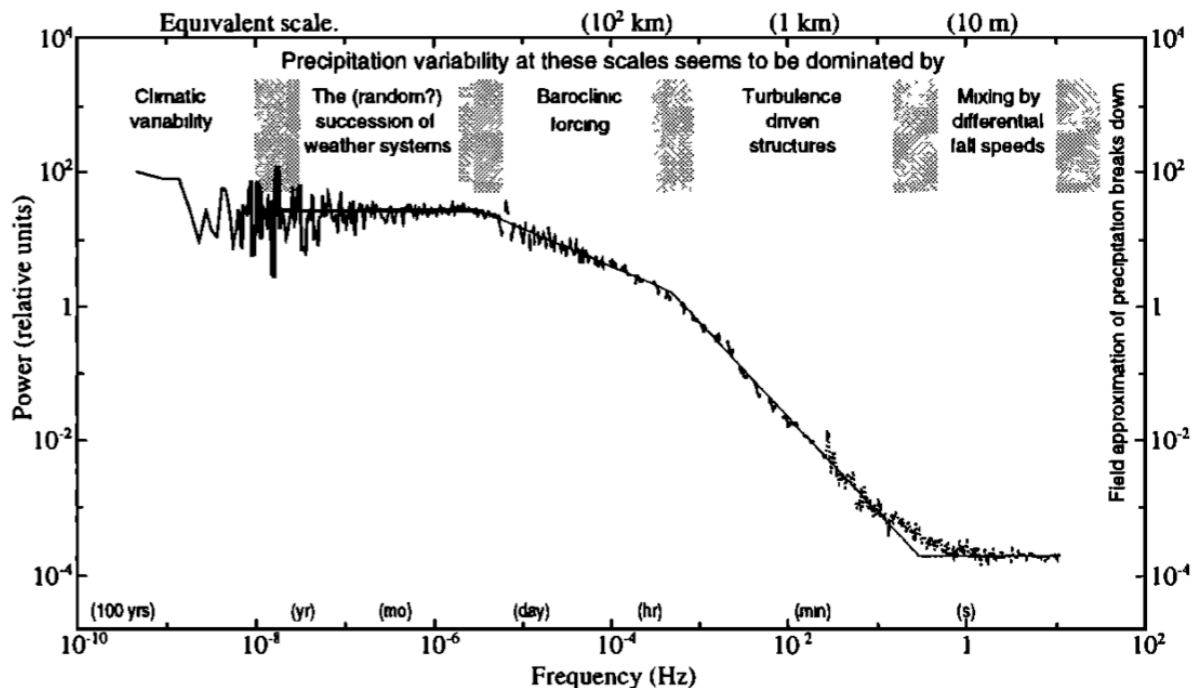


Figure 5.14: Normalized power spectrum of precipitation rates derived from three sources (*Fabry*, 1996). The range 0.1 s-10 min are results of the sonic gauge measurement in Florida and Colorado. Scales from 10 min to 1 day are given by a drop-counting rain gauge; scales larger than 1 day result from the analysis of the daily gauge at McGill. The equivalent spatial scale has been computed considering the average speed of weather echoes in Montreal as 13.2 m s^{-1} .

5.9.3 Point-rainfall scale-invariant models

5.9.3.1 Earlier studies

One of the earlier applications of multifractality to rainfall series is presented by *Hubert and Carbonnel* (1989). The fractal dimension of two rainfall series located in Burkina Faso, specifically Dédougou and Ouagadougou, was detected. A box-counting method has been applied to compute the fractal dimension.

The station of Dédougou showed scale-invariance in the range 8-128 days ($D=0.79$) and another scale-invariant regime in the range 256-16384 days ($D=1.00$). The Ouagadougou station showed three scaling regimes, 4-32 days ($D=0.22$), 64-256 days ($D=0.74$) and 512-4096 days ($D=1.00$).

The spatial rainfall analyses (*Schertzer and Lovejoy*, 1987; *Gupta and Waymire*, 1990) demonstrated that rainfall fields display an infinite number of fractal dimensions, depending on the intensity level. The possibility to apply this intriguing property to rainfall time series has been tested in numerous studies.

Hubert et al. (1993) explored the possibility to model the scaling behavior of rainfall series by using the multi-fractal model, with the advantage to deal with heavy-tailed series. The basic equation of multi-fractal fields states that

$$P(R_\lambda > \lambda_s^\gamma) \sim \lambda^{-c(\gamma_s)} \quad (5.34)$$

where λ is the scale ratio of two time intervals D/d , R_λ the intensity of the field at the scale ratio λ and γ_s the singularity order.

The analysis took into account the rain-no rain intermittency as a particular value of the multi-fractal field and determined values of the α Levy's stable parameter (Section 5.5.2.3) close to 0.5 for at least 4000 daily rainfall series all around the world. In particular, the study found scaling of point rainfall intensities up to 16 days. The scaling has been verified also for infra-daily series at Réunion island and in the Alps.

In the same year, *Ladoy et al.* (1993) described the temporal variability of rainfall observations at the Nîmes station, in southern France. They computed the parameters of the Universal Multi-fractal model (Section 5.5.2.3) in the range 12 h to 16 days, finding $\alpha = 0.45 \pm 0.05$. C_1 is estimated to be 0.6 ± 0.02 .

The multi-fractal analysis of time series has been tested for different data sets (*Fraedrich and Larnder*, 1993; *Olsson et al.*, 1993; *Kumar et al.*, 1994; *Olsson*, 1995, 1996; *Tessier et al.*, 1996; *Svensson et al.*, 1996; *Olsson*, 1998).

Fraedrich and Larnder (1993) analyzed rainfall intensities from the point of view of the Generalized Structure Function (GSF, Section 5.5.1) and of the power spectrum, finding, for continental European stations, different scaling regimes (Figure 5.15). An interesting analysis on the distribution tails confirms the hyperbolic shape of the survival probability.

Olsson et al. (1993), *Olsson* (1995) and *Olsson and Niemczynowicz* (1996) examined the rainfall scaling from a multi-fractal point of view, detecting different scale-invariance ranges depending of the analyzed location. A similar work has been proposed by *De Lima et al.* (2003). *Kumar et al.* (1994) point out that the higher moment analysis (Section 5.5.2) could be biased due to moment

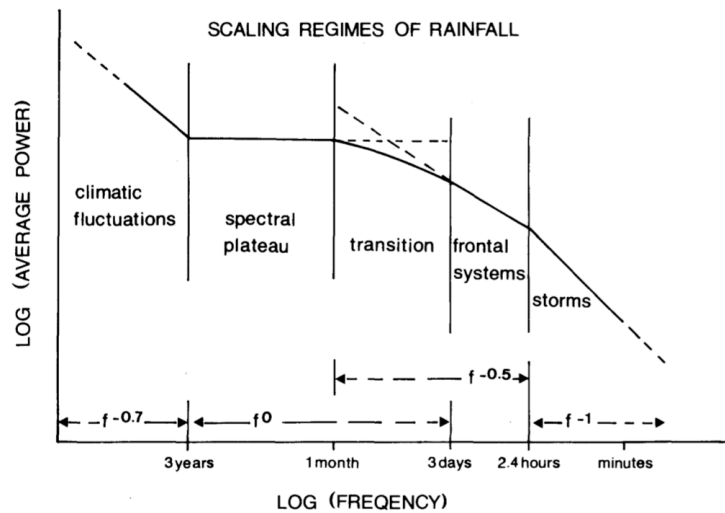


Figure 5.15: The scaling regimes of time rainfall series with respect to the accumulation duration, for the continental Europe (from *Fraedrich and Larnder (1993)*).

divergence proper of high-variable geophysical fields, and propose a Probability Weighted Moment method. *Tessier et al. (1996)* showed that it is possible to relate the multi-fractal character of time series to that of river flow series (in particular, the only changing parameter in Equation 5.18 is H), giving the first elements of the hydrological applications of fractal analysis. Other studies on the river flow multi-fractal analyses followed (as an example, *Pandey et al. (1998)*).

Svensson et al. (1996) used GSF method (Section 5.5.1) to detect the structure function $\zeta(q)$ of the rainfall fluctuations for different storm typologies. *Harris et al. (1996)* evaluated the orographic effects on the scaling behavior of rainfall series. Analyzing spectral density, Generalized Structure Function, moment analysis and the CDF tail of a number of rain gauges along a section of increasing elevation, they found a significant influence of elevation in the determination of the scaling character of rainfall.

5.9.3.2 Dealing with rain-no rain intermittency

All the studies cited above considered the rain-no rain intermittency as a particular case of the multifractal field. Discordantly, *Over and Gupta (1994)* and *Schmitt et al. (1998)* demonstrated that the rainfall variability and the rain-no rain intermittency are two separate processes and therefore an appropriate modeling of intermittency was needed. Similar evidences of the independence between intermittency and rainfall variability have been found by *Barancourt et al. (1992)* through a geostatistical approach.

These evidences prevent the application of classical multi-fractal cascades (*Tessier et al., 1993; Ladoy et al., 1993*), in which the rainfall intermittency is obtained by fixing a threshold below which the rainfall intensity is considered as zero. From a numerical point of view, this does not cause problems to the multi-fractal analysis but, as pointed out by *Schmitt et al. (1998)*, the Levy's stability index α is considerably underestimated leading to extremely heavy-tailed distributions.

The estimation of the multifractal parameters in scale-invariant processes with an atom at zero (i.e. the rain-no rain intermittency process is seen as a simple fractal process), is implemented by *Schmitt et al.* (1998) with the aim to stochastically simulate rainfall series. This parsimonious model, obtained by simply adding one parameter ruling the intermittency at different scales to the multifractal formulation (Equation 5.18), can be written as:

$$K(q) = C_\beta(q - 1) + C_1 \frac{q^\alpha - q}{\alpha - 1} \quad (5.35)$$

where C_β is the co-dimension of the rain-no rain intermittency, expressing how it increases with the resolution.

For pure multi-fractal processes, *Lavallée* (1991) demonstrated that DTM is the best method to estimate the parameters of a multi-fractal process. DTM does not yield the expected results, since

$$K(q, \eta) = \eta^\alpha K(q) - (q - 1)c_\beta, \quad (5.36)$$

The solution proposed by *Schmitt et al.* (1998) is to evaluate the moment scaling function by the TM method (Section 5.7.1.1). Figure 5.16 reports the statistical moments, the scaling moment function of wet observations and the scaling moment function of the overall process relative to the study of the rainfall series at Uccle, Belgium (from *Schmitt et al.* (1998)). Even if the moment scaling is verified (Figure 5.16-a), the moment scaling function K must take into account the effect of intermittency and Equation 5.36 must be adopted.

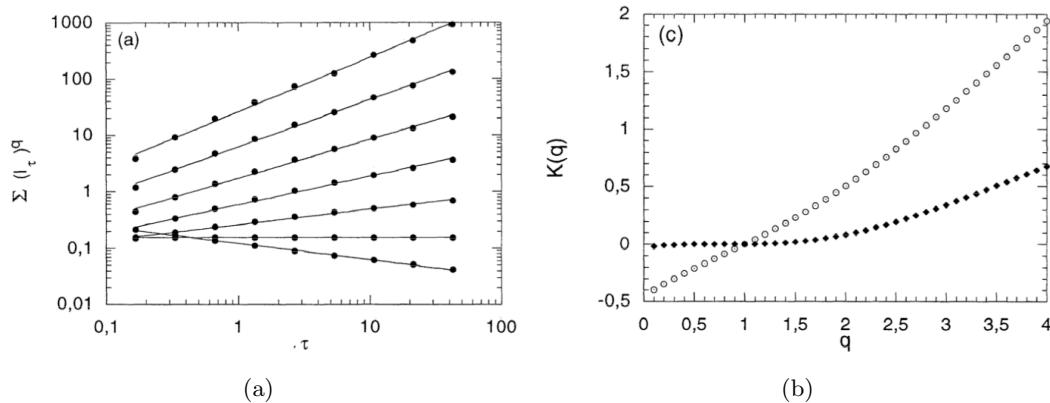


Figure 5.16: From *Schmitt et al.* (1998), Figure 7. a: Trace moments as a function of the temporal scale. From bottom to top, the moment orders increase from 0.5 to 3.5 by 0.5; b: Moment scaling function K obtained taking into account the fractal support (open circles), and moment scaling function computed taking only the positive observations.

The cascade modeling proposed by *Schmitt et al.* (1998) produces a rainfall series in which most of the rainfall statistics are strictly similar to those of a real rainfall series. Average intermittency, average intensity, variance as well as extreme behavior are correctly reproduced. Only the duration of wet events is significantly underestimated. Similar stochastic methods have been adopted by *Over and Gupta* (1996); *Olsson* (1998); *Güntner et al.* (2001); *Rupp et al.* (2009).

5.10 Scale invariance of space-time rainfall

The development and validation of scale-invariance stochastic theories of rainfall focused firstly on the spatial properties of rainfall fields (*Gupta and Waymire*, 1990, 1993; *Tessier et al.*, 1993; *Kumar and Foufoula-Georgiou*, 1993; *Over and Gupta*, 1994; *Perica and Foufoula-Georgiou*, 1996; *Marani*, 2003) and later on their temporal properties (*Hubert et al.*, 1993; *Ladoy et al.*, 1993; *Olsson*, 1995; *Schmitt et al.*, 1998; *Marani*, 2005). Rarely the research has been devoted to the space-time scaling. The space-time modeling is essential for at least two reasons:

- i. they are the only models able to disaggregate single events respecting the coherency between scales;
- ii. it is well known (since *Bras and Rodriguez-Iturbe* (1976)) that a correct space-time modeling can not deal with the spatial and temporal scales separately.

A number of studies (*Zawadzki*, 1973; *Venugopal et al.*, 1999) show that the limits of the turbulent space-time scaling of rainfall is limited in the range 30 min - 2 hours, and for spatial scales lower than 400 km². This temporal limit is explained by the increasing effect of the advection component with the increase of the accumulation time. The theory relating the spatial and temporal scales in a single scaling framework is referred to as “Frozen Turbulence” and it is detailed in the next section.

5.10.1 The “Frozen Turbulence” hypothesis

The space-time scaling as interpreted by *Zawadzki* (1973); *Gupta and Waymire* (1987); *Marsan et al.* (1996); *Venugopal et al.* (1999); *Deidda* (2000) corresponds to the Taylor hypothesis of frozen turbulence (*Taylor*, 1938). It consists in reinterpreting the temporal variations at a fixed location as spatial variations. This concept has been widely applied in turbulence to transform a time sequence of spatial measures into a 3-dimensional homogeneous and isotropic process, where a measure on the third axis λ corresponds to the time $\tau = \lambda/U$, where U is the large-scale advection velocity, supposed constant at all scales.

According to *Gupta and Waymire* (1987), a spatially homogeneous and temporally stationary random field $I(x,y,t)$, having finite second moment, satisfies Taylor’s hypothesis if the space-time variogram

$$\gamma_I(\tau, \mathbf{r}) = E[I(t, \mathbf{x}, I(t + \tau, \mathbf{x} + \mathbf{r})] \quad (5.37)$$

respects the following condition:

$$\gamma_I(\tau, \mathbf{0}) = \gamma_I(0, \mathbf{u}\tau) \quad (5.38)$$

where \mathbf{u} is a velocity vector, \mathbf{x} is a d-dimensional spatial coordinate and \mathbf{r} is a spatial lag.

The application of Taylor’s hypothesis in Equation 5.38 and regarding increments, can be extended to probability distributions (*Lovejoy and Mandelbrot*, 1985; *Bras and Rodriguez-Iturbe*, 1976), giving

$$I(t, \mathbf{x}) = V(\mathbf{x} - \mathbf{u}t) \quad (5.39)$$

where $V(\mathbf{x})$ is a spatially homogeneous random field. Equation 5.39 is referred to as the “frozen field model” for $I(t, \mathbf{x})$.

As highlighted by *Deidda* (2000), a more general formulation of the Taylor hypothesis is required in case of self-affine processes. This is the behavior we can expect in natural processes and consists of a scale-dependent velocity parameter $U_\lambda \sim \lambda^\theta$ used to rescale the time variable. In fully-developed turbulence, the velocity of a system decreases with the scale with an exponent $\theta \sim 1/3$.

5.10.2 The concept of dynamic scaling

The verification of the Taylor’s frozen-field hypothesis on empirical observations leads to the determination of the parameter relating spatial and temporal scales (*Venugopal et al.*, 1999). This parameter is usually denominated “dynamic scaling” parameter.

Assuming that the rainfall field is a multiplicative process, *Venugopal et al.* (1999) propose to analyze the fluctuations of the logarithm of the field $\Delta \ln I_{i,j,\tau}$, defined as

$$\Delta \ln I_{i,j,\tau}(L, t) = \ln(I_{i,j}^L(t + \Delta t)) - \ln(I_{i,j}^L(t)) \quad (5.40)$$

where I is the rainfall intensity at the spatial coordinates i, j , time instant t and spatial scale L . In this way the variable is approximately gaussian and independent of the background intensity.

The field $\Delta \ln I$ is studied at several spatial scales L and time lags Δt . *Venugopal et al.* (1999) study the statistical distribution of $\Delta \ln I$, which is centered at 0 being fluctuations. Therefore, the study of the statistical distribution $\Delta \ln I$ is limited to the study of its standard deviation.

The results of these studies are important: even if time-scaling and space-scaling do not hold by themselves, the standard deviation of $\Delta \ln I$ varies with the scale of analysis; $\Delta \ln I$ is constant for each couple of surface and accumulation period respecting the dynamic scaling rule $A/\Delta t^z$, where A is the surface, with z approximately estimated in the range 0.51-0.58.

The approach of *Venugopal et al.* (1999) suggests a discussion about the relations between accumulation time and time lag. This approach is not exactly space-time in the sense that the time is mainly seen as a time lag, while space is seen as an integration surface. A pure space-time study would have studied the temporal and spatial lag influence for the fluctuation fields, or the fluctuations as a function of the involved surface and accumulation duration.

In the first case, we would have seen that the decorrelation time of a storm is strictly dependent on the size of the integration window (the larger the window, the more correlated the fields in time), in agreement with *Bell* (1987) and *Bell et al.* (1990). This phenomenon can be related to the increase of the decorrelation time with the decrease of resolution that we observe for time series (for instance, in our region hourly data are generally correlated up to 12 hours in average, daily data up to 48 hours). In the second case, we would have explored the multi-scaling character of rainfall in space-time, similarly to *Marsan et al.* (1996); *Deidda* (2000) or *De Michele and Bernardara* (2005).

These hybrid approaches give us information about the decorrelation in time of spatial data, suggesting how to disaggregate rainfall when one disposes of fields separated by a time lag.

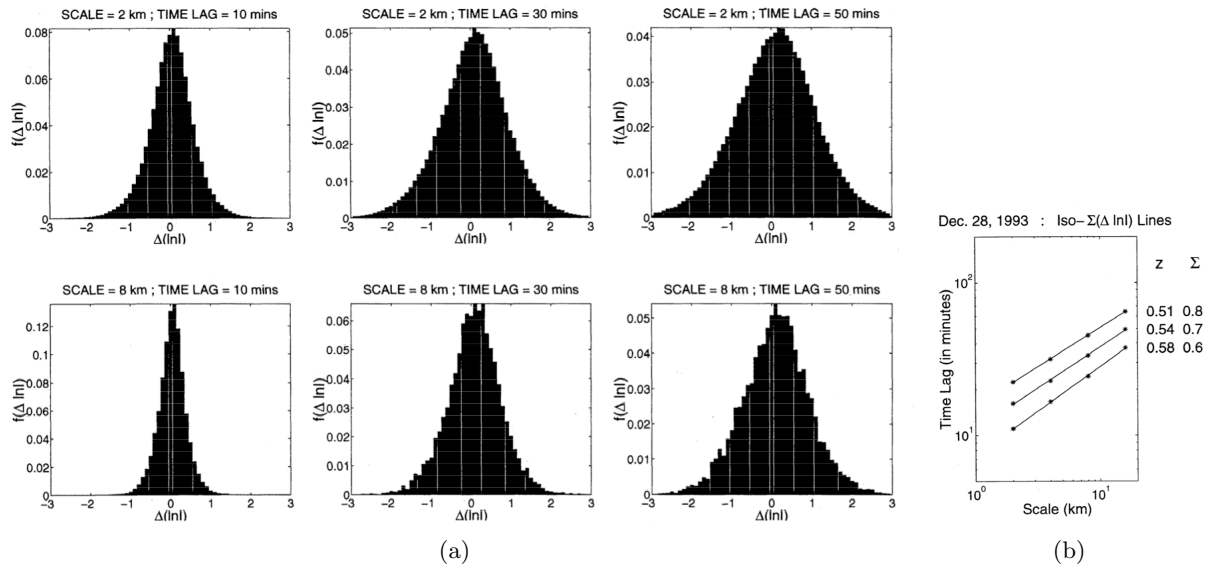


Figure 5.17: Dynamic scaling evaluation for the event of 1993, December the 28th, in Darwin, Australia. a: histograms of $\Delta \ln I$ for different spatial scales and time lag. b: Computation of the dynamic scaling parameter z for different iso-standard deviation lines.

The use of this method, except for the determination of the dynamic scaling z , is however limited to disaggregation scopes, differently from what has been observed for the *Perica and Foufoula-Georgiou* (1996) approach, for example.

Deidda (2000) supports the validity of the Taylor hypothesis in space-time rainfall fields and proposes a generalization to the self-affine case, where the scale-dependent velocity parameter is a function of the spatial scale ($U_\lambda \sim \lambda^H$). Among the proposed space-time models, the study presented by *Castro et al.* (2004) focused on the determination of the space-time singularity spectrum in a small-scale precipitation field. This approach will be discussed in the Section 6.2.4.3 relatively to the derivation of IDAF formulations. A space-time modeling based on the spectral analysis of rainfall fields has been presented by *De Michele and Bernardara* (2005), based on the concept of dynamic scaling. The approach is based on the fact that in the (x,y,z) domain, in case of isotropic random field, the 3D spectral density assumes the power-law form:

$$S(\omega_x, \omega_y, \omega_z) \propto \frac{1}{(\omega_x^2 + \omega_y^2 + \omega_z^2)^{\frac{\delta+1}{2}}} \quad (5.41)$$

where ω represents the frequency in a given dimension.

For an anisotropic field, when the Taylor's hypothesis is applied, we can define a dynamic scaling exponent and the expression becomes

$$S(\omega_x, \omega_y, \omega_z) \propto \frac{1}{(\alpha^2 \omega_x^{2z} + \beta^2 \omega_y^{2z} + \omega_z^2)^{\frac{\delta+1}{2}}} \quad (5.42)$$

where z is the dynamic scaling exponent.

5.11 Scale invariance of extreme point-rainfall: IDF curves

Intensity - Duration - Frequency curves relate the rainfall intensity I , the duration D , and the frequency of occurrence F otherwise expressed in terms of return period $T_R = \frac{1}{1-F}$. IDF curves are devoted to the estimation of hazardous rainfall or flows. Empirical IDF curves are derived by fitting rainfall intensities relative to the same return period as a function of the temporal scale. Following this methodology, *Bernard* (1932) showed that IDF relationships have similar behavior in different regions of the world. Since then, IDF expressions based on empirical analysis were adopted and used as a tool for the engineering design.

More than sixty years later, *Koutsoyiannis et al.* (1998) showed that all the different formulation proposed to describe IDF curves could be generalized through the expression

$$I = \frac{\omega}{(d^\nu + \theta)^\eta} \quad (5.43)$$

where ω , ν , θ and η are non-negative coefficients with $\nu, \eta \leq 1$. Some of these variables have mutual dependencies and restrictions, leading to a simpler formulation of IDF curves:

$$I = \frac{a(T_R)}{b(D)} \quad (5.44)$$

where T_R is the return period and D the considered duration. $b(d)$ is equal to $(D + \theta)^\eta$.

IDF curves seem to present an universal behaviour: whatever the location, extreme rainfall for a given return period is linear in double logarithmic plot for a given range of durations, approximately from 1-2 hours to 7-10 days. Figure 5.18-a presents DDF (Depth-Duration-Frequency) curves in linear scale, showing that rainfall depth increases with duration. In Figure 5.18-b we report IDF curves for Montpellier for different return periods, showing the linear behavior Figure 5.18-b in the range 2h - 7 days, and a non-linear decay for durations lower than the hour.

The linearity in log-log corresponds to a power-law behavior of IDF curves for a given return period. Usually IDF curves in the linear range are defined as a function of two parameters a and n , both dependent on the return period T_R :

$$I_{D,T_R} = a(T_R)D^{n(T_R)} \quad (5.45)$$

where the parameters $a(T_R)$ and $n(T_R)$ only depend on the return period T_R . Equation 5.45 expresses that rainfall intensity decreases when duration increases, for a fixed return-period.

Koutsoyiannis et al. (1998) provided a mathematical derivation of IDF relationships using dimensional arguments. To obtain analytical IDF expressions for any return period, the scaling properties of rainfall intensity must be coupled to the density function for rainfall extremes (block-maxima or peaks-over-threshold). This allowed to provide a series of definition for IDF as a function of the distribution chosen to model the heavy rainfall observations. The IDF formulations for Gumbel, GEV, Gamma, Log Pearson III, Log-normal and exponential distributions are provided.

The work of *Koutsoyiannis et al.* (1998) is insightful in the sense that, for the first time, a coupling of the IDF empirical model to the extreme value theory is proposed, with the aim to provide IDF expressions valid for the temporal validity range and for any return period.

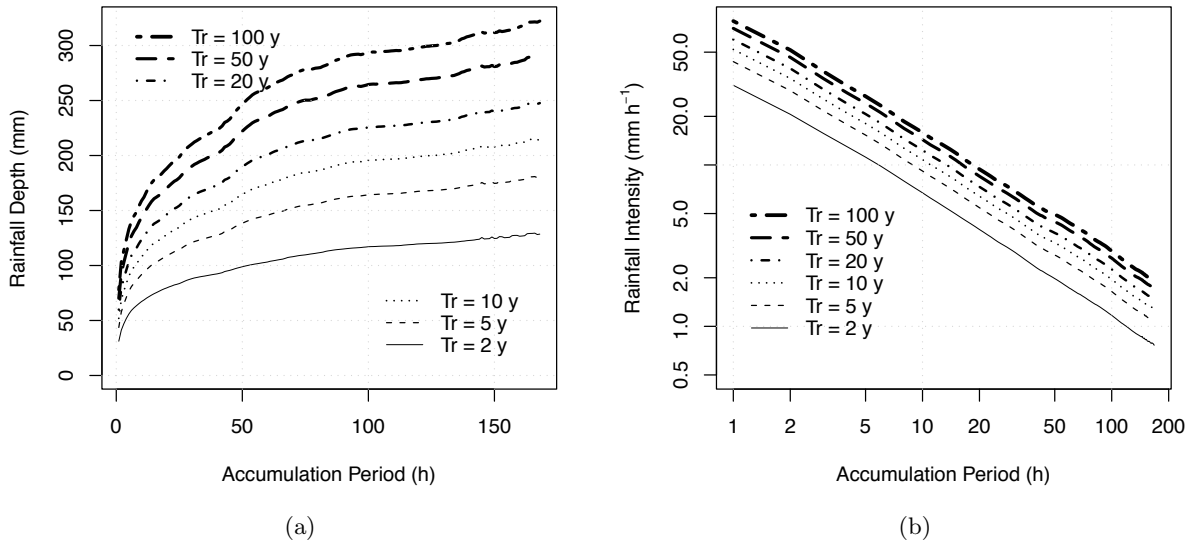


Figure 5.18: Rain gauge station of Montpellier, Southern France. a: Depth - Duration - Frequency curves. b: Intensity - Duration - Frequency curves. The data is relative to the period 1920-1972.

5.11.1 IDF scaling

The work of *Koutsoyiannis et al.* (1998) has been completed in the following years by numerous studies describing the scaling properties of IDF curves as a function of the different way in which heavy rainfall observations are sampled and modeled.

Rosso and Burlando (1990) showed that IDF formulation are implicit forms of scale-invariant relations. *Burlando and Rosso* (1996) presented a annual-maxima IDF model based on the log-normal distribution of rainfall maxima. Even though log-normal distribution does not belong to the Generalized Extreme Value distribution class for extremes, its difference between the Gumbel distribution is limited and therefore the distribution can be suitable for low return periods. With respect to the Gumbel distribution, log-normal has the advantage that it can accommodate multi-scaling in the sense of *Gupta and Waymire* (1990). A schematic plot of simple scaling and multi-scaling IDF curves is shown in Figure 5.19. On the other hand, Gumbel distribution is limited to simple-scaling scale as needed in multi-scaling. For similar reasons, the GEV distribution as well is not suitable for multiple scaling.

Bendjoudi et al. (1997) interpreted IDF curves in a multi-fractal sense. Based on the concept of Universal Multi-fractal (Section 5.5.2.3), they demonstrate that the classical IDF formulation of the type:

$$I_{T_R} = \frac{KT_R^m}{D^n} \quad (5.46)$$

where KT_R^m corresponds to $a(T_R)$ in Equation 5.43, being a power-law relation can be studied under the point of view of multifractality.

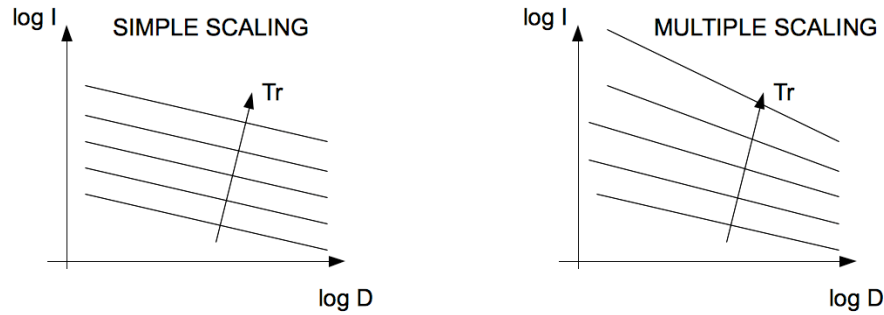


Figure 5.19: Schematic representation of simple and multiple scaling IDF curves. Each curve represents the rainfall intensity I as a function of the accumulation duration D for a specific return period T_R .

The advantage with respect to the IDF formulations for maxima is that, in this context, the whole sample is considered, providing an explanation of the relationship between common and heavy rainfall.

By performing a multi-fractal analysis on the rainfall series (by using Trace Moments or Double Trace Moments method, Section 5.7.1.1 and 5.7.1.2 respectively), the multi-fractal parameters C_1 , α_s and the non-conservativity Hurst exponent H can be computed.

For moments $q > q_D$ a “multi-fractal phase transition” (Schertzer and Lovejoy, 1992) cause the moment scaling function $K(q)$ to be linear with the moment order q .

After mathematical development, Bendjoudi *et al.* (1997) derives multi-fractal IDF curves as:

$$\ln(I_{T_R}) = \frac{1}{q_D} \ln(T_R) - \ln(D) + c \quad (5.47)$$

where c is a constant and q_D the minimum moment order for which moment divergence is obtained.

Menabde *et al.* (1999) show that simple-scaling of rainfall intensity maxima holds in temporal ranges of hydrologic interest. Their approach associates the IDF concept described in Koutsoyiannis *et al.* (1998) with the GEV-I (Gumbel) framework, in a context of scale-invariance.

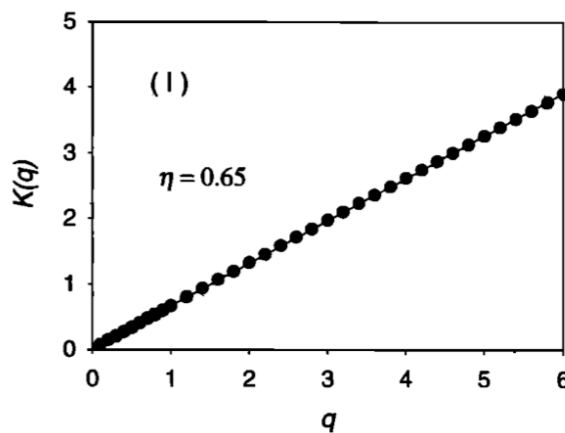


Figure 5.20: Simple scaling behavior of rainfall maxima at Melbourne, Australia in the range 30 min - 48 hours. From (Menabde *et al.*, 1999)

Let us assume that rainfall maxima follow EV-I (Gumbel) distribution and feature temporal simple-scaling. The scaling exponent $K(q)$ of the statistical raw moments (Equation 5.14) is a linear function of the moment order q :

$$K(q) = H \cdot q \quad (5.48)$$

with constant H . This behavior has been verified in *Menabde et al.* (1999) on two significantly different rain gauges in Australia and South Africa.

The Gumbel distribution is defined by two parameters, the location μ and the scale σ . As stated in *Menabde et al.* (1999) and *Borga et al.* (2005), the strict sense scaling relationship (Equation 5.14) allows to determine the distribution parameters at any duration D from a first guess at a reference duration D_{ref} :

$$\mu_D = \left(\frac{D}{D_{ref}} \right)^{-K(1)} \mu_{D_{ref}} \quad (5.49)$$

$$\sigma_D = \left(\frac{D}{D_{ref}} \right)^{\frac{-K(2)}{2}} \sigma_{D_{ref}} \quad (5.50)$$

In case of simple scaling of rainfall maxima (Equation 5.48), $K(1) = K(2)/2$ since $K(q)$ is linear with null intercept. From a practical point of view, the Gumbel parameters μ and σ are estimated thanks to the sample mean $E[x]$ and standard deviation $s[x]$ following the relations:

$$\hat{\mu} = E[x] - \gamma \hat{\sigma} \quad \text{and} \quad \hat{\sigma} = \frac{s[x] \sqrt{6}}{\pi} \quad (5.51)$$

where $\gamma \sim 0.5772$ is the Euler constant.

Combining equations 5.50 and 5.51, the knowledge of the mean and standard deviation of maxima at a particular scale is sufficient to estimate the Gumbel parameters at any duration.

In particular, if simple scaling holds, the maximum rainfall intensity at any duration D and return period T_R can be expressed by (*Menabde et al.*, 1999):

$$I_{D,T_R} = \frac{\mu(D_{ref}) - \sigma(D_{ref})\varepsilon}{\left(\frac{D}{D_{ref}} \right)^{-n}} \quad (5.52)$$

where $n = K(1) = K(2)/2$.

Borga et al. (2005) adopt the same scale-invariant Gumbel framework to build regional maps of the extreme rainfall behaviour. A similar scaling model for annual maxima has been used by *Gerold and Watkins* (2005). A GEV scale-invariant model is proposed in Section 8.

5.12 Application of Multiplicative Cascades

The downscaling techniques allow to reproduce the fine-scale variability of a random field starting from a large scale information. These techniques are used in hydro-meteorology in at least three applications:

- increase the resolution of meteorological models. Due to computational constraints, the operational meteorological models provide precipitation forecasts on scales of about 100 km^2 and few hours;
- disaggregate the rainfall satellite estimation, limited in spatial and temporal resolution;
- reproduce the variability of the rainfall observed by ground measurements for the hydrological simulations. A deterministic hydrological simulation running based on average rainfall depth do not provide reliable results due to the non-linearities of the basin response. This is one of the aims of the MEDUP project.

5.12.1 Other downscaling techniques

A possible approach for filling in the scale gap between operational needs and resolution of the measurements is based on the use of stochastic downscaling models. Downscaling is based on the implementation of a stochastic disaggregation algorithm able to reproduce some required features of rainfall distributions such as its actual small-scale variability. The fine-scale distribution has to be consistent with the known statistical properties of the small-scale rainfall distribution. The field resulting from disaggregation is not the deterministic result of a physical downscaling, but rather a realization of a statistically consistent rainfall field. Statistical downscaling models generate ensemble of realizations at a relatively low computational cost, leading to the possibility of ensemble prediction.

The stochastic rainfall models are generally grouped in three categories:

- i. individual precipitation cells;
- ii. auto-regressive processes (geostatistical simulation);
- iii. multiplicative cascades;

Individual Precipitation Cells: The earlier models of spatio-temporal precipitation downscaling were based on stochastic point processes, following the fundamental work of *Le Cam* (1961). They were based on the generation of individual rain cells, characterized by a given spatial structure and a Poisson-distributed arrival time. Studying a series of apparently dissimilar storms, *Austin and Houze* (1972) observed that they were composed of clearly definable patterns of precipitation area. Four sizes of precipitation areas have been detected, from the synoptic to the small-scale cells (Figure 5.21). Based on the observations of *Austin and Houze* (1972), *Waymire et al.* (1984) presented a stochastic model constituted by four embedded levels of precipitation areas. *Rodriguez-Iturbe et al.* (1986) investigated the total rainfall depth generated at a point by simulating storms whose maximum intensity was exponentially distributed and whose inter-arrival time was Poisson-distributed or clustered. *Eagleson et al.* (1987) compared three stochastic models showing the utility of point process models for storms that are stationary in space. *Northrop* (1998) simulates rain-fields assuming elliptical rain cells and Poisson arrival time. Similar spatial rainfall simulators are used by *Wheater et al.* (2000), *Willems* (2001) and *Cowpertwait et al.* (2002).

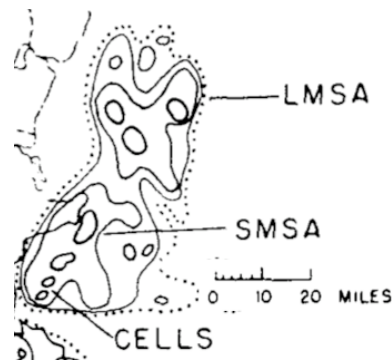


Figure 5.21: Precipitation patterns: Large Meso-scale Areas (LMSA), small meso-scale areas (SMSA) and cells. Taken from *Austin and Houze (1972)*.

Geostatistical Simulations: The geostatistical simulations (or auto-regressive processes) yield stochastic rainfall fields featuring a given spatial correlation. This can be done i) directly, by imposing a correlogram function or a variogram function and then applying a geostatistical simulation method such as Turning Bands Methods (*Matheron, 1973; Mantoglou and Wilson, 1982; Wood and Chan, 1995*); ii) indirectly, by applying an inverse Fourier transform to a 2D spectrum with a given amplitude distribution (circulant embedding methods, *Dietrich and Newsam (1993); Chan and Wood (1997)*). In both cases, the methods provide linearly correlated Gaussian fields. To transform Gaussian fields into the desired distribution, appropriate static nonlinear transformation can be used (anamorphosis). Since derived by a transformation of Gaussian fields, the obtained fields are called meta-Gaussian³. An example of use of meta-Gaussian geostatistical simulations is given in *Guillot and Lebel (1999)*. In case of space-time simulations, the earlier separable space-time models (see *Kyriakidis and Journel (1999)* for a review) were not able to model the space-time interactions. Non-separable space-time models must respect additional constraints. *Stein (2005)* overviews the space-time models proposing a new class of space-time covariance models.

5.12.2 Multiplicative Cascades

The third main category of statistical downscaling techniques is known as multiplicative cascades. Multiplicative cascade models have been proposed to describe the variability of natural processes exhibiting scale-invariance. These models have been used in disparate fields, such as turbulence (e.g. *Kolmogorov (1941); Mandelbrot (1974); Frisch and Parisi (1985); Meneveau and Sreenivasan (1987)*), internet traffic (*Feldmann et al., 1998*), stock prices (*Mandelbrot, 1997*), river flow (*Gupta and Waymire, 1990*) and rainfall (*Over and Gupta, 1996; Olsson, 1998; Menabde and Sivapalan, 2000; Güntner et al., 2001; Veneziano et al., 2002; Badas et al., 2006; Gaume et al., 2007; Rupp et al., 2009*). They consists of splitting each large-scale observation in b sub-grid observations, where b is the multiplicity of the cascade. Each sub-grid observation is actually obtained by multiplying the large scale-observation by a random coefficient, called “cascade weight”. The random weights must respect properties concerning the mass conservation (the weights sum to 1), and probability distribution. In the following paragraphs we provide an overview of the main features of Multiplicative Cascades.

³In Appendix E an example of log-normal field generation is reported based on the package `RandomFields` of R.

5.12.3 Bare and dressed quantities

By these colorful adjectives, *Schertzer and Lovejoy* (1987) depicted the two possible results of cascades. The “bare” and “dressed” quantities originate from consistently different concepts. The bare quantities are the result of a pure disaggregation process (Figure 5.22-a shows dyadic disaggregation scheme for generating bare quantities). Assuming that natural processes are originated from multiplicative cascades, observation of natural phenomena corresponds to dressed quantities, i.e. the result of the aggregation of finer-scale processes at the instrument resolution. The output of multiplicative cascades can be dressed quantities if an aggregation process follows the disaggregation process, as shown in Figure 5.22-b. In these examples, the k_i observations are split through multiplication by the weights $\omega_i^{(m)}$, where m is the cascade step.

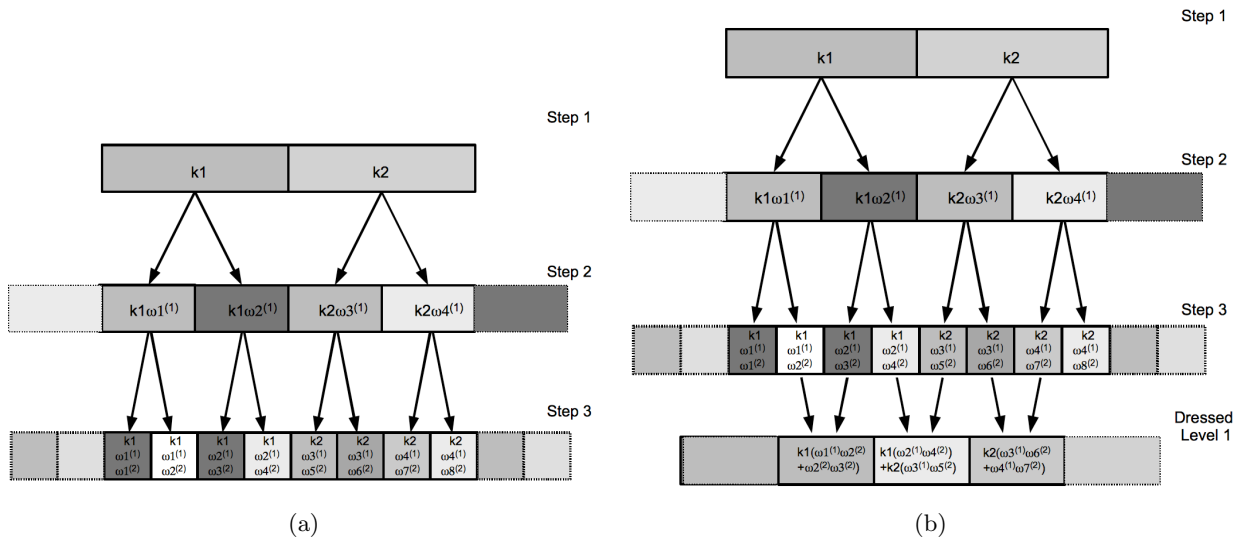


Figure 5.22: Schematic of bare cascades (a) and dressed cascades (b) in a dyadic cascade scheme.

5.12.4 Canonical and Micro-canonical cascades

The difference between canonical and micro-canonical cascades is the way in which the mass is conserved through the scales: at each level, the expected value of the sum of weights is equal to unit in canonical cascades (i.e. the overall mass is conserved, but the local aggregations will not correspond to the original values). This scheme is typically used in prediction, where initial mass acts as an initialization value and the aim is to obtain a statistically correct behavior at the finer scales. The canonical cascade scheme is depicted in Figure 5.23-a, where one can see that no constraint is imposed for the choice of the weights except the overall unit average.

For the disaggregation of historical series, when the mass should be conserved throughout the scales, micro-canonical cascades are used, in which the weights are chosen such that their sum is equal to the unit. In dyadic cascades, it can be written as $\omega_1 + \omega_2 = 1$. The micro-canonical cascade scheme is represented in Figure 5.23-b. The advantage of microcanonical cascades, as reported by (*Mandelbrot, 1974*), is the moments convergence.

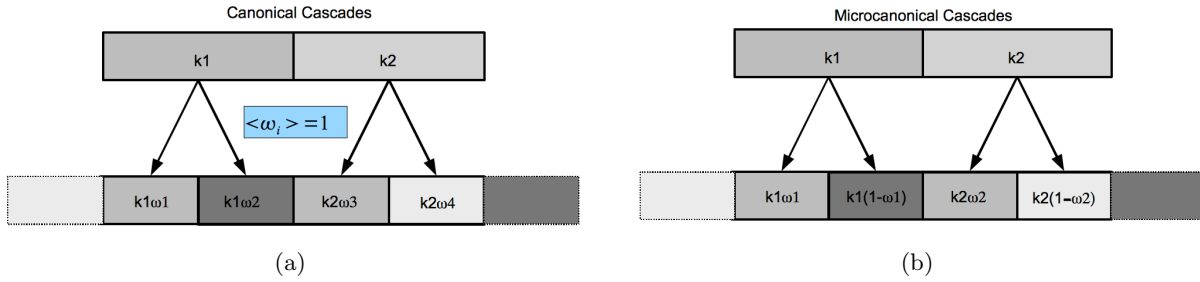


Figure 5.23: Schematic of canonical and micro-canonical cascades.

5.12.5 Discrete and continuous cascades

The multiplicative process is usually imagined as a discrete cascade. At each disaggregation step, the number of element in which original observations are split is equal to the multiplicity level m . In general, when $m \geq 2$, the cascade is discrete. Dyadic cascades have $m = 2$. The first cascade models proposed by *Gupta and Waymire (1993)*; *Over and Gupta (1996)* were discrete. They assume that the cascade multiplicity does not affect the behavior of the cascade after a large number of steps. *Schertzer et al. (1995)* proposed a continuous cascade process, i.e. the multiplicity of the cascade m tends to 1. The number of steps is densified, reaching a continuum of scales. The real advantage of this kind of disaggregation is to remove the spurious correlation encountered for small lags and highlighted by *Carsteanu and Foufoula-Georgiou (1996)* as a consequence of the independent weight choice. The drawbacks of continuous cascades are their complexity and the limitation to canonical processes: fixing an arbitrary multiplicity $m \rightarrow 1$, it is not possible to exactly split an observation in m observations, therefore continuous cascades can only be canonical.

5.12.6 Cascade Implementation

Since the cascade weights distribution has influence in the fine-scale distribution of the disaggregated field, the implementation of the cascade requires the choice of an appropriate weights distribution, respecting the actual rainfall distribution. *Gupta and Waymire (1993)* proposed cascade models by choosing a number of different distributions. For the application of multiplicative cascades, only distributions satisfying the following conditions can be adopted:

- i. the weights distribution must have unit mean, to respect the mass conservation;
- ii. the distribution must be close to the variable distribution;
- iii. the distribution should be stable or (relaxed case) infinite-divisible.

The Gaussian distribution is stable and it is the most frequent distribution encountered in nature. However, many natural processes (among them, rainfall) do not follow Gaussian distribution. The accuracy of the operational point measurement devices (e.g. rain gauges, see Section 2) does not help in detecting the actual distribution of rainfall values. Depending on the scale and on the climatology of the site, the rainfall distribution can be closer to an exponential (*Todorovic and Woolhiser, 1975*),

to a 2-parameters Gamma (*Buishand, 1978*), to a mixed exponential (*Chapman, 1998*) rather than to a Weibull mixture (*Suhaila and Jemain, 2007*).

The recent introduction of radar imagery for the quantitative estimation of rainfall can be useful for the analysis of spatial rainfall. Thanks to the GATE program, *Kedem and Chiu (1987)* were able to study radar-measured rainfall fields. They verified that rainfall-rates approximately follow a log-normal distribution at spatial scales ranging from 16 to 1600 km^2 .

This finding has large importance for the interpretation of the mechanisms involved in the rainfall generation: log-normal distribution possesses intriguing properties. In particular, the conservation of log-normality of a variable through a range of scales is the indication that the process is multiplicative. Additive processes are generated by the sum of the effect of various fluxes, while multiplicative processes are generated by the product of the various fluxes (e.g. energy) at various scales. Multiplicative processes that conserve the probability distribution throughout the scales are self-similar, i.e. the distribution of the rainfall at one scale can be derived from the distribution at another scale by means of scaling relationships.

To be scale-invariant and additive, a process should follow, at all scales, a stable-upon-addition distribution. A random variable has stable distribution if the linear combination of n independent copies of the variable X preserves the same distribution:

$$X_1 + X_2 + \dots + X_n \stackrel{d}{=} c_n X + d_n \quad (5.53)$$

where c_n and d_n are two constants, and $c_n = n^{1/\alpha_s}$ where α_s is the Levy's stable parameter, defined in the range $0 < \alpha_s \leq 2$. The symbol $\stackrel{d}{=}$ means equality in distributions. Stability upon addition is a restrictive property. The Levy's Stable is the only distribution class possessing this feature. The normal distribution is the simplest case of Levy's stable process and it is defined as

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (5.54)$$

where μ and σ are the mean and the variance of the distribution, respectively.

The stable distribution, featuring heavy tails, has no explicit form but is defined by its characteristic function⁴ (*Nolan, 2009*):

$$f(x, \alpha_s, \beta_s, c_s, \mu_s) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp[it\mu_s - |c_s t|^\alpha_s (1 - i\beta_s \operatorname{sgn}(t)\phi_s)] \quad (5.55)$$

In Equation 5.55, $\phi_s = \tan(\pi\alpha_s/2)$ when $\alpha_s \neq 1$ and $\phi_s = -(2/\pi)\log|t|$ when $\alpha_s = 1$. The α_s parameter is the Levy's stability index, expressing the heavy-tail behavior of the distribution. The distribution hyperbolically decays as $1/\alpha_s$ for large x except when $\alpha_s = 2$ (that is the normal distribution). The case $\alpha_s = 1$ is the well-known Cauchy distribution, widely used in physics.

The stability-upon-multiplication is verified when the product of n independent copies of the variable X preserves the same distribution:

⁴The characteristic function provides an alternative way for describing a random variable. In case the variable admits a density function, the characteristic function is the Fourier transform of the density function. Differently from the density function, the characteristic function of a distribution always exists. For a random variable X , it is defined as $\phi_X(t) = E[e^{itX}]$

$$Y_1 \cdot Y_2 \cdot \dots \cdot Y_n \stackrel{d}{=} Y^{e_n} + \exp(f_n) \quad (5.56)$$

It is easy to verify that if Y_n is a function of the type $\exp(X)$, one can take the logarithms obtaining

$$X_1 + X_2 + \dots + X_n \stackrel{d}{=} e_n X + f_n \quad (5.57)$$

where e_n and f_n are two constants. A log-stable distribution can be thus derived from the exponentiation of a stable distribution. The extension of the classical central limit states that the normed product of a set of random variables (characterized by having finite variance) will asymptotically approach a log-stable distribution as the number of variables increases.

The semilogarithmic plot in Figure 5.24-a shows that stable distributions have heavier tails compared to Gaussian distributions, i.e. very large extremes are relatively frequent. If we represent a random walk, a Levy walk results in extremely large jumps compared to the Gaussian case (Figure 5.24-b).

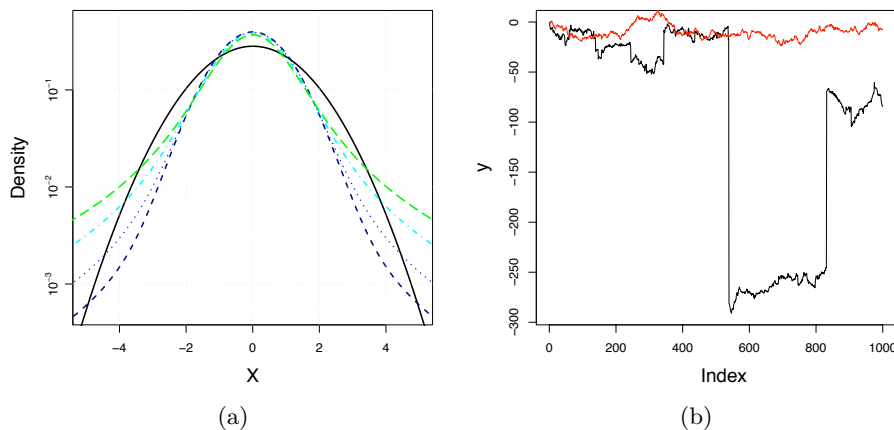


Figure 5.24: Comparison Gaussian versus max-stable distributions. a: Density plot of max-stable distribution for various α_s values, $\beta_s = 0$ and $c_s = 1$. Legend: solid black, $\alpha_s = 2$ (Gaussian distribution), dashed dark blue, $\alpha_s = 1.9$, dotted blue, $\alpha_s = 1.8$, dash-dotted cyan, $\alpha_s = 1.6$, long-dashed green $\alpha_s = 1.4$. b: 1D random walk for the Gaussian case ($\alpha_s = 2$), red line, and for a max-stable process with $\alpha_s = 1.4$, black line.

The stability is a restrictive property. Moreover, issues concerning the use of Levy's stable distribution arise. In particular, the variance diverges for any Levy's stable distributed variable. This means that, due to the "wild" character of such series, whatever the sample length a new observation can sensibly modify the second moment of the distribution. Another indicator of the "wildness" of this distribution concerns the divergence of the mean for $\alpha < 1$. The character of such series is extremely singular, meaning that the use of Levy's stable series is limited. The estimate of the Levy's stable parameters is awkward, and few reliable estimator techniques seem to be available. Among these, the maximum-likelihood function proposed by *Nolan* (2001).

Infinite divisibility can be suitable for scale-invariance modeling (*Waymire and Gupta*, 1981). It is a relaxed property compared to stability-upon-multiplication. A random variable X is said to be

infinitely divisible if for every $n \in \mathbb{N}$ the following equality in distribution is verified (*Steutel and Van Harn, 2003*):

$$X \stackrel{d}{=} X_{n,1} + \dots + X_{n,n} \quad (5.58)$$

where $X_{n,1}, \dots, X_{n,n}$ are independent with $X_{n,j} \stackrel{d}{=} X_n$ for all j and some X_n , the n -th order factor of X .

Distributions exhibiting infinite divisibility are the Poisson, the exponential, the negative binomial, the geometric, the Gamma, the Student's t distributions, besides normal and stable distribution families. Some of them have been used in geophysical applications (*She and Waymire, 1995; Deidda, 2000*), even though with ambiguous results (*Schertzer et al., 1995*).

5.12.7 Implementation of the multiplicative cascades

Gaussian and Levy's stable cascades The mathematical properties of Levy's stable distributions have been widely studied in *Zolotarev (1986)* and in *Samorodnitsky and Taqqu (1994)*. *Veneziano and Langousis (2005)* propose a parametrization suitable to express the parameters of the Levy's stable generator as a function of the multi-fractal parameters C_1 and α_s . If the cascade has multiplicity m , the generator of log-stable cascade must respect the following parametrization:

$$\begin{cases} \alpha_s = \alpha_s \\ \beta = -1 \\ \mu = \frac{C_1}{1-\alpha_s} \\ \sigma = C_1^{1/\alpha_s} [\ln(m)]^{(1-\alpha_s)/\alpha_s} \left[\frac{\cos(\frac{\pi\alpha_s}{2})}{1-\alpha_s} \right]^{1/\alpha_s} \end{cases} \quad (5.59)$$

when $\alpha_s \neq 1$. The β parameter is fixed to -1 in cascades.

In the particular case $\alpha_s = 1$, the μ and σ parameters become

$$\begin{cases} \mu = -C_1 \ln \ln(m) \\ \sigma = \frac{\pi C_1}{2} \end{cases} \quad (5.60)$$

In case of normal generators ($\alpha_s = 2$) the parameters are:

$$\begin{cases} \mu = -C_1 \\ \sigma = 2C_1 \ln(m) \end{cases} \quad (5.61)$$

In the Gaussian case, β has no influence while σ corresponds to the standard deviation of the distribution.

Poisson cascade As reported in Section 5.12.6, Poisson distribution is not stable but it possesses a relaxed property: the infinite divisibility. *She and Waymire (1995)* and *Deidda (2000)* generated Poisson cascades. The Poisson random generator η is:

$$\eta = e^A \beta^y \quad (5.62)$$

where A and β are constant parameters and y is a Poisson distributed random variable. A poissonian variable y is defined as a function of Λ , that is the expected number of occurrences in an interval. The probability that there are exactly m occurrences (m being a non-negative integer, $m = 0, 1, 2, \dots$) is equal to

$$p(m = \Lambda) = \frac{\Lambda^m e^{-\Lambda}}{m!} \quad (5.63)$$

where e is the base of the natural logarithm ($e \sim 2.71828$), $m!$ is the factorial of k and Λ is a positive real number, equal to the expected number of occurrences that occur during the given interval.

5.13 Conclusion

The rainfall generation process is a complex atmospheric phenomenon whose multi-scale mechanisms are still partially unknown. The behavior of rainfall distributions at different scales is still an open subject and different answers have been addressed to this problem in literature. The fact that rainfall is the result of the action of several atmospheric phenomena pushes to find analogies with the small-scale turbulence, for which we have seen that the power-spectrum is log-linear, meaning that no characteristic scale of turbulence exists. The scale-invariant approaches offer the possibility to guess the distribution of a random field at scales that differ from the observation scale.

The main objective of scale-invariant analysis is to determine the singularity spectrum, i.e. the function allowing to rescale each quantile of a statistical distribution. Two main approaches are available: i) the Generalized Structure Function approach, consisting in analyzing the p^{th} power of the fluctuations of a field at increasing lags l , and in determining the scaling function $\zeta(p)$; ii) the moment analysis, consisting in analyzing the q th power of a subsequently degraded field and in finding the scaling function $\tau(q)$ or $K(q)$ that defines the moment scaling. We have seen that the two approaches (i) and (ii) may lead to the same singularity spectrum $D(h)$, and that therefore the two formalisms are substantially equivalent. On the other hand, depending on the kind of available data, one technique could be recommended respect to the other. In case of non-stationary fields, the Generalized Structure Function could be biased by the drift, similarly to what happens for the variographic analysis in Section 4.4.2. The moment scaling analysis is more general than GSF since the $\tau(q)$ function exactly corresponds to the Legendre transform of $D(h)$. The moment analysis can be strongly biased for large q by heavy observations, and for small q by the measure precision. A relatively new approach analyzes the scaling of a derived product of the wavelet analysis, the Wavelet Transform Modulus Maxima (WTMM), to compute the singularity spectrum. Due to the complexity of the method, few analyses in this sense have been done up to this moment.

The implementation of these methods allows to diagnose the scale invariant features of: i) spatial rainfall ii) temporal rainfall iii) space-time rainfall. One of the most interesting results has been the verification of the Taylor hypothesis of “frozen turbulence” on space-time rainfall fields, for limited spatial and temporal scales.

In conclusion, the state of the art on the rainfall scaling highlighted that:

- scale-invariance in space is usually verified for small scales (maximum 400 km^2) and therefore its application for the region of interest (16000 km^2) could be strongly limited; its application for the analysis of singular rainfall events may allow to obtain precipitation fields reproducing the fine-scale variability at resolutions higher than the observation scale;
- scale-invariance in time is usually found for singular rainfall events for durations lower than 5 minutes; long series often exhibit scale-invariance in the range 2h-15 days, which can be of interest for determining the behavior of high-resolution rainfall based on daily data.
- space-time rainfall is verified for singular storms and in temporal windows for which the stationarity of the mean and variance is ensured; fixing a dynamic scaling parameter, it is possible to treat the temporal coordinate as a third spatial coordinate, simplifying the parametrization and allowing to manage the space-time anisotropy; this behavior is however broken for temporal scales higher than 40 minutes and 100 km^2 .
- in many regions of the world the rainfall maxima seem to scale with a simple scaling relation; this property allows to build models coupling the laws of extremes with scale-invariance, providing a robust framework for the modeling of Intensity-Duration-Frequency (IDF) curves.

In the final part of the chapter, we introduced the stochastic modeling methods and in particular the scale-invariant stochastic modeling. We highlighted the advantages and drawbacks of the use of Levy's stable, Gaussian and Poisson cascade generators.

6 Spatial Rainfall Extremes

Résumé

*Le chapitre présente l'état de l'art des principales méthodes empiriques et semi-empiriques utilisées pour quantifier la relation entre les plus fortes intensités de pluie et la surface d'agrégation (**Areal Reduction Factors ou ARF**). La plupart de ces méthodes sont fondées sur le fait que la pluie était à l'origine mesurée par un réseau pluviométrique au sol. Ce type de mesure est encore aujourd'hui la méthode la plus fiable pour la mesurer. L'estimation quantitative de la pluie par radar (QPE), en effet, est aujourd'hui encore sujette à nombreuses sources d'incertitude. En général, toutes les méthodes pour définir les ARF suivent les mêmes contraintes: i) **en augmentant la surface d'agrégation, l'intensité du maximum baisse**; ii) **en augmentant la durée d'accumulation, le rapport entre pluie spatiale et pluie ponctuelle va vers 1**. En particulier, nous décrivons la méthode présentée par De Michele et al. (2001) qui met en relation l'échelle d'agrégation spatiale et temporelle par le biais d'un coefficient de "**dynamique d'échelle**". L'origine de ce concept vient de la physique des petites échelles (Zawadzki, 1973; Venugopal et al., 1999; Deidda, 2000) et à été adapté à des surfaces de 1 à plusieurs milliers km^2 et à des échelles de temps de l'ordre de 10 minutes au jour.*

6.1 Introduction

This chapter presents a state of the art of the main empirical and semi-empirical methods for quantifying the relations between heavy rainfall intensities and the aggregation surface (**Areal Reduction Factors** also known as **ARF**). All these methods originate from the fact that rainfall was originally measured by means of a ground measurement network; this is still, at the present time, the most reliable measurement of rainfall. The quantitative rainfall estimation by radar, in fact, is submitted to a series of uncertainty sources. In general, all the methods follow at least two constraints related to the rainfall behavior: **the larger the surface, the lower the rainfall intensity; the larger the accumulation period, the closer the areal maximum is to the point maximum**. In particular, we describe the method presented by De Michele et al. (2001) consisting in relating the spatial and temporal aggregation scales through a "**dynamic scaling**" ratio. This concept, originated from the small-scale physics (Zawadzki, 1973; Venugopal et al., 1999; Deidda, 2000), is adapted for surfaces and time scales of hydrologic interest (1 to 10000 km^2 and 10 min to 24 hours).

6.2 From point to spatial maxima: ARF

6.2.1 Background

The estimation of the frequency of spatial rainfall events is necessary in the engineering design: the design of hydraulic structures requires the estimation of the rainfall amount that can possibly fall over a catchment. Dealing with point-rainfall, the intensity of the design rainfall event is provided by the rainfall frequency analysis, such as IDF curves (Section 5.11). The occurrence frequency of a given rainfall amount over an area larger than the rain gauge sampling area ($\sim 1000\text{cm}^2$) is more difficult to estimate as the small-scale rainfall variability is not known and the point-rainfall network density is usually insufficient to obtain reliable estimates of the spatial rainfall amounts (Section 2.4). Several studies (*Obled et al. (1994)*; *Arnaud et al. (2002)*, among the others) highlighted that the knowledge of the average rainfall depth could be sufficient for hydrological purposes, the influence of rainfall spatial variability being of lower importance dealing with extreme events .

Therefore, a tool for the evaluation of the rainfall depth associated with a given return period falling over a surface can be sufficient for hydrological purposes. This motivates the creation of empirical curves referred to as Areal Reduction Factors.

Areal Reduction Factors, as defined by the Natural Environmental Research Council (*NERC, 1975*), are factors applied to point rainfall values characterized by a specified duration and return period to obtain the areal rainfall values characterized by the same duration and return period.

Two types of Areal Reduction Factors are commonly used: storm-centered ARFs, defined for concentric windows as the ratio of the average areal rainfall to the maximum point rainfall (the point center) for given storms; fixed-area ARFs, where the same ratio is computed scanning a surface with a moving-window. The example reported in Figure 6.1-a shows the spatial rainfall maxima of a given rainfall field recorded in a storm-centered and in fixed-area approach, respectively. While in the storm-centered approaches the spatial maximum locations are concentric to the point maximum, in the fixed-area approaches the spatial maxima are selected within a climatically homogeneous region. The ARF that derive from these two approaches can be consistently different, as schematically shown in Figure 6.1-a.

Storm centered implementation could be difficult, for example in presence of multi-cellular storms (*Asquith and Famiglietti, 2000*). Its statistical significance is questioned by *Omolayo (1993)*, since point and areal rainfall maxima are often generated by different storm types: point maxima are generated by deep convective events while areal maxima are consequence of stationary convective systems. *Sivapalan and Blöschl (1998)* argue that storm-centered ARFs are systematically smaller than fixed-area ARF: scanning the entire window for seeking the maximum, it is almost sure that a higher rainfall depth, not concentric to the storm, could be observed.

According to *Omolayo (1993)*, ARF can be calculated by dividing the maximum areal rainfall by the average T-year rainfall of the gauges within the same area:

$$ARF(A, D, T_r) = \frac{I_A(D, T_r)}{\frac{1}{\sum_i w_i} \sum_i w_i I_0(D, T_r)} \quad (6.1)$$

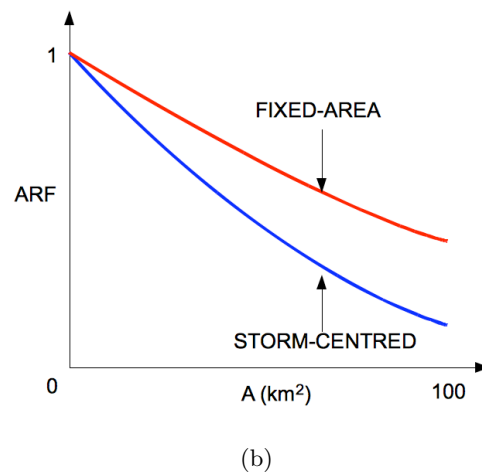
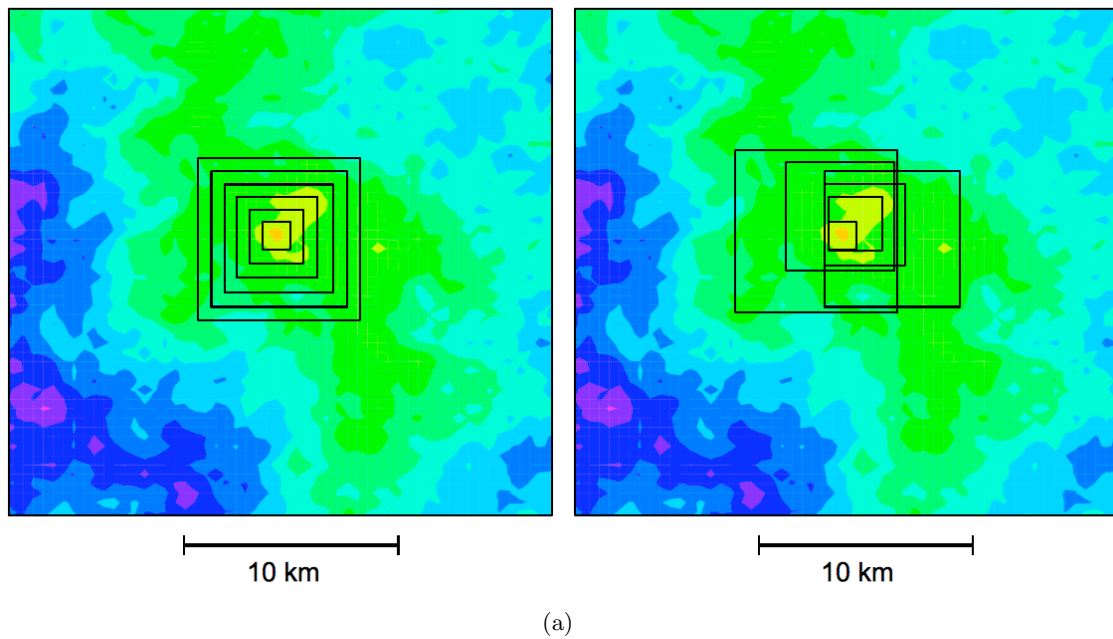


Figure 6.1: a: Maxima-selection process for a rainfall field with the aim of determining ARF curves. Left: storm-centered spatial rainfall selection. Right: fixed-area spatial rainfall selection. b: Comparison of the ARF curves obtained for the same field with the two methods.

The U.S. Weather bureau (*U.S. Weather Bureau*, 1958) developed an empirical method, disregarding the return period effect on the ARF, relating the mean of the annual maximum areal rainfall series to the mean of the annual maximum point measurements at all stations. Indicating with I_0 the point maxima and with I_A the concurrent point measurements that yield the spatial maximum, ARF are expressed, for a database of J years of observations at I rain gauges, as:

$$ARF(A, D) = \frac{\sum_j \sum_i w_i I_{A,ij}}{\sum_j \sum_i I_{0,ij}} \quad (6.2)$$

where $j = 1, \dots, J$ refers to the year and $i = 1, \dots, I$ is the rain gauge index. The weights w_i were originally determined by the Thiessen polygons method.

The fixed-area method adopted in the UK (*NERC*, 1975) considers the effect of the return period, for extreme events, of secondary importance for the ARF determination. For each region of area A and for each duration D , the ARF is computed according to the following expression

$$ARF(A, D) = \frac{1}{IJ} \sum_j \sum_i \frac{I_{A,ij}}{I_{0,ij}} \quad (6.3)$$

The method is a computational simplification of 6.2 and consists in evaluating the ARF as the average ratio between the areal rainfall maxima and average point rainfall. In the report, the ARF return period is indicated to be approximately 2-3 years (it depends on the law of extremes; the return period of the average annual value for a gumbel-distributed variable is ~ 2.35).

In Figure 6.2 two plots showing the ARF results for UK are shown. In Figure 6.2-b the ARF are multiplied by 100, expressing the percentage of the areal rainfall with respect to the point rainfall.

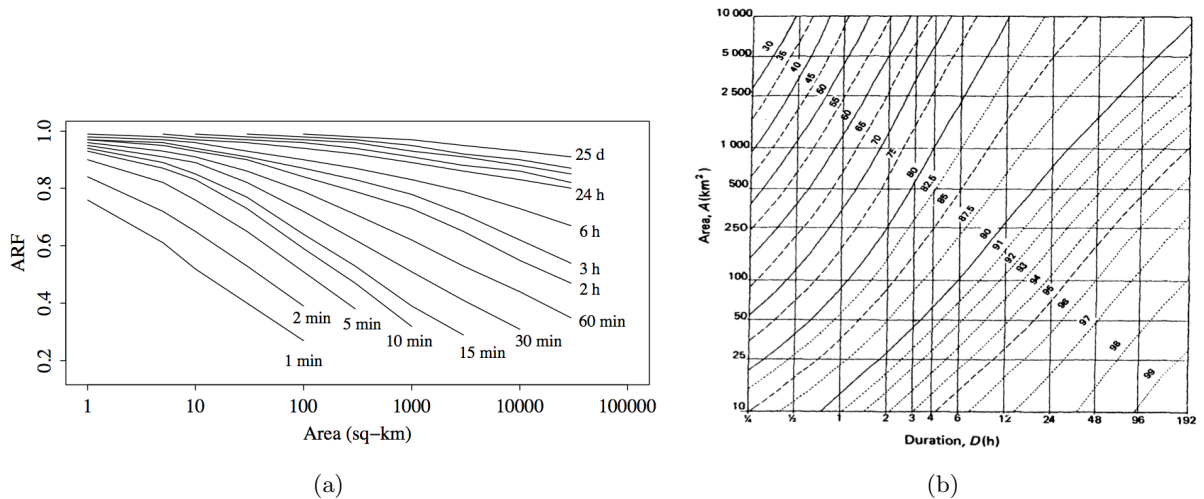


Figure 6.2: ARF curves according to *NERC* (1975), taken (a) from *Svensson* (2007) and (b) from *Langousis* (2005).

Figure 6.2-b shows that ARF are approximately constant for $D \sim A^{0.7}$. *Koutsoyiannis* (1997) derived an analytical expression for the NERC ARFs:

$$ARF(A, D) = \max \left(1 - \frac{0.048A^{0.36-0.01 \ln(A)}}{D^{0.35}}, 0.25 \right) \quad (6.4)$$

The *Bell* (1976) approach is the first fixed-area method in which the ARF are computed with respect of the return time. ARFs are computed as the ratio of the areal rainfall of return period T_R to the weighted average of the point rainfall of return period T_R :

$$ARF(A, T_R) = \frac{I_{A, T_R}}{\frac{1}{N} \sum_{i=1}^N w_i I_{i, T_R}} \quad (6.5)$$

The above-shown approaches are totally empirical. The computation of ARF curves can be done only at given durations and the semi-analytical expressions are derived in a totally empirical way. In the next sections we analyze the main methods proposed in literature for the determination of semi-empirical ARF curves.

6.2.2 Geostatistically-based ARF approaches

In this section we describe three approaches consisting in combining statistics of point-rainfall series and geostatistics of spatial fields with the aim to build semi-empirical ARF curves. The first contribution in this context has been provided by *Roche* (1963) who constructed the bivariate probability density function for pair of rain gauges separated by a given distance.

The basic hypothesis to use geostatistical approaches are: i) second-order stationarity of data; ii) the variance of the rainfall fields exists iii) point and spatial distribution of rainfall intensities belong to the same distribution class. Concerning the latter point, *Rodriguez-Iturbe and Mejía* (1974) assumed normally-distributed intensities, *Lebel and Laborde* (1988) derived ARF for Gumbel-distributed rainfall maxima, *Sivapalan and Blöschl* (1998) assumed exponentially distributed point rainfall and associated to it Gamma-distributed spatial rainfall.

Even though geostatistical approaches are attractive frameworks to combine the spatial structure of events with the point distribution of rainfall series, the resulting ARF relations are calibrated based on a correlation structure computed event-by-event, and therefore these methods can be classified as storm-centered. Storm-centered methods are of easy computation but they are i) not conservative (they systematically underestimate ARF compared to fixed-area methods) ii) not statistically significant (*Omolayo*, 1993).

6.2.2.1 Rodriguez-Iturbe and Mejía approach

Rodriguez-Iturbe and Mejía (1974) assumed that the rainfall field is a stationary Gaussian process. The stationarity implies that the areal rainfall average is equal to the average of point measurements; the Areal Reduction Factor can be expressed as a variance reduction factor κ . They show that:

$$\kappa(A) = \sqrt{E[\rho(u, u')]} \quad (6.6)$$

where $\kappa(A)$ is the value to which the point variance σ_p should be multiplied to obtain the variance of the areal rainfall. $\rho(u, u')$ is the correlation coefficient between two generic points u and u' contained within the area A .

ARF are proven to be dependent on the duration D and on the return period T_R . In this approach, no explicit dependences are shown. Nonetheless, it is well-known that the correlation structure of the field changes with the accumulation duration (the range of a rainfall field passes from about 20 km to about 100 km for the 1-hour and 24-hour rainfall, respectively) and with the return period (the higher the intensity of the event, the smaller its spatial extent and the steeper the ARF curves). The main issues concern the stationarity of the field in case of extreme events, which is not ensured, and the gaussianity of the field (denied by the analysis of radar images, not available at that time).

6.2.2.2 Lebel and Laborde approach

Lebel and Laborde (1988), similarly to the *Rodriguez-Iturbe and Mejía* (1974) approach, coupled geo-statistical analysis of rainfall fields with a model for rainfall extremes. The annual and monthly rainfall maxima are supposed to be Gumbel distributed. Assuming Gumbel-distributed maxima, the sample mean and variance completely define the maxima distribution. To describe spatial maxima, the Gumbel distribution is coupled to a covariance function in which, similarly to *Rodriguez-Iturbe and Mejía* (1974), the variance of the areal process decreases as a function of the integration surface.

An innovation in this context is the introduction of the climatological variogram (Section 4.4.4) as a tool for determining the average correlation structure of a series of fields characterized by different magnitude (and therefore different asymptotic variance).

The average of areal rainfall maxima and areal-maxima variance are given as a function of the point maxima average μ_z and variance σ_z^2 by the relationships:

$$\begin{cases} \mu_A = \mu_z \\ \sigma_A^2 = \sigma_z^2 \int \int_A \gamma(u, u') dud u' \end{cases} \quad (6.7)$$

where u and u' are two generic points within the domain A , and γ is the variogram function expressing the covariance between two points.

Lebel and Laborde (1988) derived the spatial-maxima probability distribution assuming that the point as well as spatial maxima distribution are Gumbel. The Gumbel spatial parameters Θ_{1A} and Θ_{2A} are defined as:

$$\begin{cases} \Theta_{1A} = 0.78 \left(\sigma_z^2 - \frac{1}{a^2} \int \int_A \gamma(u, u') dud u' \right) \\ \Theta_{2A} = \mu_A - 0.577 \Theta_{1A} \end{cases} \quad (6.8)$$

leading to the following distribution of spatial rainfall maxima:

$$F(z) = \exp^{-\exp-(z-\Theta_{1A})/\Theta_{2A}} \quad (6.9)$$

The variance reduction factor for the Gumbel distribution is evaluated as:

$$\kappa(A, T_R) = \frac{C_V^{-1} + r(A)(0.78(z - \Theta_2)/\Theta_1 - 0.45)}{C_V^{-1} + 0.78(z - \Theta_2)/\Theta_1 - 0.45} \quad (6.10)$$

in which $r(A)$ is a function of surface and of the climatological variogram:

$$r(A, \gamma) = \sqrt{1 - \frac{1}{A^2 \sigma_z^2} - \int \int_A \gamma(u, u') du du'} \quad (6.11)$$

6.2.2.3 Sivapalan and Blöschl approach

The approach of *Sivapalan and Blöschl* (1998) is similar to the *Lebel and Laborde* (1988) one. The main difference is that *Sivapalan and Blöschl* (1998) model the overall rainfall distribution instead of the extreme distribution. They consider that point rainfall is an exponentially-distributed variable featuring an exponentially-decaying correlation structure, and that the areal rainfall is Gamma distributed. The Gamma distribution is a generalization of the exponential case whose parameters are directly related to the exponential distribution parameters through the variance reduction factor previously seen in *Rodriguez-Iturbe and Mejía* (1974) and *Lebel and Laborde* (1988).

The Areal Reduction Factors with this approach can be computed as:

$$ARF(\kappa^2(A/\lambda^2), T_R, D) = \frac{b(T_R)c(T_R)\kappa^2 f_2(\kappa^{-2}) - \frac{\kappa}{f_1(\kappa^{-2})} \ln[\ln(\frac{T_R}{T_R-1})]}{b(T_R)c(T_R) \ln[\ln(\frac{T_R}{T_R-1})]} \quad (6.12)$$

where b and c are empirically determined as a function of the return period and f_1, f_2 are empirically determined as a function of the parameter k_A of the gamma distribution, knowing that k_A is related to the variance reduction factor according to the relationship $k_A = \kappa^{-2}$.

In Figure 6.3, the ARF curves for a particular weather regime in Austria are shown. Figure 6.3 shows that the ARF decrease with the increase of the return period; for return periods higher than 10 years the ARF dependence on the return period is weaker.

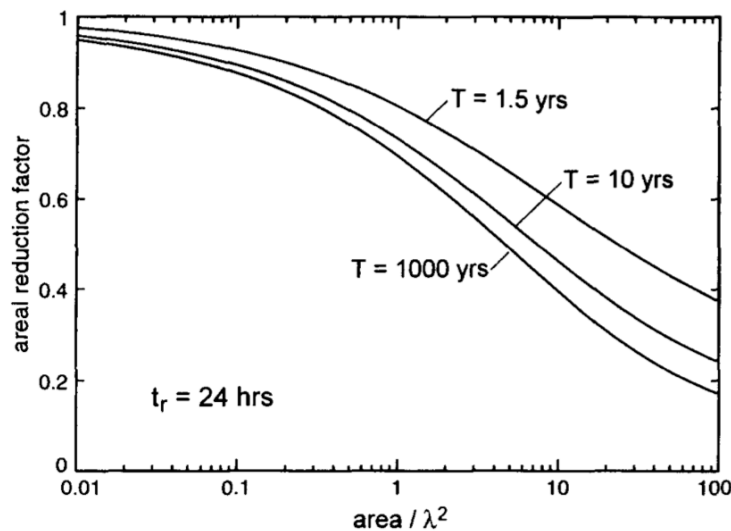


Figure 6.3: ARF curves for the K-45 weather regime in Austria, from *Sivapalan and Blöschl* (1998).

6.2.3 Stochastic approaches

6.2.3.1 Bacchi and Ranzi ARF approach

Bacchi and Ranzi (1996) present a stochastic derivation of the ARF based on the crossing properties of random fields. With the term “crossing properties”, the authors refer to the local behaviour of the spatial and temporal derivatives of the fields at points where a given threshold is exceeded. The number of exceedances of threshold levels is assumed to converge to the Poisson-distribution and a hyperbolic tail of the probability of exceedances of rainfall intensity has been adopted.

This approach assumes stationarity of the rainfall fields and homogeneity of the crossings in space. The ARF expression presents a number of fitting parameters that have to be inferred from data. The derived ARF curves assume power-law decay with respect to integration surface and duration of the storm.

Bacchi and Ranzi (1996) found a small decrease of the ARF with increasing T_R .

6.2.3.2 Asquith and Famiglietti

The annual-maxima centered approach proposed by *Asquith and Famiglietti* (2000) specifically considers the distribution of concurrent precipitation surrounding an annual-precipitation maxima. The advantage of the approach is the ease of computation, since it does not require the estimation of spatial rainfall averages or the determination of spatial correlation coefficients.

The proposed ARF model is less conservative than the TP-29 approach (*U.S. Weather Bureau*, 1958), in which the dependence on the return period T_R is taken into account.

The method focuses on point annual rainfall maxima. The ratio between these values and the precipitation depth recorded at the surrounding (“concurrent”) stations is evaluated.

The mean ratio in concentric rings surrounding the maxima is evaluated for the entire database. The method requires stationarity of the rainfall-maxima moments (stationary mean and variance). Although the method is easy and computationally cheaper, it is close to storm-centered approaches. It is useful for determining a design storm rather than to evaluate a statistically-significant value of spatial-rainfall.

6.2.4 Scale-invariant ARF

In this section we describe continuous ARF models based on the space-time scale invariance of rainfall intensity. The models previously discussed were mainly semi-empirical and scale-dependent. They do not allow to define an ARF parametrisation as a function of duration.

6.2.4.1 Dynamic scaling of space-time rainfall maxima

The approach proposed by *De Michele et al.* (2001) consists in adopting the concept of “dynamic scaling” proposed by *Venugopal et al.* (1999) at scales of hydrological interest. This hypothesis, physically consistent for scales lower than 45 min and 100 km^2 as a consequence of the Taylor’s frozen hypothesis (*Taylor*, 1938), cannot be extended to larger scales, because the advection component at larger temporal scales is not negligible. The dynamic scaling assumption is taken in this context in a statistical sense: for particular ratios of temporal and spatial scales, the probability distribution of rainfall maxima is scale-invariant.

The analysis starts from the concept that the distribution of rainfall maxima is scale-invariant in a statistical sense:

$$I(\lambda^a D, \lambda^b A) \stackrel{d}{=} \lambda^{-H} I(\lambda^a D, \lambda^b A) \quad (6.13)$$

where a and b are two parameters expressing the rainfall intensity decay increasing the temporal and spatial aggregation, respectively. According to *Gupta and Waymire* (1990) (Section 5.5.2), the equality is also valid in “wide sense”, that is in the sense of moments and quantiles. The expected value of annual maxima, for example, could be expressed as:

$$E[I(\lambda^a D, \lambda^b A)] \stackrel{d}{=} \lambda^{-H} E[I(\lambda^a D, \lambda^b A)] \quad (6.14)$$

where H is linear with q in case of simple-scaling.

De Michele et al. (2001) propose to describe the Intensity-Duration-Frequency-Area curves IDAF as the product of IDF and Areal Reduction Factor:

$$I(D, A, T_R) = I(D, T_R) \cdot ARF(A, D, T_R) \quad (6.15)$$

where the IDF formulation takes the form:

$$I(D, T_R) = a_1(T_R) D^{-v(T_R)} \quad (6.16)$$

The IDAF expression is determined by respecting the following asymptotic considerations:

- i. the general ARF form should be function of A^a and D^b ;
- ii. when $A \rightarrow 0$ (approaching the rain gauge scale), $ARF \rightarrow 1$;
- iii. when $A \rightarrow \infty$, the mean rainfall intensity tends to zero;
- iv. when $T \rightarrow \infty$, the areal rainfall equates the point rainfall.

By means of dimensional analyses, *De Michele et al.* (2001) found a suitable expression for IDAF curves:

$$I(D, A, T_R) = a_1(T_R) D^{-v(T_R)} \left[1 + \omega \left(\frac{A^z}{D} \right)^b \right]^{-v(T_R)/b} \quad (6.17)$$

where $z = a/b$ is the dynamic scaling ratio. Since a and b represent the power-law exponent defining the ARF decay with area and duration, respectively, z can be seen as the ratio between the spatial and temporal decay rate of ARF curves.

The influence of the return period is implicit in v , scaling exponent of the IDF formulation. Removing, in Equation 6.17, the IDF part (the same as Equation 6.16), one obtains the ARF expression:

$$ARF(D, A, T_R) = \left[1 + \omega \left(\frac{A^z}{D} \right)^b \right]^{-v(T_R)/b} \quad (6.18)$$

De Michele et al. (2002) propose to couple this IDAF model with a distribution for extremes. To take into account multi-scaling, according to *Burlando and Rosso* (1996), a log-normal distribution for extremes has been chosen. In a multiple scaling context, v varies with the return period. The log-normal density function accommodates the multiple scaling requirements, but it is not a density function for extreme value distributions. The multi-scaling IDAF model is therefore suitable for low return periods, far from the asymptotic conditions where extreme value theory is applicable. For higher return periods, maxima are GEV distributed (Section 3.5). GEV distribution only fits with the simple-scaling framework.

The model of *De Michele et al.* (2001) has been applied to the region of Milan, a region featuring 16 rain gauges for a surface of 300 km^2 (one rain gauges approximately every 20 km^2). The 8-years data series has been spatialized by kriging interpolation and the annual maxima for each window size have been recorded. The IDF model parameters, and subsequently the ARF model parameters have been estimated, allowing to build IDAF curves. The results obtained by *De Michele et al.* (2001) are reported in Figure 6.4-a.

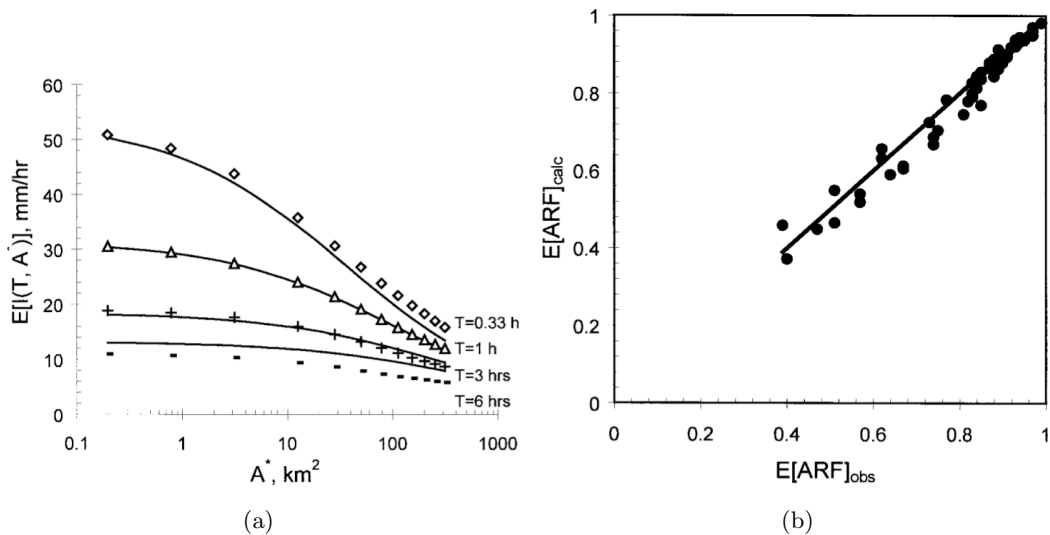


Figure 6.4: a: Intensity - Duration - Area - Frequency curves for $T_R = 2.35$ (corresponding to the average annual rainfall maxima) for the Milan urban area. b: Comparison of empirical ARF proposed by *NERC* (1975) with the result of *De Michele et al.* (2001) (the graphs are both taken from *De Michele et al.* (2001)).

According to the evidences shown in Section 2.4 it seems that the rain gauge density is too sparse to evaluate the spatial rainfall for areas lower than 10 km^2 . To catch the structure of the storm, a rain gauge network must have density considerably higher than the characteristic storm size. For this reason, it seems more appropriate to fit this model for durations larger than 1 hour (for which the storm characteristic size is larger than the inverse of the rain gauge density) and for surfaces higher than the inverse of the rain gauge density, so that at least two rain gauges are taken into account in the computation of the spatial rainfall.

The comparison with the results empirically determined by *NERC* (1975) show a good agreement of the two approaches (Figure 6.4-b).

6.2.4.2 Advection effects in ARFs

As stated in Section 6.2.4.1, the advection component is not negligible for durations higher than 40 min and areas larger than 100 km^2 (Deidda, 2000). Therefore the Taylor's frozen hypothesis is not valid beyond these scales. If physical assumptions behind this model are not applicable for singular storms, the "dynamic scaling" concept can be successfully applied in a statistical sense (De Michele et al., 2001), assuming that incoming directions of storms advected in the study domain are homogeneous in space, leading to a null-average advection velocity.

If one wants to analyze the effects of single storms within a basin, the scale-invariant climatological ARF curves, assuming no advection, are not really useful. In single storms, advection affects the decay of rainfall depth with the area implying anisotropy of ARFs and dependence of ARFs on the basin shape.

Veneziano and Langousis (2005) deal with this issue, proposing an IDAF model obtained by coupling the scale-invariant IDF model proposed by Veneziano and Furcolo (2002) with perfect multifractality in space-time. They consider that the advection velocity, similarly than in Taylor's frozen turbulence, can be used as a factor to rescale the temporal coordinate in order to obtain a space-time isotropic model. They propose equations describing IDAF and ARF curves for very elongated basins and regularly shaped basins.

The results Veneziano and Langousis (2005) have been successfully compared to the empirical results of NERC (1975) and (Bell, 1976) and with the theoretical model presented by De Michele et al. (2001). However, it seems that the basic hypothesis of Taylor's frozen turbulence that allows to rescale the temporal dimension, considering it as a third spatial dimension, can not be assumed for scales larger than the turbulence one.

6.2.4.3 IDAF in small space-time rainfall

A small-scale experiment has been conducted by Castro et al. (2004) to study the space-time structure of rainfall in order to derive IDAF relationships. Rainfall has been filmed in a spatial cube of 1.2 m 1.2 m 1.2 m, at a rate of 30 images/s. Drop sizes have been inferred from vertical velocities. The 2Dx1D singularity spectrum has been evaluated considering rainfall as a multi-fractal process in (x,y,t).

The rainfall intensity i is considered dependent on duration D , area A and return period T_R following the relation:

$$i(A < D < T_R) \propto D^{-1} A^{-z/2} T_R^\delta \quad (6.19)$$

fitting Equation 6.19 with experimental data allows to determine the following empirical formula:

$$i \propto T_R(I(d, A) > i)^{1.227} d^{-1} A^{-0.581} \quad (6.20)$$

6.3 Max-stable spatial-maxima modeling

A relatively new framework for the modeling of rainfall extremes in space is the Max-Stable theory. It represents a generalization of the extreme value theory valid for point series (Section 3) and bivariate

extreme analysis (*Herr and Krzysztofowicz, 2005; Renard and Lang, 2007*) to multivariate data, such as the rainfall measured at a network of rain gauges.

The extreme value theory (*Smith, 1990*) states that a sample of i.i.d. point-maxima is GEV-distributed (Section 3.5). Differently from what has been assumed in *Lebel and Laborde (1988); Sivapalan and Blöschl (1998)*, spatial extremes are distributed according to Max-Stable distributions, describing the joint distribution of maxima at several locations.

To apply max-stable theory, the rainfall maxima Y at one location, distributed following GEV, should be transformed into a Unit Fréchet random variable Z through the transformation:

$$Z = \left(1 + \xi \frac{Y - \mu}{\sigma}\right)^{1/\xi} \quad (6.21)$$

The max-stable theory aims to model the joint distribution $P(Z(x_1) \leq z, \dots, Z(x_n) \leq z)$.

Two interpretations are given to the concept of continuous max-stable fields, the Smith's model (*Smith, 1990*) and the Schlatter model (*Schlatter, 2002*).

The Smith's model defines a max-stable random process $Z(x)$ as:

$$Z(x) \stackrel{D}{=} \max_k (U_k f(x - T_k)) \quad (6.22)$$

where U_k can be interpreted as the storm sizes, f the shape of the storm centered at T_k . Figure 6.5 shows an intuitive interpretation of the Smith's model as the envelop of the absolute maxima of different storms.

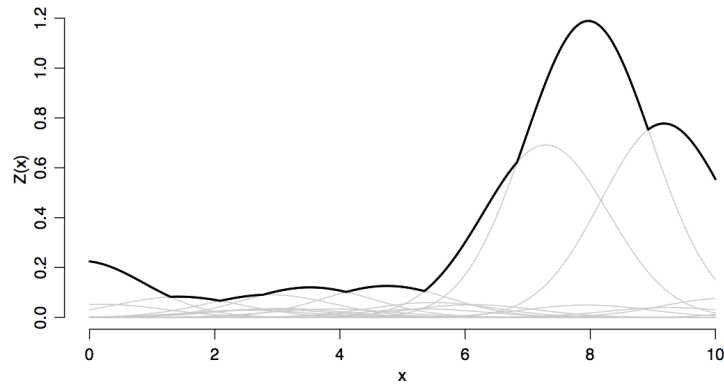


Figure 6.5: Representation of the Smith's model for a max-stable field.

When Z is a unit-Fréchet point-process, its cumulative density function is:

$$P(Z \leq z) = \exp\left(-\frac{1}{z}\right) \quad (6.23)$$

The cumulative joint probability is given by:

$$P(Z(x_1) \leq z, \dots, Z(x_n) \leq z) = \exp\left(-\frac{\theta}{z}\right) \quad (6.24)$$

where θ is called “extremal coefficient”, indicating the degree of mutual dependence of the N stations. θ takes values between 1 and N and provides a measure of the degree of spatial dependence between stations¹. More precisely, the extremal coefficient θ depends on the correlation function and has two different analytical expressions for the Smith’s and the Schlater’s model. It can be computed starting from variograms, but a more general estimation requires the computation of the first order generalized structure function, called “madogram” (Cooley *et al.*, 2006; Matheron, 1987). The madogram computation only requires that the first order moment (i.e. the mean) of the random field to be finite.

The madogram is defined as

$$\nu(h) = \frac{1}{2}E[|Z(x) - Z(x+h)|] \quad (6.25)$$

where Z is a stationary random field with finite mean and h the separation lag between stations.

Cooley *et al.* (2006) derived the expression relating the extremal coefficient θ and the madogram ν :

$$\theta(h) = \begin{cases} \mu_\beta(\mu + \frac{\nu(h)}{\gamma(1-\xi)}) & \text{if } \xi < 1 \\ \exp\left(\frac{\nu(h)}{\sigma}\right) & \text{if } \xi = 0 \end{cases} \quad (6.26)$$

where μ, σ, ξ are the GEV distribution parameters, γ the Gamma function and μ_β is given by:

$$\mu_\beta = \left(1 + \xi \frac{x - \mu}{\sigma}\right)^{1/\xi} \quad (6.27)$$

The method, theoretically consistent, represents without doubt a significant advance in the field of spatial extremes estimation. Nonetheless, a series of remarks should be highlighted:

- the extremal coefficient θ is an indicator of the spatial dependence between all stations. On the other hand, the madogram, used for estimating the extremal coefficient, is a pairwise indicator of the correlation. A consistent estimator based on a multivariate indicator should be preferred.
- the madogram is evaluated on maxima that are not concurrent (therefore it is not physically based);
- the normalization of maxima to Unit Fréchet is not sufficient to get rid of their spatial heterogeneity. The approach is, again, limited to homogeneous (i.e. relatively small) area.
- the transformation of GEV into Unit Fréchet and the fitting of a max-stable model are two necessary steps to determine the spatial extremes behaviour. Padoan (2008) proposes a direct fitting of GEV and spatial dependence in a single step so that the estimation error could be determined, but its use is limited to gauged sites. The use of response surface (describing the spatial dependence of the GEV parameters on the location) could help in solving this problem in regions where the drifts are easy to be modeled. In Cévennes-Vivarais region, the GEV parameters are strongly influenced by the orography, preventing the use of response surfaces.

¹Poorly speaking, it can be seen as the number of independent stations as a function of the lag h .

6.4 Conclusion

This chapter provides an overview of the main methodologies described in literature to deal with the extreme rainfall intensity decrease as a function of the spatial scale.

In the earlier approaches, the so-called Areal Reduction Factor (ARF) curves have been empirically determined. Basically, it is expressed as the ratio between the spatial rainfall associated to a given return period to the point rainfall associated to the same return period. ARFs express the factor respect to which the areal rainfall should be divided to obtain the point rainfall characterized by the same occurrence.

In the context of storm-centered approaches, semi-empirical models have been proposed, based on the spatial correlation of rainfall extremes. They rely on the definition of an average or climatological variogram, that should correctly model the average spatial structure of extremes. The statistical significance of storm-centered ARFs has been questioned by *Omolayo* (1993). In addition, the law of spatial extremes can not be seen as a simple extension of the law of extremes for point rainfall.

A statistically self-similar approach has been presented by *De Michele et al.* (2001). Of easy application, it relies on the relation between spatial and temporal scales under the concept of “dynamic scaling”. As it has been defined, at scales beyond the validity range of the Taylor frozen hypothesis (Section 5.10.1), the “dynamic scaling” concept is intended in a purely statistical sense.

Finally, we briefly summarized the main points of the max-stable maxima modeling, representing the extension of the extreme value theorem for spatial rainfall.

Part III

Results

7

Heavy tails of rainfall distributions

Résumé

Le but de cette étude est d'évaluer les propriétés d'invariance d'échelle temporelle des précipitations extrêmes ponctuelles. Dans la région d'intérêt, les queues des distributions de probabilité des pluies horaires montrent linéarité en log-log, une caractéristique des distributions à queues lourdes. La conservation de cette propriété aux différentes échelles est compatible avec l'invariance d'échelle des extrêmes, et a été vérifiée pour une longue série pluviométrique dans la gamme d'accumulations 1-24 h. Dans l'article on propose une méthode objective pour estimer les deux paramètres qui définissent la queue hyperbolique d'une distribution. L'application de méthodes d'estimation non biaisées nous a permis d'attendre une réduction significative de la variabilité dans l'estimation des paramètres de la queue hyperbolique. Ce résultat nous permet d'attendre deux objectifs: i) vérifier la présence d'invariance d'échelle en analysant les queues hyperboliques d'une série pour différentes durées d'accumulation; ii) obtenir une structure spatiale cohérente et parvenir à la régionalisation de ces paramètres par méthodes géostatistiques. Les cartes montrent que les queues de distribution sont hyperboliques dans la zone de plaine du Rhône entre le bord de mer et les contreforts. De plus, l'exposant de la loi de puissance est pratiquement constant avec la durée, alors que dans le piémont et sur la crête du massif des Cévennes les queues sont plutôt exponentielles (réjet de la loi de puissance). Nous discutons les raisons physiques pour les résultats et les conséquences pour la modélisation statistique de fortes pluies. Un point de vue novateur pour la compréhension des précipitations extrêmes dans un cadre d'invariance d'échelle est offert.

L'article tente de fournir des éléments pour élucider deux des questions liées au comportement des fortes pluies:

- i. est le comportement de fortes précipitations différent en fonction de l'échelle temporelle?*
- ii. sont les queues de précipitations hyperboliques ou exponentielles?*

Examinons ces points en détail. Il est bien connu que les maxima journaliers sont généralement modélisés avec la distribution de Gumbel. Cette hypothèse, raisonnable dans la plupart des cas, correspond à supposer les extrêmes distribués de façon exponentielle. L'avantage est de simplifier l'estimation de la loi des extrêmes (le modèle est défini par 2 paramètres au lieu de 3). Cette hypothèse

ne peut pas se vérifier directement sur les données, en raison de la longueur limitée des séries. Un deuxième point pertinent est que, lorsque l'on tente de modéliser les maxima infra-journalier, les queues de distribution sont systématiquement plus lourdes que prévu par le modèle Gumbel. Habituellement, les données infra-journalières sont modélisées en utilisant une distribution Fréchet (GEV-II). Toutefois, l'utilisation de deux lois distinctes pour modéliser les maxima journaliers et infra-journaliers est en contraste avec les évidences de la continuité du comportement des précipitations (moyenne, variance, quantiles) dans la gamme d'accumulations 1 h - 15 jours. Comme il n'y a aucun moyen d'estimer le comportement des queues directement à partir de la distribution des maxima, nous proposons dans cet article d'examiner la totalité de la distribution de probabilité et de vérifier si cette continuité à travers des échelles est vérifiée ou non.

En considérant toute la distribution, il est possible d'avoir davantage de données, mais les hypothèses pour l'application de la théorie des valeurs extrêmes ne sont plus valables (notamment les hypothèses de i.i.d.). Par conséquent, l'analyse de la distribution de probabilité peut servir qu'à l'évaluation du comportement hyperbolique ou exponentiel des données et à la vérification des propriétés d'invariance d'échelle, ne permettant pas une estimation directe des quantités extrêmes.

7.1 Introduction

The article deals with heavy rainfall observation and tries to provide elements to elucidate two of the questions related to the behavior of heavy rainfall events:

- i. is the behavior of heavy rainfall different at various temporal scales?
- ii. are the rainfall tails hyperbolic or exponential?

Let us examine these points in detail. It is well known that daily maxima are usually modeled with the Gumbel distribution. This assumption, reasonable in the most of cases, corresponds to consider exponentially-tailed extremes. The advantage is to simplify the estimation of the law of extremes (the model is defined by 2 parameters instead of 3), but this assumption could not be directly verified on data, due to the limited series length. A second point is that, when one attempts to model infra-daily maxima, the distribution tails are systematically thicker than exponential. Usually, infra-daily data are modeled using a Fréchet (GEV-II) distribution. However, the evidences of the continuity of the rainfall behavior (mean, variance, quantiles) throughout the scales 1 h - 2 weeks make surprising that two consistently different laws model must be used to model daily and infra-daily maxima. Since there is no way to estimate the tails behavior directly from the maxima distribution, we propose in this article to examine the parent probability distribution and to check whether this continuity throughout the scales is verified or not.

Taking the entire distribution, it is possible to have more data, but the assumptions for the application of extreme value theory such as the i.i.d. hypothesis do not hold anymore. By consequence, the analysis of probability distribution may serve only for the assessment of the hyperbolic or exponential behavior of data and for the verification of the scale-invariance properties and does not allow a direct estimation of extreme quantities.

**7.2 Article: Scaling properties of heavy rainfall at short durations:
a regional analysis**

Water Resources Research, VOL. 46, W09531, 2010

Scaling properties of heavy rainfall at short duration: a regional analysis

D. Ceresetti, G. Molinié, J.-D. Creutin

LTHE - Laboratoire d'Étude des Transfert en Hydrologie et Environnement - Université de Grenoble (CNRS, INPG, IRD, UJF), France.

Abstract. The aim of this paper is to assess the scaling properties of heavy point rainfall with respect to duration. In the region of interest, the probability distribution tails of hourly to daily rainfall display log-log linearity. The log-log linearity of tails is a feature of fat-tailed distributions. The conservation of this property throughout the scales will be investigated in the framework of scale-invariant analysis. Evidences of the scaling of heavy rainfall are shown for one particularly long rainfall series through the conservation of the survival probability shape at durations in the range 1-24 hours. An objective method is implemented to estimate the hyperbolic-tail parameters of rainfall distributions. This method is automatized and detects the lower bound above which the distributions exhibit power-law tails and determines the power-law exponent α using a maximum likelihood estimator. The application of unbiased estimation methods and scale-invariant properties for the estimation of the power-law exponent provides a significant reduction of the inter-gage power-law variability. This achievement is essential for a correct use of geostatistical approaches to interpolate the power-law parameters at ungauged sites. The method is then applied to the raingage network in the Cévennes-Vivarais region, a Mediterranean mountainous region located in Southern France.

The maps show thicker rainfall-distribution tails in the flat area between the sea shore and the foothill. It is shown that in a flat region closer to the Mediterranean Sea the rainfall distribution tails are hyperbolic and the power-law exponent is quasi constant with duration whereas, over the mountain, the power-law behavior is less defined. The physical reasons for such results and the consequences for the statistical modeling of heavy rainfall are then discussed, providing an innovative point of view for the comprehension of the rainfall extremes behavior at different temporal scales.

1. Introduction

During the last thirty years, a considerable body of investigations analyzed the scale-invariance of rainfall, demonstrating that rainfall fields have intrinsic scaling properties within a specified range of scales. A physical process is scale-invariant if its probability distribution, once applied a rescaling factor, does not change under scale magnification or contraction within a given range.

Frisch and Parisi [1985] provided fundamental insights into the multi-scaling behavior of processes. Analyzing the average value of the q -th power of the change in the turbulent velocity for different time lags, they found that $|v(h) - v(h+l)|^q$ varies as the power-law $l^{\zeta(q)}$, where ζ is non linear with q . The non-linearity of $\zeta(q)$ indicates that the velocity fluctuations display multifractal scaling. A Legendre transform allows to switch from the moment scaling function to the codimension function $c(\gamma)$, describing the scaling in terms of probability distribution. The singularity order γ in the codimension function expression is the dual of the moment order q in the moment scaling function.

A particular case of scaling, referred to as simple-scaling, occurs when the scaling exponent $\zeta(q)$ is linear with q . In simple-scaling processes the probability distribution is rescaled from a scale to another by means of a single scaling exponent, while in multifractality the scaling exponent

depends on the degree of singularity of the process. The distribution equality between two probability distributions at different scales is referred to as “strict sense scaling”. A weaker property is usually adopted for assessing the scaling behavior of a process: the equality of moments, referred to as “wide sense scaling” [*Gupta and Waymire*, 1990].

First evidences of the multiscaling behavior of meteorological fields were shown by *Schertzer and Lovejoy* [1987] analyzing meteorological radar reflectivities. *Gupta and Waymire* [1990] evaluated and detailed the different types of scaling of instantaneous radar rainfall with respect to the surface. The multiscaling concept can be also applied to time series of raingage data over a wide range of temporal scales. *Ladoy et al.* [1993] analyzed a pluviometric series located in Nîmes (France) covering 50 years of data characterized at a 12 hours time resolution. They have been able to determine the multifractal parameters of the rainfall series finding scale-invariance in the range 12 hours - 16 days. *Hubert et al.* [1993] analyzed data from different regions at temporal scales of 6-min at the Reunion Island (Indian Ocean), of 15-min in the French Alps, of one day at Nîmes and at Dédougou (Burkina Faso, West Africa). They found out multiple scaling behavior of point rainfall rates from one to several days (16 - 30 days depending on the location).

The scale-invariance of rainfall maxima has been the topic of several studies. *Bendjoudi et al.* [1997] derived a multifractal based Intensity-Duration-Frequency formulation showing that a multi-fractal phase transition implies algebraic tails above a given singularity level. From the direct analysis of rainfall data series, *Burlando and Rosso*

[1996]; *Menabde et al.* [1999]; *Borga et al.* [2005] showed that the annual maxima are approximately simple scaling in the range 0.5-24 hours. It is worth mentioning that *Burlando and Rosso* [1996] and *Menabde et al.* [1999] processed data series from different climatic regions. In *Burlando and Rosso* [1996], one raingage station is located in a flat area whereas the second is a mountainous station. *Menabde et al.* [1999] dealt with a midlatitude temperate region of Australia and with a semiarid region of South Africa.

The usual approach for the analysis of heavy rainfall is based on the extreme value theory. This theory considers events exceeding a given threshold (Peak Over Threshold - POT) or maxima during a given period (annual or shorter periods, see *Kotz and Nadarajah* [2000] for more details), resulting in two distribution classes: Generalized Pareto Distributions (GPD) and Generalized Extreme Value (GEV) distributions. The GEV distribution class involves three types of maxima: hyperbolic-tailed (GEV-II), exponential (GEV-I) and bounded maxima (GEV-III). Maxima rainfall rates usually follow GEV-I (Gumbel) or GEV-II (Fréchet) distributions [*Kottegoda and Rosso*, 1997], depending on the decay of the probability distribution (respectively exponential and hyperbolic tailed, hyperbolic tails being thicker than exponential ones). Daily rainfall maxima are often modeled with Gumbel distributions [*Gumbel*, 1958; *Koutsoyiannis et al.*, 1998]. Nevertheless, infra-daily rainfall can show thicker tails and the improper use of GEV-I distribution leads to a generalized underestimation of extreme events for high return periods. The choice between the two approaches (hyperbolic or exponential tailed-distributions) has been justified only by empirical evidence on the distribution of maxima.

Few studies on heavy-rainfall scaling focused on the properties of the underlying probability distributions. *Hubert and Bendjoudi* [1996] studied the distribution of heavy rainfall in Dédougou over scales ranging from one day to one year. Analyzing the power-law exponent α in double-logarithmic plot, they showed that the hyperbolic tail of the pdf does not change with the accumulation period, resulting in simple-scaling of rainfall extremes. Approximate simple scaling is also illustrated in Table 3 of *Sivakumar* [2000] analyzing the hyperbolic tail of the probability distributions for accumulation durations between 6 hours and 7 days for two rainfall series at Singapore and Leaf River Basin (Mississippi - USA).

In this paper, our purpose is to give a regional description of heavy rainfall statistics. According to our knowledge, no study focused on the inter-gage properties of the probability distribution tails. Many papers focused on the behavior of single raingages, not representative of a complex-relief region. Applying an objective method for determining the power-law exponent α , and a scale-invariant relationship that involves the hyperbolic tail of the distribution, our aim is to reduce the inter-gage variability of the power-law parameters. According to our experience, the determination of the power-law parameters using an arbitrary-threshold based method prevents a robust parameter estimation and thus the comparison between gages and the regionalization of the variable. Applying the method to about two hundred stations, we show that a coherent interpolation process is now possible as well as the prediction of the tail behavior at ungauged sites.

The paper is structured as follows: we first show evidence of hyperbolic behavior of heavy rainfall at specific stations and we describe the objective method implemented in order to determine the hyperbolic model parameters (Section 2). In Section 3, we use a reference rainfall series (50 years) to check for the simple scaling properties of distribution tails between 1 and 24 hours. The goal of Section 4 is to map the model parameters of heavy rainfall. A comparison of the power-law exponent α at accumulation durations from 1 h to 8 h allows us to identify sub-regions where α is approximately constant. The results related to the rainfall-forcing processes and to the extreme-value theory are discussed in Section 5.

2. Heavy point rainfall behavior

In this section, we characterize the positive rainfall rates by their survival probability (complement to 1 of the cumulative distribution function). The decay of the survival probability gives information about the underlying law of extremes. An example of the survival probability of hourly point rainfall is plotted in Figure 2 for the raingage station of Cognac, about 30 km SW of Alés, France (see Figure 1). As we are interested in heavy rainfall, the plot is limited to the upper 5% of the observations. The survival probability tail is hyperbolic. It can be parametrized by a decay rate α and by a lower bound x_{min} . In practice, the decay rate is usually estimated by fitting a power-law to the data (i.e. a straight line on a log-log plot) and calculating its slope. However, this particular fitting process is influenced by the empirical estimation of the survival probability. One can see that the empirical survival probability of the highest observations on the log-log plot in Figure 2 diverges from the straight line. As illustrated in the following simple exercise, this can be interpreted as a consequence of the mode of computation of the highest empirical frequencies (outliers). Taking $N=100$ realizations of a random variable X , the empirical survival probability can be defined, if we take the Weibull plotting-position expression, by:

$$P(X > x) = 1 - \frac{i}{N+1}, \quad (1)$$

where i is the rank of a sorted sample x , varying from 1 to N . Let us consider to add one further observation to the series. If this observation is the highest of the sample, it will be ranked 101 in the sorted sample. The survival probability of the 100-ranked sample will be modified by 100%, passing from 0.01 to about 0.02. At the same time, the survival probability of the 50-ranked sample will be modified by only 0.5%, passing from about 0.5 to about 0.495. This exercise highlights that empirical survival probabilities are robustly estimated when looking at usual probability levels, but strongly biased when dealing with extreme quantiles. The bias of a plotting-position formula depends on the probability distribution of the sample. More adapted expressions for the determination of the plotting position in positively skewed data are available for several probability distributions [*Kottegoda and Rosso*, 1997], and the exercise above can be easily generalized for any of these.

This example highlights that any power-law fitting method based on the plotting of empirical cumulative density function is affected by large uncertainties, which increase in presence of outliers in the probability distribution. *Goldstein et al.* [2004] showed the inaccuracy of some of these graphical methods by calculating the bias in the estimation of the power-law exponent α of samples composed by 10000 realizations. They found out that the Maximum Likelihood Estimator (MLE) provides a better estimate $\hat{\alpha}$ than other methods, including Least-Squares Linear Regression (LSq). The MLE estimator (Equation 2) is equivalent to the Hill estimator adopted in extreme value theory.

$$\hat{\alpha} = 1 + n \left[\sum_{i=1}^n \ln \frac{x_i}{X_{min}} \right]^{-1}. \quad (2)$$

The main advantage using MLE with respect to LSq is that the method provides an unbiased estimate of the exponent $\hat{\alpha}$, independently of the empirical cumulative distribution. We performed complementary simulations to extend the numerical experiment of *Goldstein et al.* [2004] to shorter series (sets of about 1000 realizations) drawn from a Pareto distribution:

$$P(X \geq x) = \left(\frac{x}{x_{min}}\right)^{-\alpha}, \quad (3)$$

for all $x \geq x_{min}$, where x_{min} is the so-called scale parameter and α the shape parameter. One hundred series with $x_{min} = 10$ and $\alpha = 3$ have been generated with N , the number of realizations, ranging from 100 to 10000. In Figure 3 the box-plots summarize the distributions of the estimated $\hat{\alpha}$ computed using, respectively, the LSq and MLE methods. Considering a set of 10000 samples, we notice that LSq provides far more scattered estimations of $\hat{\alpha}$ than MLE, in agreement with *Goldstein et al.* [2004]; the dispersion of $\hat{\alpha}$ remains of the same order of magnitude whatever the sample set size. Moreover, the average of $\hat{\alpha}$ estimated by MLE remains close to 3 (the actual value), while it is more fluctuating when estimated by LSq. This confirms that MLE is a more consistent estimator than LSq. Two reasons have been highlighted in literature. First, LSq is more sensitive to the presence of outliers in the distribution tails. Second, the residuals of the linear fitting of log-reduced variables do not follow a Gaussian distribution [*Clauset et al.*, 2009].

After having defined an unbiased method for the estimation of the power-law exponent of hyperbolic distributions, the second major problem is to define the scale parameter (x_{min} in Equation 3) i.e. the lower bound above which the power-law holds. Considering an arbitrary bound, as we did in Figure 2, is obviously not satisfactory. The determination of $\hat{\alpha}$ depends on the choice of the lower bound \hat{x}_{min} . This bound can be different from one raingage to another, because of the heterogeneity of the rainfall regime in the region. *Clauset et al.* [2009] derived a method to estimate the lower bound \hat{x}_{min} .

The probability density function of a variable y assuming discrete values and distributed as a power-law is defined as: (Equation 4 [*Goldstein et al.*, 2004]):

$$p(x) = \frac{x^{-\alpha}}{\zeta(\alpha, x_{min})} \quad (4)$$

where α is the power-law exponent and $\zeta(\alpha, x_{min})$ is the generalized Zeta function, defined as:

$$\zeta(\alpha, x_{min}) = \sum_{n=0}^{\infty} (n\Delta x + x_{min})^{-\alpha} \quad (5)$$

where x_{min} is the lower bound and Δx is the raingage accuracy (0.1 mm of rain depth for the analyzed raingage database).

The estimated lower bound \hat{x}_{min} is determined by means of the Kolmogorov-Smirnov (KS) statistics. *Clauset et al.* [2009] have shown that this objective method is among the most efficient for comparing two distributions. The D statistics of the KS test is defined in Equation 6:

$$D = \max_{x \geq x_1} |S(x) - P(x)| \quad (6)$$

where $S(x)$ and $P(x)$ are the cumulative probability distributions of the observed samples and of the model, above a lower bound x_1 . Figure 4 shows the D statistics as a function of x_1 for the raingage of Colognac. The value of x_1

corresponding to the minimum of D provides the estimated \hat{x}_{min} , 7.2 mm h^{-1} in the case shown in Figure 4. Therefore, $\hat{\alpha}$ is estimated applying MLE to the X realizations higher or equal to \hat{x}_{min} .

Figure 5 illustrates the sensitivity of $\hat{\alpha}$ to x_1 . For x_1 higher than 20 mm h^{-1} the $\hat{\alpha}$ sensitivity to x_1 , as well as the estimation uncertainty, is consistently high. This is the result of the rapid decrease of the sample set size, and in this method, as well as in other methods, α is never taken in this range. For values of x_1 close to the optimal value \hat{x}_{min} , i.e. within the range $0.1 - 15 \text{ mm h}^{-1}$, the sensitivity of $\hat{\alpha}$ is considerably lower, varying of some decimals. However, it is worthy to notice that small variations of $\hat{\alpha}$ can have relevant influence in the estimation of rainfall for very high quantiles.

3. Scaling behavior of heavy rainfall

In the previous section, we stated that the tail distribution of hourly rainfall behaves as a power-law at many raingage stations and we described a method to estimate the power-law parameters. In the current section, we investigate the conservation of this property for temporal resolutions ranging from 1 to 24 hours, for the longest hourly-raingage series of the region, located in Montpellier (see Figure 1). This raingage collected over 50 years of hourly data, in the period 1920-1972. This raingage has been used for testing some of the properties that we assume throughout the paper. Rainfall rates for four durations (1, 4, 10 and 24 hours) have been computed by aggregation within non-overlapping windows. To make possible their scale-free inter-comparison, the sample sets are firstly normalized by the mean rainfall rate, subsequently, for each duration, a sample with fixed size is chosen (2000 non-zero samples).

The assumption of data stationarity is often required to analyze the heavy-rainfall behavior. On the other hand, the sample set size has to be as long as possible to improve the robustness of the statistics. These two requirements could be incompatible. The stationarity of the rainfall intensities of the 50-year long data set of Montpellier is thus questionable. Therefore, we have checked this stationarity in computing the survival probabilities displayed in Figure 6 for two consecutive sub-periods lasting 25 years each. The two sub-periods do not show considerable differences.

Figure 6 shows that the empirical survival probability exhibits hyperbolic tails at durations of 1, 4, 10 and 24 h. The lower bounds x_{min} and slopes α are computed using the method described in Section 2. Figure 6 emphasizes that, at any duration between 1 and 24 h, the hyperbolic tail has an approximately constant slope, while the variability of the series with respect to the mean decreases with the accumulation duration. The lower bound x_{min} above which the power-law behavior holds depends on the rainfall duration. If x_{min} is the limit of the hyperbolic tail and the simple scaling holds at this point, x_{min} should scale as a function of the accumulation duration such that the absolute quantile is a constant.

Therefore, the highest rainfall rates of this long series display simple scaling properties for durations in between 1 and 24 hours. If the rainfall rate is a random process $X(t)$ ($t \in \mathfrak{R}$), we are able to magnify or contract by a factor λ the highest rates without modifying the distribution shape [*Sornette*, 2004, p. 148]. As stated by *Gupta and Waymire* [1990], we can compute a scale function $\lambda^{\frac{d}{\theta}} \geq 0$ such that:

$$X(\lambda t) \stackrel{d}{=} \lambda^{\frac{d}{\theta}} X(t). \quad (7)$$

The equality in distribution (Equation 7) is referred to as “strict sense simple scaling”. It is obvious in Fig. 6 that the strict sense simple scaling does not apply to the whole rainfall rate distribution (also stated by *Gupta and Waymire* [1990]) which is rather multifractal ([*Hubert et al.*, 1993; *Tessier et al.*, 1993]). However, this is not incompatible with the simple scaling behavior observed for the highest rainfall rates. Several studies showed evidences of the simple scaling behavior of very high quantiles, such as annual maxima of the rainfall rate [*Burlando and Rosso*, 1996; *Bendjoudi et al.*, 1997; *Menabde et al.*, 1999; *Borga et al.*, 2005] while other authors reported a change in the high rainfall-quantile behavior that *Schertzer and Lovejoy* [1992] define as “Multifractal Phase Transition”.

4. Regionalization of the power-law exponent

4.1. Study region and data

The Cévennes-Vivarais region is located in the South-East of France (see Figure 1). This region is prone to heavy rainfall events causing flash floods [*Jacq*, 1994; *Delrieu et al.*, 2005]. Typical meteorological conditions have been detected as triggering conditions for flash-floods, mainly the advection of warm-humid air from the South.

The region is southerly bounded by the Mediterranean sea providing warm and humid air masses. The Alps massif to the East and the Massif Central to the West channel the flow in the Rhône River valley (eastern boundary of the study region). The Massif Central mountain range, approximately oriented north-northwest, is impacted by low level air masses from South and favors their lifting. The North-Western part of the study region, usually less concerned by severe rainfall events, is constituted by flat highlands.

The raingage network in the region has been installed at the beginning of the previous century. However, digitized hourly rainfall data are available only since 1993. In this study, we used data from 1993 to 2008 provided by the French Meteorological Service Météo-France. From 1993 to 2000 about 150 raingages were available; this number increased to about 200 after the year 2000 (date of implementation of the Hydrometeorological survey service: OHMCV, *Delrieu* [2004]). The raingage density is very fluctuating from one place to another (see Fig. 1) and the mean raingage density is approximately one per 150 square kilometers.

4.2. Methodology and implementation

The hyperbolic behavior and self-similarity of the distributions of heavy rainfall intensities cumulated over periods from 1 to 24 hours have been empirically assessed in Section 3. In the current section, we regionalize the parameters characterizing the self-similarity of heavy rainfall rates at different durations.

The steps involved in the estimation of the power-law exponent α at a raingage are the following:

1. selection of a raingage having at least 2000 non-zero observations at the duration $D = 1h$;
2. cumulate the raingage observations over higher accumulation durations through a fixed-window process; we cumulated at 2, 4 and 8 hours.
3. for $D=1h$, estimate x_{min} by minimizing the D statistics of the Kolmogorov-Smirnov test (for each value of x_1 , a value of α is computed and the statistics D is returned);
4. estimate the quantile of x_{min} in the complete rainfall series (both zero and positive values)
5. for each duration $D > 1h$, X_{min} is computed as the value of x_D corresponding to the same quantile as for $D=1h$;
6. compute α with the method proposed by *Clauset et al.* [2009] taking X_{min} as the lower bound, following Equation 2.

The interpolated exponent $\hat{\alpha}_1$ of the point rainfall for the 1-hour duration is mapped in Figure 7-a, clearly showing elongated structures corresponding to the mountain ridge. The regionalization of $\hat{\alpha}$ is obtained by interpolation, performed only if the variable has a definite correlation structure. The Universal Kriging method (described by *Chiles and Delfiner* [1999]) has been chosen to interpolate the values of $\hat{\alpha}$ at different accumulation times. We emphasize that even though the value of $\hat{\alpha}_1$ has been spatially interpolated, it remains a local measure. Since different mechanisms are involved, the integration of α over a surface does not correspond to the areal power-law exponent.

The interpolated $\hat{\alpha}_1$ values can be altered by two kinds of errors. One is due to the interpolation process; the second is due to the assumption of hyperbolic behavior of rainfall distribution tails and their fitting. The former is evaluated through the kriging standard deviation displayed in Figure 9-a. This Figure shows that except in the domain fringe, the kriging standard deviation is lower than 10% of $\hat{\alpha}_1$ which we consider as acceptable in comparison, for instance, to the variation of 20% of $\hat{\alpha}$ across the region. The confidence interval $\Delta\hat{\alpha}$ efficiently assesses the latter error type (i.e. the reliability of the point $\hat{\alpha}$ estimation). In Figure 8-a, the confidence interval $\Delta\hat{\alpha}$ for the 95% confidence level is mapped for the region of interest. Figure 8-b shows that the confidence interval roughly varies between 14% in the plain region to 26 % of $\hat{\alpha}$ in the northern part of the region of study. The α estimation is the most reliable in the southern part of the study region. The lowest $\hat{\alpha}_1 \sim 2.6$ are located at the lowest altitude and increase gradually with altitude up to the Cévennes-Vivarais mountain ridge ($\hat{\alpha}_1 \sim 3.6$) and the Alps. In the Rhône river valley the gradient is weaker. We have to point out that, in the mountainous sub-region, the power-law model is less adapted to the series, as shown by analyzing the confidence interval (Figure 8-b). In the next section, we will evaluate the α exponent for the accumulation periods of 2, 4 and 8 hours.

4.3. Regional rainfall scaling assessment

Following methodology described in the previous paragraph, the assessment of the simple scaling assumption is undertaken in the whole study region by evaluating the α behavior at different accumulation periods. At each raingage, rainfall rates are aggregated over 2, 4 and 8-hour periods using non-overlapping windows. The 4-hour limit guarantees sufficiently long rainfall series (> 500) while, for the 8-h interval, most of gages had been discarded due to the poor sample set size.

Since the quantile of x_{min} is assumed to be scale-invariant (Section 2), this property has been used to retrieve its values at the 2-, 4- and 8-hour durations from x_{min} computed at the 1-hour duration. Using the Maximum Likelihood estimation method (2), $\hat{\alpha}_2$, $\hat{\alpha}_4$ and $\hat{\alpha}_8$ are estimated (Section 2) and mapped (Section 4.2, Figure 7-b,c,d). The interpolation variance associated to the $\hat{\alpha}_2$, $\hat{\alpha}_4$ and $\hat{\alpha}_8$ kriging is almost identical in pattern and displays increasing values of the estimation error with duration (Figure 9-b,c,d). This is due to the decreasing sample set size and to the reduced number of available gages. Despite those sources of uncertainties, $\hat{\alpha}_2$, $\hat{\alpha}_4$ and $\hat{\alpha}_8$ remain approximately constant in the sub-region corresponding to the lowest altitudes. On the contrary, the fluctuations seem more consistent near the mountain ridge and in the northwest plateau. This evidence validates the scaling behavior of heavy rainfall for short duration in this flat sub-region in agreement with *Hubert and Bendjoudi* [1996]. In the sub-region at the North of the Mont Lozère, α varies consistently with duration. The simple scaling hypothesis does not hold in this area.

Figure 10 shows the map of x_{min} for $D=1h$. In the plain region and over the Massif Central, the lower bound x_{min}

is the lowest ($\sim 4 \text{ mm h}^{-1}$) and increases towards the North-West over the mountain slope until the ridge up to about 9 mm h^{-1} and towards the North up to 8 mm h^{-1} . Mountainous and north-eastern raingages show the smallest proportion of events lying in the hyperbolic part of the distribution. This is the main evidence of the effect of orography on rainfall from the point of view of the probability distribution.

5. Conclusive remarks

The paper has shown that an objective method can be used to characterize the heavy rainfall distribution featuring hyperbolic tails. The objective and unbiased determination of the power-law exponent α is necessary for the regionalization of the power-law behavior of rainfall series. This process has been repeated at different accumulation periods, leading to the definition of a sub-region where the power-law exponent is approximately constant. Considering the relatively short observation period (16 years), we assume that no significant trends affecting the stationarity of the raingage series are present.

In the next sub-section, we will interpret the results from the point of view of the physical processes generating rainfall within the study region, and we will discuss the current findings regarding the statistical analysis of extreme rainfall events.

5.1. Physical interpretation of the results

The behavior of rainfall distribution tails is heterogeneous in the region of interest. The regionalization of the rainfall variability (α values, Fig. 7), the agreement between the power-law model and the tail distribution shape (confidence interval, Fig. 8) and the proportion of observations concerned by the power-law behavior (x_{min} values, Fig. 10) delineate the differences between the flat area and the mountainous region. In the South-East sub-region (between Alès, Nîmes and Montpellier) the rainfall variability, i.e. the ratio between maximum and average rainfall, is the highest at short durations: it is the area where the power-law model is the most adapted (lower confidence interval in Fig. 8) and the number of observations exhibiting hyperbolic tails is the highest (lower x_{min} in Fig. 10). The power-law exponent in the region shows a relief-oriented gradient: both the Central Massif and the Alps exhibit high α , corresponding to lower rainfall variability compared to the flat areas. The signature of the Rhône Valley is sharp for small accumulation periods, decreasing for high accumulation periods.

Several studies [Sénési et al., 1996; Ducrocq et al., 2002; Ricard, 2002; Ducrocq et al., 2003; Delrieu et al., 2005; Nuissier et al., 2008] have shown that the heaviest rainfall are yielded by mesoscale-systems entering the region from South and South-East. Grossly speaking, the relatively warm and humid air masses coming from the Mediterranean sea are lifted upward by an orographic barrier, the Massif Central slopes, and by thermodynamical mechanisms (cold pool, Nuissier et al. [2008]) which block the heaviest rainfall in the South-East of the study region.

The South-North gradient, displayed by the statistical properties of heavy rainfall in the Northern part of the study area especially for accumulation periods higher than 2 hours, is less linked to the relief. Both valley and mountain slopes are present in the region. The lack of references concerning the rainfall events occurring in this region allows only hypothetical reasoning. The average increase of α values in this region corresponds to a general decrease of rainfall variability compared to the southern portion. An interpretation may be that, besides the relief effect, the distance from the storm-triggering zone plays an important role on the weakening of the storm convection, due to the ground friction. In conclusion, the sheltering effect generated by the relief is not the only factor limiting the rainfall variability.

5.2. Consequences for the extreme modeling

The representativity of the power-law exponent α for the description of the variability of heavy rainfall is demonstrated by two main results. First, α has a well-determined spatial structure (Fig. 7). The interpolation process has been easily performed since the variable has a definite empirical variogram at any duration. Interesting properties of α are that whatever the accumulation duration, α is always lower in the flat area; in addition, in the Southern portion of the domain, α is approximately constant with the accumulation duration, satisfying the necessary conditions for the simple scaling of heavy rainfall. As recalled in Section 1, the cumulative probability distributions of extreme rainfall intensities are usually modeled either by Generalized Extreme Value (GEV) or Generalized Pareto Distributions (GPD) [Kottegoda and Rosso, 1997; Sornette, 2004] depending on the selection of heavy rainfall events (maxima or Peaks Over Threshold). Depending on their parameters, both the GEV and GPD distributions may display exponential or power-law tails. Extreme rainfall analyses related to design rainfall assessment in the South-East of France [Guillot and Duband, 1967; Slimani and Lebel, 1986; Nguyen Thao et al., 1993; Cernesson et al., 1996] and elsewhere Zhang and Singh [2007], for instance, assumed that infra-daily extreme rainfall intensities follow Gumbel distributions (GEV-I) i.e. exponential survival probability tails. Thick tail distributions have been found in space-time rainfall fluctuations [Kumar and Foufoula-Georgiou, 1993; Perica and Foufoula-Georgiou, 1996].

Koutsoyiannis [2003] pointed out some reasons of the GEV-I popularity in hydrology: design-rainfall studies are based on maxima analyses, GEV-I exhibit linearity on Gumbel diagrams. on the log-log plot in Figure 2 Moreover, GEV-II (Fréchet) distribution has one additional parameter respect to GEV-I, giving larger uncertainties with limited sample sets. However, several studies reported that GEV-I underestimates actual extreme rainfall intensities (see Koutsoyiannis [2003] for a detailed description). In this study most of raingages shows hyperbolic tails at various durations: this is an evidence of the Fréchet behavior of maxima; the exponential behavior of survival probability tails typical of Gumbel distribution is not in agreement with our findings concerning the Southern part of the study region. Considering the proportion of hyperbolically distributed samples at the accumulation time of 1 hour, we observed that in most of raingages they are no more than 5 % of the positive rainfall values. This value corresponds to the 0.1 - 0.3% of the whole observations, meaning that in a year, between 8 and 20 observations are hyperbolically-distributed. However, the strong inter-dependence of most of these values further limits the number of independent data lying in the hyperbolic tail. In addition, increasing the accumulation period, the number of observations per year decreases, reducing the number of hyperbolically-distributed samples. In the case of the 24-h accumulation period, we should observe in average a hyperbolic sample every three years. In our opinion, this explains why a limited sample set of daily data better fits a GEV-I distribution rather than a GEV-II one.

The most important result in this paper is the assessment of the variability of α with duration (Fig. 11). The variance among the α values at 1h, 2h, 3h, 4h is used as indicator of the variability of the temporal scaling properties of heavy rainfall. The 8-hour accumulation period has not been considered in this computation: many stations have been discarded due to the poor sample set size. The lower the variance, the higher the reliability in the temporal

simple-scaling behavior. Figure 11 shows that, in a large sub-region covering both the flat area and a portion of the foothill (Southern part of the study region), the variability of α with the accumulation period is small (variance lower than 0.06). In this zone α can be considered as a constant with respect to duration. Since in the same sub-region the confidence interval of the power-law estimation is lower than 20% of the value of α (Fig. 8), in these areas the rainfall distribution can be considered approximately self-similar in the power-law distributed part. In the framework of the usual extreme value analysis, this means that the “extreme value index” ξ , (which is 0 in GEV-I and equal to $\frac{1}{\alpha}$ in GEV-II), would always be higher than 0 in this sub-region and, even more noteworthy, it is constant with duration. In regions where time simple-scaling of heavy rainfall holds, the derivation of one of the three parameters of GEV-II by means of scale-invariance relations can therefore determine a considerable improvement in the fitting of limited samples series with GEV-II distribution.

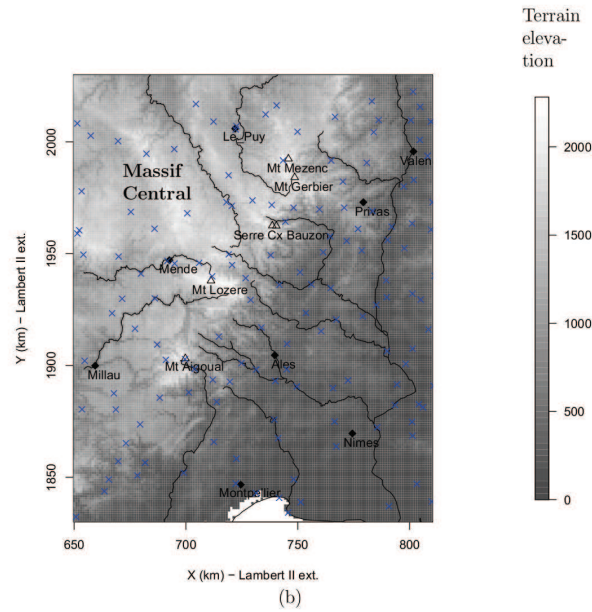
Acknowledgments. This work has been supported by the Vulnerability, Environment and Climate (VMC) program of the French National Agency for Research (ANR). The project is entitled “Forecast and Projection in Climate Scenario of Mediterranean Intense Events : Uncertainties and Propagation on Environment” (MedUP). Data have been collected by the French weather service Météo-France. The authors are grateful to the reviewers for their constructive remarks.

References

- Bendjoudi, H., P. Hubert, and D. Schertzer, Interpretation multifractale des courbes intensité-durée-fréquence des précipitations, *C. R. Acad. Sci. Paris*, 2, 323–326, 1997.
- Borga, M., C. Vezzani, and G. Dalla Fontana, Regional rainfall depth-duration-frequency equations for an alpine region, *Natural Hazards*, 36(1-2), 221–235, 2005.
- Burlando, P., and R. Rosso, Scaling and multiscaling models of depth-duration-frequency curves for storm precipitation, *Journal of Hydrology*, 187, 45–64, 1996.
- Cernesson, F., J. Lavabre, and J. Masson, Stochastic model for generating hourly hyetographs, *Atmos. Res.*, 42, 149–161, 1996.
- Chiles, J., and P. Delfiner, *Geostatistics modeling spatial uncertainty*, New York: Wiley, 1999.
- Clauset, A., C. Shalizi, and M. Newman, Power-law distributions in empirical data, *SIAM Review*, (51), 661–703, 2009.
- Delrieu, G., L’Observatoire Hydro-météorologique Méditerranéen Cévennes-Vivarais (the Cévennes-Vivarais Mediterranean Hydro-meteorological Observatory), *La Houille Blanche*, 6(1), 83–88, 2004.
- Delrieu, G., et al., The catastrophic flash-flood event of 8–9 september 2002 in the Gard region, France: a first case study for the Cévennes-Vivarais Mediterranean Hydro-meteorological Observatory, *J. Hydrol.*, 6(1), 34–52, 2005.
- Ducrocq, V., D. Ricard, J. Lafore, and F. Orain, Storm-scale numerical rainfall prediction for five precipitating events over France: on the importance of the initial humidity field, *Weather Forecasting*, 17, 1236–1256, 2002.
- Ducrocq, V., G. Aullo, and P. Santurette, Précipitations intenses et les inondations des 12 et 13 novembre 1999 sur le sud de la France, *La Météorologie*, (42), 18–27, 2003.
- Frisch, U., and G. Parisi, Fully developed turbulence and intermittency, *Turbulence and Predictability in Geophysical Fluid Dynamics and Climate Dynamics*, pp. 84–88, 1985.
- Goldstein, M., S. Morris, and G. Yen, Problems with fitting to the power-law distributions, *European Physics Journal*, B(41), 255–258, 2004.
- Guillot, P., and D. Duband, The gradex to compute extreme flood probability from the observation of rainfall, in *International Hydrology Symposium*, 1, 506–515, Colorado State University, Fort Collins, Colorado, 1967.
- Gumbel, E., *Statistics of Extremes*, Columbia University Press, New York, 1958.
- Gupta, V., and E. Waymire, Multiscaling properties of spatial rainfall and river flow distributions, *Journal of Geophysical Research*, 95(D3), 1999–2009, 1990.
- Hubert, P., and H. Bendjoudi, Introduction à l’étude des longues séries pluviométriques, in *Actes des XIIèmes Journées Hydrologiques de l’Orstom*, Montpellier, 1996.
- Hubert, P., Y. Tessier, S. Lovejoy, P. Ladoy, J. Carbonner, S. Violette, and I. Desurosne, Multifractals and extreme rainfall events, *Geophysical Research Letters*, 20(10), 931–934, 1993.
- Jacq, V., Inventaire des situations à précipitations diluviennes sur les régions Languedoc-Roussillon, PACA et Corse, période 1958–1994 (inventory of extreme precipitation events in the Languedoc-Roussillon, PACA and Corse regions in the period 1958–2004). phénomènes remarquables., *Tech. Rep. 3*, Météo-France (SCEM), 1994.
- Kottegoda, N., and R. Rosso, *Probability and Reliability for Civil and Environmental Engineers*, Blackwell Pub, 1997.
- Kotz, S., and S. Nadarajah, *Extreme value distributions. Theory and applications*, Imperial College Press, 2000.
- Koutsoyiannis, D., On the appropriateness of the gumbel distribution for modelling extreme rainfall, in *ESF LESC Exploratory Workshop*, Bologna, Italy, 2003.
- Koutsoyiannis, D., D. Kozonis, and A. Manetas, A mathematical framework for studying rainfall intensity - duration - frequency relationships, *Journal of Hydrology*, 206, 118–135, 1998.
- Kumar, P., and E. Foufoula-Georgiou, A new look at rainfall fluctuations and scaling properties of spatial rainfall using orthogonal wavelets, *Journal of Applied Meteorology*, 32, 209–222, 1993.
- Ladoy, P., F. Schmitt, D. Schertzer, and S. Lovejoy, Variabilité temporelle multifractale des observations pluviométriques à Nîmes, *C.R. Acad. Sci. Paris*, 317(2), 775–782, 1993.
- Menabde, M., A. Seed, and G. Pegram, A simple scaling model for extreme rainfall, *Water Resources Research*, 35(1), 335–339, 1999.
- Nguyen Thao, T. P. N. G., P. Bois, and J. A. Villasenor, Simulation in order to choose a fitting method for extreme rainfall, *Atmos. Res.*, 30, 13–36, 1993.
- Nuissier, O., V. Ducrocq, D. Ricard, C. Lebeaupin, and S. Anquetin, A numerical study of three catastrophic precipitating events over western mediterranean region (Southern France): Part I: Numerical framework and synoptic ingredients, *Quart. J. Roy. Meteor. Soc.*, 134(630), 111–130, 2008.
- Perica, S., and Foufoula-Georgiou, Model for multiscale disaggregation of spatial rainfall based on coupling meteorological and scaling descriptions, *Journal of Geophysical Research*, 101(D21), 26,347–26,361, 1996.
- Ricard, D., Initialisation et assimilation de données à méso-échelle pour la prévision à haute résolution des pluies intenses de la région cévennes-vivarais (in french), Ph.D. thesis, Université Pierre Sabatier-Toulouse III (France), 2002.
- Schertzer, D., and S. Lovejoy, Physically based rain and cloud modeling by anisotropic, multiplicative turbulent cascades, *Journal of Geophysical Research*, 92, 9692–9714, 1987.
- Schertzer, D., and S. Lovejoy, Hard and soft multifractal processes, *Physica A*, 185, 187–194, 1992.
- Sénési, S., P. Bougeault, J. Chèze, P. Cosentino, and R. Thepenier, The vaison-la-romaine flash flood: Mesoscale analysis and predictability issues, *Weather Forecasting*, 11(4), 417–442, 1996.
- Sivakumar, B., Fractal analysis of rainfall observed in two different climatic regions, *Hydrological Sciences Journal*, 45(5), 727–738, 2000.
- Slimani, M., and T. Lebel, Comparison of three methods of estimating rainfall frequency parameters according to the duration of accumulation., in *International symposium on flood frequency and risk analyses*, vol. Hydrologic frequency modeling, pp. 277–291, Singh, J.V.; Reidel, D. (eds.) ; D. Reidel Publishing Company, 1986.
- Sornette, D., *Critical Phenomena in Natural Sciences*, Springer-Verlag Berlin, 2004.
- Tessier, Y., S. Lovejoy, and D. Schertzer, Universal multifractals: Theory and observations for rain and clouds, *Journal of Applied Meteorology*, 32(2), 223–250, 1993.
- Zhang, L., and V. P. Singh, Gumbel–hougaard copula for trivariate rainfall frequency analysis, *Journal of Hydrologic Engineering*, 12(4), 409–419, doi:10.1061/(ASCE)1084-0699(2007)12:4(409), 2007.



(a)



(b)

Figure 1. a: Localisation of the region of interest. b: Elevation map (gray shaded area in m above sea level) in the region of interest. The crosses indicate the hourly raingage network. The full line indicates the main hydrographic network. The main river in the region is the Rhône river. It roughly represents the eastern boundary of the region. The Mediterranean shore is the southern boundary and the mountain ridge, oriented North-Northwest, is the southern limit of the Massif Central plateau.

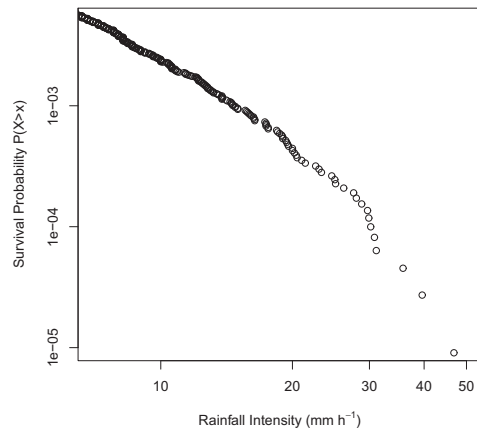
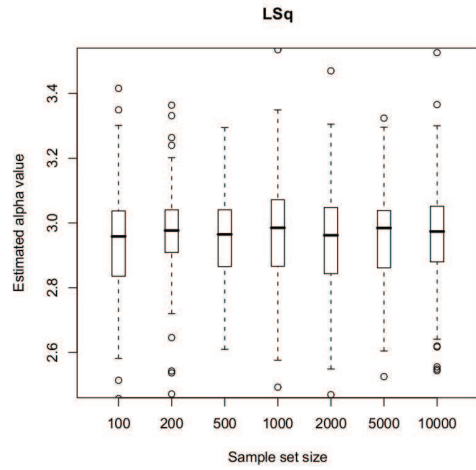
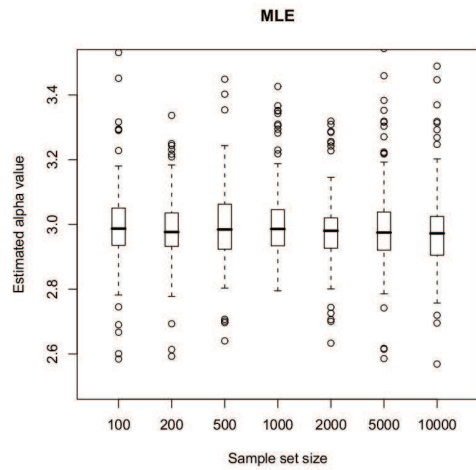


Figure 2. Log-log plot of the survival probability distribution of hourly rainfall for the raingage station of Cognac, France.



(a)



(b)

Figure 3. Box-plots of the distribution of $\hat{\alpha}$ using the Least Squares Fitting (a) and the Maximum Likelihood Estimator (b) on 100 samples of different size. All the samples were distributed following a Pareto distribution, $x_{min}=10$, $k=3$.

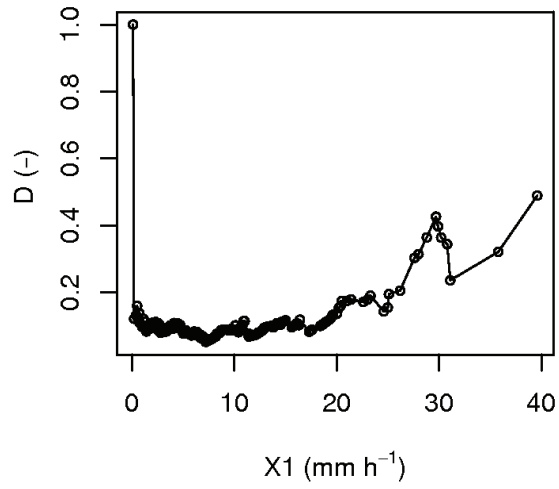


Figure 4. Plot of the Kolmogorov-Smirnov D statistics as a function of the lower bound x_1 for the hourly rain gauge station of Cognac, France (see Fig. 2). The minimum of D corresponds to x_{min} , which is used in turn to determine α .

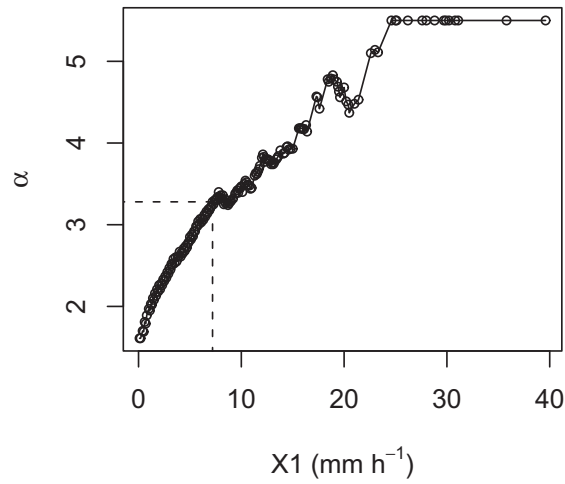


Figure 5. Power-law exponent α as a function of the lower bound x_1 for the rain gauge station of Cognac, France (see Figure 2).

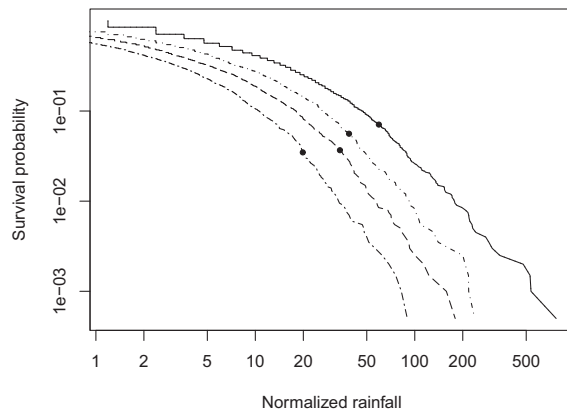
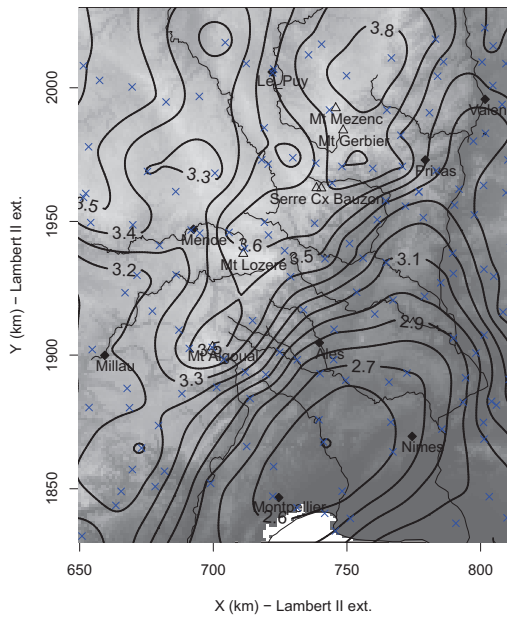
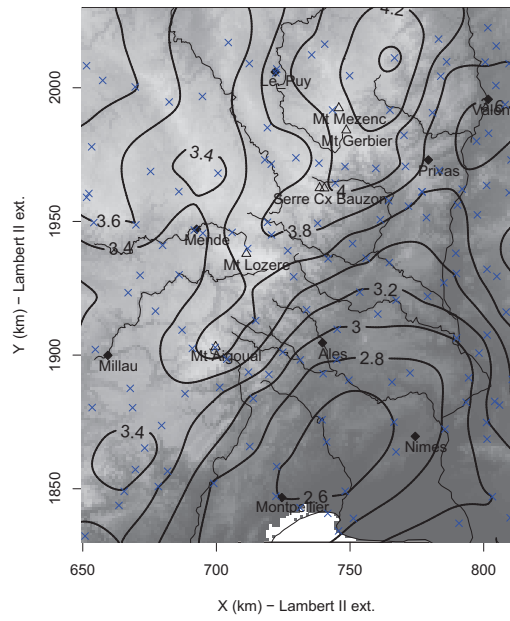


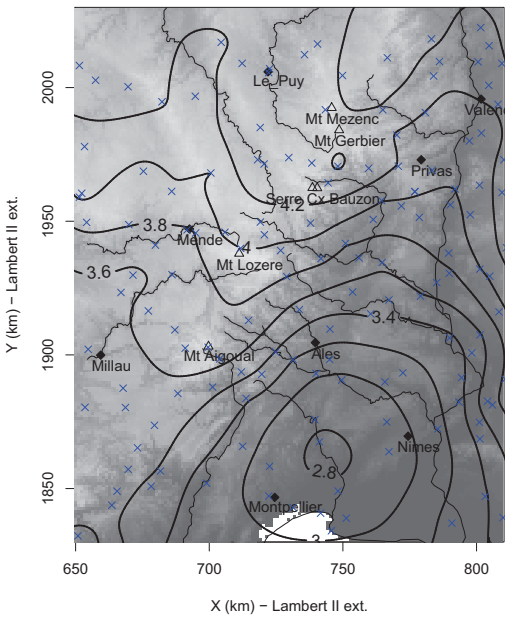
Figure 6. Log-log plot of survival probability of the normalized rainfall rate for durations of 1 h (solid line), 4 h (dash-dotted line), 10 h (dashed line) and 24 h (doubly dashed line) hours at the raingage station of Montpellier - Bel-Air. A solid circle marks the lower bounds x_{min} .



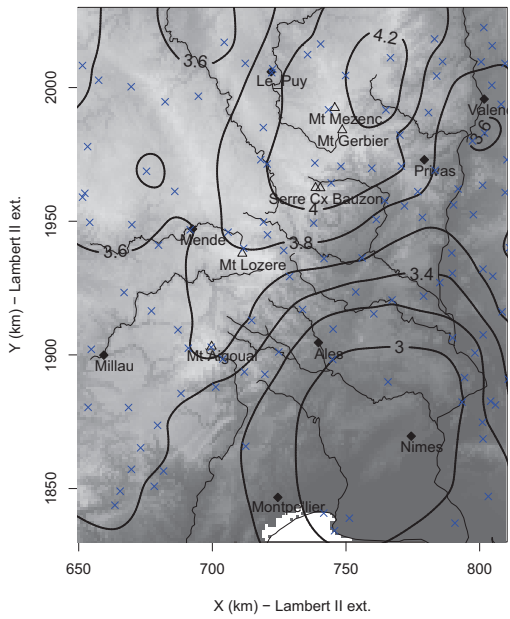
(a)



(b)



(c)



(d)

Figure 7. Power-law exponent ($\hat{\alpha}$) map in the region of interest for different accumulation periods. a: $D=1h$. b: $D=2h$. c: $D=4h$. d: $D=8h$. The crosses represent the considered raingauge network at the corresponding duration. See Figure 1 for details on the background.

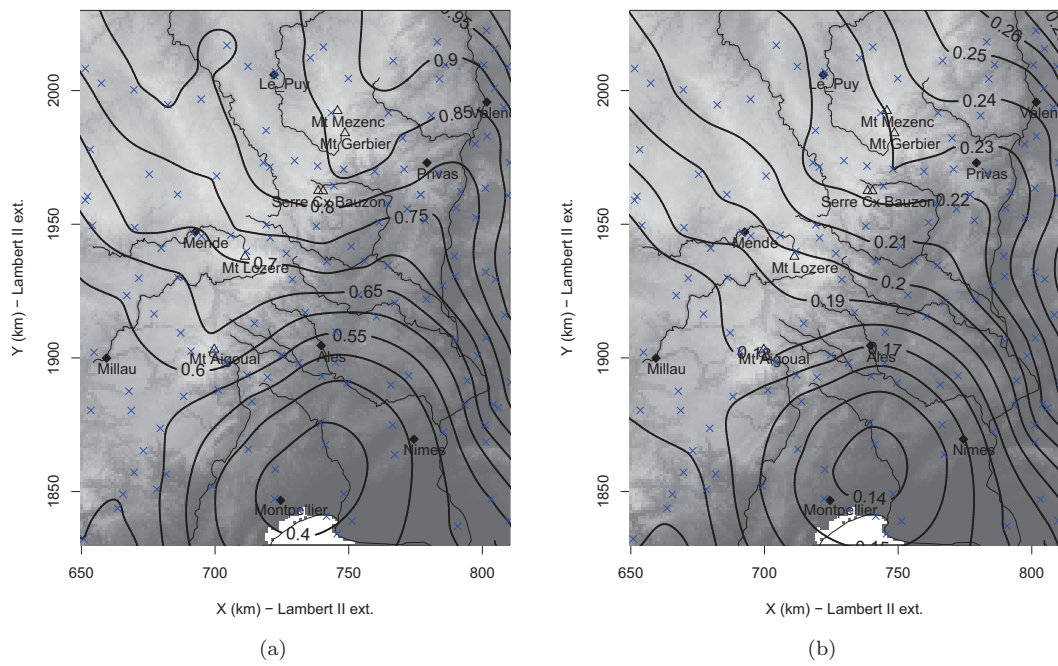


Figure 8. Confidence interval of $\hat{\alpha}_1$ for the 95% confidence level. a: absolute confidence interval; b: relative to $\hat{\alpha}_1$ confidence interval. The maps have been obtained by kriging interpolation.

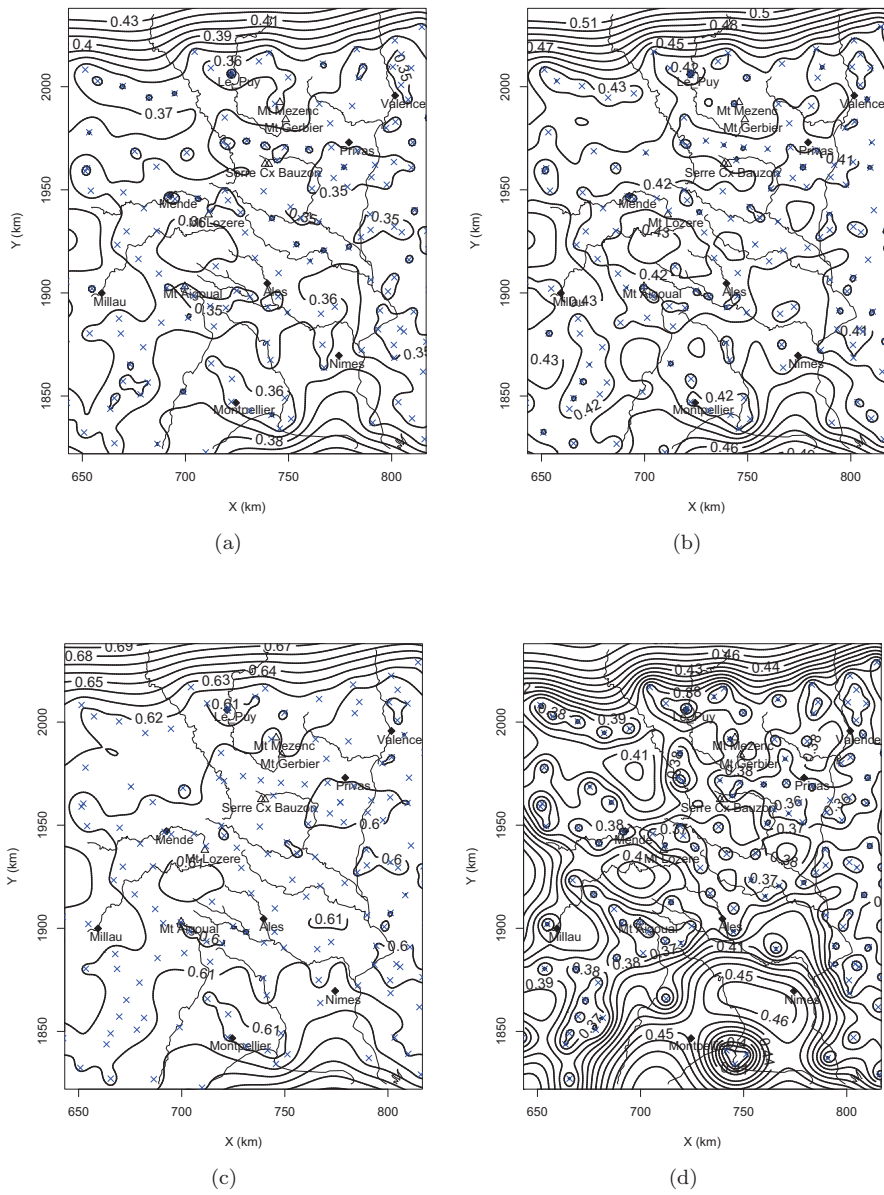


Figure 9. Kriging standard deviation map for the α exponent. a: $D=1h$. b: $D=2h$. c: $D=4h$. d: $D=8h$.

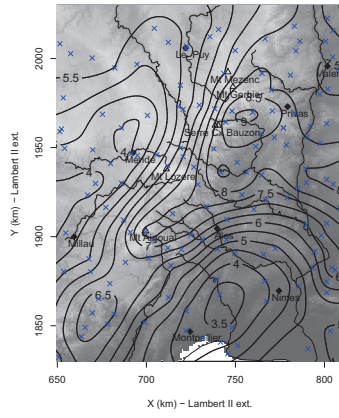


Figure 10. Map of the power-law lower bound \hat{x}_{min} for $D=1h$.

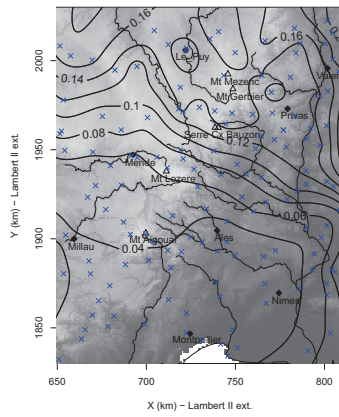


Figure 11. Stability of the power-law exponent: variance of the $\hat{\alpha}$ values at 1, 2, 3, 4 h.

A scale-invariance Intensity-Duration-Frequency model

Résumé

*Dans cet article nous appliquons un modèle invariant d'échelle pour l'estimation des précipitations extrêmes dans une région du sud de la France. Nous utilisons des observations de pluie journalière pour montrer que la distribution des maxima ponctuels peut présenter à la fois des queues exponentielles et hyperboliques, et par conséquent elle s'adapte à la distribution des valeurs extrêmes généralisée (GEV) et non pas à une plus restrictive distribution de Gumbel. Grâce aux pluviomètres à enregistrement (horaires), nous montrons que les précipitations maximales dans la zone d'étude sont invariantes d'échelle au moins dans la gamme de durées 4 h à 100 h, et que l'invariance d'échelle est du type “**simple scaling**”. Le modèle couple les courbes intensité-durée-fréquence (IDF) avec la distribution généralisée des valeurs extrêmes. Le paramètre de forme de la loi des extrêmes (constante en “simple scaling”) a une influence importante dans la détermination des précipitations pour les périodes de retour élevées. L'amélioration de son estimation passe par l'adoption d'une méthode des excès (**Peaks-Over-Threshold**). Les quatre paramètres du modèle sont ensuite cartographiés, et l'estimation des pluies extrêmes pour la période de retour de 50 et 100 ans est évaluée. Nous mettons en évidence les différences avec le modèle Gumbel simple scaling. Nous discutons enfin sur les phénomènes météorologiques sous-jacentes qui peuvent mener au comportement des précipitations extrêmes dans la région d'étude.*

8.1 Introduction

In this paper we apply a scale-invariant model for the estimation of rainfall maxima in a region of Southern France. We use daily rainfall observations to show that the point distribution of rainfall maxima over the study region can present exponential or hyperbolic tails, and consequently fit with the Generalized Extreme Value (GEV) distribution framework (see also *Ceresetti et al. (2011)*). Thanks to recording (hourly) rain gauges, we show that rainfall maxima in the study region display the simple-scaling property at least over the range 4 h- 100 h. The model derives Intensity-Duration-Frequency

(IDF) curves coupling the temporal scale-invariance with the Generalized Extreme Value distribution. The shape parameter (a constant in simple scaling) has large influence in the rainfall determination for large return periods. We improve its estimation by adopting a Peaks-Over-Threshold method. The 4 model parameters are then mapped, and the return level rainfall intensities for the return period of 50 and 100 years are computed and analyzed, highlighting the differences with the Gumbel simple scaling model. We finally discuss about the underlying meteorological phenomena that may lead to the extreme rainfall behavior over the region of study.

The paper relies on the considerations concerning the self-similarity of random variables (Section 5.5.2) and the scaling of Intensity-Duration-Frequency curves (Section 5.11). To facilitate the article reading, in the following we define L-moments. The estimation by L-moments (*Hosking, 1990*) is more adapted than classical moments when dealing with high-variability observations such as rainfall maxima. L-moments estimation, in fact, only requires that the mean of the distribution is finite; the higher-order moments do not need to be finite. In addition, the computation of standard error only requires that the distribution has finite variance. Furthermore, L-moments are less sensitive to outlying data values (*Vogel and Fennessey, 1993*).

The L-moments are defined (*Hosking, 1990*) for a real valued random variable X with cumulative distribution function $F(x)$, for a random sample of size n drawn from the distribution of X , as:

$$L_r = r^{-1} \sum_{k=0}^{r-1} (-1)^k \binom{r-1}{k} E[X_{r-k:r}] \quad (8.1)$$

where $X_{1:n} \leq X_{2:n} \leq \dots \leq X_{n:n}$ is the order statistics of the random sample and $r = 1, 2, \dots$ is the moment order.

The first 4 L-moments are:

$$\begin{aligned} L_1 &= E[X] = \int_0^1 x(F) dF \\ L_2 &= \frac{1}{2} E[X_{2:2} - X_{1:2}] = \int_0^1 x(F) (2F - 1) dF \\ L_3 &= \frac{1}{3} E[X_{3:3} - 2X_{2:3} + X_{1:3}] = \int_0^1 x(F) (6F^2 - 6F + 1) dF \\ L_4 &= \frac{1}{4} E[X_{4:4} - 3X_{3:4} + 3X_{2:4} - X_{1:4}] = \int_0^1 x(F) (20F^3 - 30F^2 + 12F - 1) dF \end{aligned} \quad (8.2)$$

The L-moment ratios τ_r are defined as

$$\tau_r = \frac{L_r}{L_2} \quad (8.3)$$

The Gumbel parameters are given as a function of the L-moments (*Hosking, 1990*):

$$\begin{aligned} \mu &= L_2 / \log(2) \\ \sigma &= L_1 - \gamma\sigma \end{aligned} \quad (8.4)$$

The GEV estimation requires the use of the Gamma function Γ :

$$\begin{aligned} \mu &= L_1 + \sigma \frac{1 - \Gamma(1 - \xi)}{\xi} \\ \sigma &= \frac{-L_2 \xi}{(1 - 2\xi) \Gamma(1 - \xi)} \end{aligned} \quad (8.5)$$

where the Gamma function is defined as $\Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t} dt$.

The ξ parameter can be estimated from a data sample by inverting the equation in terms of the third L-moment ratio τ_3

$$\tau_3 = 2 \frac{(1 - 3^\xi)}{1 - 2^\xi} - 3 \quad (8.6)$$

The latter equation can not be explicitly expressed as a function of ξ , however an approximate formula exists.

8.2 Article: Intensity-Duration-Frequency curves in a GEV scale-invariant framework

Submitted to Water Resources Research, September 2010

Intensity - Duration - Frequency curves in a GEV scale-invariant framework

D. Ceresetti, G. Molinié, J.-D. Creutin

LTHE, Laboratoire d'Etudes des Transferts en Hydrologie et Environnement - (CNRS,UJF,IRD,INPG), Grenoble, France

Abstract.

In this paper we apply a scale-invariant model for the estimation of rainfall maxima in a region of Southern France. We use daily rainfall observations to show that the point distribution of rainfall maxima over the study region can present exponential or hyperbolic tails, and consequently fit with the Generalized Extreme Value (GEV) distribution framework. Thanks to recording (hourly) raingages, we show that rainfall maxima in the study region display simple-scaling property at least over the range 4 h- 100 h. The model derives Intensity-Duration-Frequency (IDF) curves coupling the temporal simple-scaling with the Generalized Extreme Value distribution. The shape parameter (a constant in simple scaling) has large influence in the rainfall determination for large return periods. We improve its estimation by adopting a Peaks-Over-Threshold method. The four model parameters are then mapped, and the estimation of the extreme rainfall for the return period of 50 and 100 years is presented, highlighting the differences with the Gumbel simple scaling model. We finally discuss about the underlying meteorological phenomena that may lead to the extreme rainfall behavior over the region of study.

1. Introduction

The problem of the frequency estimation of extreme hydro-meteorological events needs very long and stationary series that are seldom available. Climate change (anthropogenic or natural) gives doubts about the stationarity of existing data bases. Practical problem always made the availability of such data bases somewhat problematic. The appropriate temporal scales for analyzing rainfall events are not necessarily in agreement with the temporal resolution of the rainfall records. The almost exclusive use of daily raingages around the world makes impossible to correctly assess the frequency of infra-daily events. In regions where long series of infra-daily rainfall intensities are available, empirical determinations of IDF relationships for different durations can lead to incoherence between durations [Borga *et al.*, 2005], due to sampling effects on the aggregation of rainfall intensities.

In the 90's, following the dissemination of fractal science in the hydro-meteorological research community, rainfall scale invariance issues have been addressed by various authors [Schertzer and Lovejoy, 1987; Hubert and Carbonnel, 1989; Ladoy *et al.*, 1993; Marsan *et al.*, 1996; Bendjoudi *et al.*, 1997; Venugopal *et al.*, 1999; Deidda *et al.*, 1999; Harris *et al.*, 2001]. Pragmatic studies [Burlando and Rosso, 1996; Menabde *et al.*, 1999; Veneziano and Furcolo, 2002; Borga *et al.*, 2005] have used the temporal scale-invariance property of rainfall series to infer analytical expressions of IDF relationships. Burlando and Rosso [1996] and Bendjoudi *et al.* [1997] were the firsts to demonstrate that the usual IDF relationships were expressions of the rainfall scale invariance. At about the same time, thanks to the extreme value (EV) theory, Koutsoyiannis *et al.* [1998] described the mathematical framework for studying rainfall Intensity-Duration-Frequency relationships. Menabde *et al.*

[1999] showed that the simple-scaling framework (a particular case of the multifractal one) could be coupled with IDF relationships, and presented a scaling IDF model assuming Gumbel-distributed (EV-I) rainfall maxima.

Recently, the availability of long series allowed to show that rainfall maxima may exhibit tails thicker than the Gumbel model [Koutsoyiannis, 2003]. In a previous paper we showed the need to generalize the behavior of rainfall maxima in the framework of the Generalized Extreme Value (GEV) distribution in the study region [Ceresetti *et al.*, 2010]. In this paper we propose a derivation of IDF relationships in the context of GEV-distributed maxima by means of a scale-invariant approach. Practically, we verify the simple-scaling assumption for durations in the range 1-100 h using hourly observations; then, using a daily rainfall database, we determine the parameters of the scale-invariant model.

The paper is organized in six sections. Section 2 focuses on the classical expression of Intensity - Duration - Frequency curves and on the implicit scale-invariance assumption that this model implies taking a fixed return period. After a description of the simple-scaling EV-I (Gumbel) model for IDF curves (Section 2.1), the formulation of a simple-scaling GEV model is presented including a L-moments formulation (Section 2.2). A brief overview of the study region and of the raingage networks follows in Section 3. The GEV scale-invariant model is applied to the raingage network in Section 4 with the aim to provide a continuous description of the extremes behavior for accumulations over 1 and 100 h. In Section 5 a discussion concerning the advantages of this parsimonious model, the implications and limits of the simple-scaling assumption are reported, together with the comparison of rainfall maps built with the Gumbel and the GEV model for return periods T_R of 50 and 100 years.

2. Intensity - Duration - Frequency curves and scale-invariance

Intensity - Duration - Frequency curves relate the rainfall intensity I , the duration over which the intensity is considered D , and the corresponding frequency of occurrence F

otherwise expressed in terms of return period $T_R = \frac{1}{1-F}$. IDF curves are devoted to the estimation of design rainfall or flows. Empirical IDF curves are derived by fitting the rainfall intensity relative to the same return period as a function of the temporal scale. Following this methodology, *Bernard* [1932] showed that IDF relationships have similar behavior in different regions of the world. Since then, IDF expressions based on empirical analysis were adopted in engineering design. More than 60 years later, *Koutsoyiannis et al.* [1998] provided a mathematical derivation of IDF relationships using dimensional arguments:

$$I_{D,T_R} = a(T_R)D^{n(T_R)} \quad (1)$$

where $a(T_R)$ and $n(T_R) < 0$ are two parameters depending only on the return period T_R . Equation (1) expresses that rainfall intensity decreases when duration increases, for a fixed return-period.

To obtain analytical IDF expressions for any return period, the scaling properties of rainfall intensity must be coupled to an extreme-value distribution (block-maxima or peaks-over-threshold). A number of models coupling extreme-value distributions and scale invariance of rainfall maxima have been proposed [*Burlando and Rosso*, 1996; *Bendjoudi et al.*, 1997; *Menabde et al.*, 1999; *Borga et al.*, 2005]. These models follow the concept of scaling proposed by *Gupta and Waymire* [1990] that we briefly recall hereafter.

A process Y is strict sense scaling when its probability distribution at the scale λ can be derived from the distribution of the process at a reference scale λ_{ref} through the relationship:

$$Y_\lambda \stackrel{d}{=} \left(\frac{\lambda}{\lambda_{ref}} \right)^{-H} Y_{\lambda_{ref}} \quad (2)$$

When the scaling concept is applied to the maximum rainfall intensity I ($mm h^{-1}$) as a function of the duration D (h), Equation (2) becomes:

$$I_D \stackrel{d}{=} \left(\frac{D}{D_{ref}} \right)^{-H} I_{D_{ref}} \quad (3)$$

Gupta and Waymire [1990] report that the scaling in distribution (strict sense) implies a similar relation for the statistical moments (property known as “wide-sense simple scaling”):

$$E[I_D^q] = \left(\frac{D}{D_{ref}} \right)^{-qH} E[I_{D_{ref}}^q] \quad (4)$$

where q is the moment order. Equation (4) is valid if the moments of the maximum rainfall intensity are defined. If the scaling exponent H depends on q , the process is said “simple scaling” otherwise it is “multiple scaling”.

A “wide sense” scaling relation is also available for quantiles [*Burlando and Rosso*, 1996]:

$$I_{D,k} = \frac{D}{D_{ref}}^{-H} I_{D_{ref},k} \quad (5)$$

where k is the k -th quantile.

The identification of Equations (1) and (5) for a given quantile leads to $a(T_R) = I_{D_{ref},T_R}$. Thus $a(T_R)$ is the rainfall intensity for the reference duration D_{ref} and the return period $T_R = \frac{1}{1-F}$. The identification also indicates that

$n(T_R) = H$. It follows that if the scaling exponent n of Equation (1) varies with the probability level F (i.e. with the return period T_R) we have multiple scaling and, if it is constant, we have simple scaling. In both cases, the plot I versus D displays straight lines in log-log scales (one for each T_R).

In order to explicit the return-period dependence in Equation (1), we introduce an extreme-value distribution for the maximum rainfall intensity. The distributions of independent and identically distributed (i.i.d.) maxima can be modeled with the Generalized Extreme Value density function (GEV, *Kotz and Nadarajah* [2000]). Three particular expressions (EV-I, EV-II and EV-III) are derived from the GEV following the random variable property. The most relevant property of GEV distribution in this context is that its third and fourth moments (skewness and kurtosis) only depend on the shape parameter and are constant in EV-I, i.e. the distribution is unable to modify its shape according with the multi-scaling framework; this implies incompatibility of GEV with the multi-scaling modeling of rainfall maxima.

Rainfall maxima show slight deviations from the simple scaling behavior; *Burlando and Rosso* [1996] proposed a multi-scaling model of log-normally distributed maxima. This model is adapted to define Intensity-Duration-Frequency curves for small return periods, where the asymptotic conditions underlying the extreme-value theorem are not completely fulfilled.

In our application, the deviation from simple scaling is small, as shown in Section 2.2. Therefore, a theoretically consistent simple-scaling model will be adopted. In this framework, the shape parameter remains constant whatever the scale.

In the two following paragraphs, IDF formulations are derived in the simple-scaling framework, for Gumbel (Section 2.1) and, more generally, GEV (Section 2.2) distributed maxima.

2.1. Gumbel Simple-scaling IDF model

Let us assume that rainfall maxima follow an EV-I (Gumbel) distribution and feature temporal simple-scaling. The moment scaling function $K(q)$ of the statistical raw moments (Equation 4) is a linear function of the scaling exponent H and of the moment order q :

$$K(q) = H \cdot q \quad (6)$$

where H is a constant in simple scaling. The Gumbel distribution, whose cumulative function is

$$F(x, \mu, \sigma) = e^{-e^{-(x-\mu)/\sigma}} \quad (7)$$

is defined by two parameters, the location μ and the scale σ . As stated in *Menabde et al.* [1999]; *Borga et al.* [2005], the strict sense scaling relationship (Equation 4) allows to determine the distribution parameters at any duration D from a first guess at a reference duration D_{ref} :

$$\mu_D = \left(\frac{D}{D_{ref}} \right)^{-K(1)} \mu_{D_{ref}} \text{ and } \sigma_D = \left(\frac{D}{D_{ref}} \right)^{\frac{-K(2)}{2}} \sigma_{D_{ref}} \quad (8)$$

In case of simple scaling of rainfall maxima, Equation (6) applies and then $K(1) = K(2)/2$. This implies that the coefficient of variation CV, ratio of the standard deviation to the mean, is independent of the duration. From a practical

point of view, the Gumbel parameters μ and σ are estimated thanks to the sample mean $m(x)$ and sample standard deviation $s(x)$ following the relations:

$$\hat{\mu} = m(x) - \gamma\hat{\sigma} \quad \hat{\sigma} = \frac{s(x)\sqrt{6}}{\pi} \quad (9)$$

where $\gamma \sim 0.5772$ is the Euler constant.

Combining Equations (8), (9) and knowing the mean and standard deviation of the maxima at a particular scale allow to estimate the Gumbel parameters at any duration. In particular, if simple scaling holds, the maximum rainfall intensity at any duration D and return period T_R can be expressed by [Menabde et al., 1999]:

$$I_{D,T_R} = \frac{\mu(D_{ref}) - \sigma(D_{ref})\varepsilon}{\left(\frac{D}{D_{ref}}\right)^{-n}} \quad (10)$$

where $n = K(1) = K(2)/2$ and

$$\varepsilon = -\log\left(1 - \frac{1}{T_R}\right) \quad (11)$$

Comparing the IDF formulation in Equation (10) with Equation (1), the numerator corresponds to $a(T_R)$ (i.e. the value of the variable for the reference duration and for the given return period) while the denominator is equal to D^{-n} if $D_{ref} = 1h$. In simple scaling, n does not depend on T_R . The same expression has been derived by Borga et al. [2005] considering, instead of the Gumbel location and scale parameters, the expected value of the annual maxima at D_{ref} and the coefficient of variation CV:

$$I_{D,T_R} = m(x) \left\{ 1 - \frac{CV\sqrt{6}}{\pi} [\gamma + \log(\varepsilon)] \right\} \left(\frac{D}{D_{ref}}\right)^n \quad (12) \quad I_{D,T_R} = E[x] \left(1 + \frac{CV [-(-\log F)^{-\xi} - \Gamma(1-\xi)]}{\sqrt{\Gamma(1-2\xi) - \Gamma(1-\xi)^2}} \right) \left(\frac{D}{D_{ref}}\right)^n \quad (18)$$

As pointed out by the same authors, the Gumbel parameter estimation by L-moments [Hosking, 1990] is more adapted than classical moments when dealing with high-variability observations such as rainfall maxima. Different from the above-quoted method involving classic statistical moments, the L-moments estimation only requires that the mean of the distribution is finite; the higher-order moments do not need to be finite. In addition, the computation of standard error only requires that the distribution has finite variance. Furthermore, L-moments are less sensitive to outlying data values [Vogel and Fennessey, 1993].

The Gumbel parameters are given as a function of the L-moments by Stedinger et al. [1993]. With respect to Equation (12), the first two sample moments $m(x)$ and $s(x)$ are substituted by the first and second order L-moments L_1 and L_2 . The ratio of the two first order L-moments, called τ_2 , has the same significance as the CV. Therefore, in the case of maxima following a Gumbel distribution whose parameters are estimated using L-moments, the IDF relationship becomes [Borga et al., 2005]:

$$I_{D,T_R} = L_1 \left(1 - \frac{\tau_2}{\log(2)} (\gamma + \log(\varepsilon)) \right) \left(\frac{D}{D_{ref}}\right)^n \quad (13)$$

where n and ε are the same as Equation (10).

2.2. GEV Simple-scaling IDF model

The Gumbel distribution is a particular case (EV-I) of the more general GEV distribution. Recently Koutsoyianis [2003] has demonstrated the appropriateness of the GEV distribution in modeling daily rainfall maxima. The GEV

simple-scaling model can be derived from the inverse function of the cumulative probability density function:

$$F(x, \mu, \sigma, \xi) = \exp \left\{ - \left[1 + \xi \left(\frac{x - \mu}{\sigma} \right) \right]^{-1/\xi} \right\} \quad (14)$$

and from the expression of the first two GEV moments:

$$\hat{\mu} = m(x) + \frac{\hat{\sigma}}{\xi} - \frac{\hat{\sigma}}{\xi} \Gamma(1 - \xi) \quad (15)$$

and

$$\hat{\sigma}^2 = \frac{s(x)^2 \xi^2}{\Gamma(1 - 2\xi) - \Gamma(1 - \xi)^2} \quad (16)$$

where $\Gamma(z) = \int_0^\infty t^{z-1} \exp(-t) dt$ is the Gamma function, and $\xi \neq 0$. The case $\xi = 0$ corresponds to the Gumbel density function detailed in the previous paragraph. From Equation (4) (also in Salvadori and De Michele [2001]), the GEV parameters μ , σ and ξ scale as:

$$\begin{cases} \mu_D = \left(\frac{D}{D_{ref}}\right)^{-H} \cdot \mu_{D_{ref}} \\ \sigma_D^2 = \left(\frac{D}{D_{ref}}\right)^{-2H} \cdot \sigma_{D_{ref}}^2 \rightarrow \sigma_D = \left(\frac{D}{D_{ref}}\right)^{-H} \cdot \sigma_{D_{ref}} \\ \xi_D = \xi_{D_{ref}} \end{cases} \quad (17)$$

The simple-scaling GEV model is defined, as a function of $E[x]$, CV, ξ and n by:

The use of L-moments [Smithers and Schulze, 2004] improves the parameters estimation as in the case of Gumbel model. The first two GEV parameters are defined as a function of the first two L-moments (L_1 and L_2) as:

$$\hat{\mu} = L_1 + \sigma \frac{1 - \Gamma(1 - \xi)}{\xi} \quad (19)$$

$$\hat{\sigma} = \frac{-L_2 \xi}{(1 - 2\xi) \Gamma(1 - \xi)} \quad (20)$$

The combination of Equations (14) and (19) leads to the simple scaling expression of the IDF relationship for GEV distributed maxima:

$$I_{D,T_R} = L_1 \left(1 + \tau_2 \frac{\left(1 - \frac{(-\log F)^{-\xi}}{\Gamma(1-\xi)} \right)}{(1 - 2\xi)} \right) \left(\frac{D}{D_{ref}}\right)^n \quad (21)$$

Described in terms of classical moments or L-moments, the GEV scale-invariant model provides a mathematically coherent coupling between the power-law form that IDF curves exhibit for each return period and the asymptotic distribution for block-maxima, the GEV distribution. In the next sections, we apply this model to the rainfall database located in the Cévennes-Vivarais region, in Southern France, with the aim to better understand the behavior of extreme rainfall events for ranges of accumulation of hydrological interest.

3. Data

The studied region is located in the South-East of France. It covers a window of about $160 \times 200 \text{ km}^2$, bounded by the Rhone River to the East and by the Mediterranean Sea to the South (Figure 1). The South-Eastern half is rather flat and close to the sea level. The N-W half of the region is part of the Massif Central. Oriented 30° , the Massif Central ridge reaches maximum elevations between 1500 and 1800 m.

The rainfall regime of the region combines the effects of Mediterranean Sea and the mountainous topography. In this context, the location and magnitude of the rainfall extremes vary with the accumulation period [Ceresetti et al., 2010; Molinié et al., 2010]. Jacq [1994] reported 144 rain events with daily precipitation amounts greater than 190 mm during the 1958-1994 period. They mainly involve the southeasterly oriented foothills. On the other side, the flat area is submitted to intense but relatively short showers (up to 150 mm h^{-1}). All the region is prone to flash floods.

The database is provided by the French meteorological service Météo-France and managed by the OHM-CV (Mediterranean Hydro-meteorological Observatory Cévennes-Vivaraïs). The available raingages are reasonably well distributed in the area. The average rain gage density is about 2 stations per 100 km^2 (Figure 2). Only the 2% of the area with elevation above 1600 m is undersampled.

The hourly rainfall database is composed of about 150 tipping-bucket raingages series covering the period 1993-2008.

The daily database includes about 300 series covering the period 1958-2008. In order to better assess the rainfall extreme behavior, this study uses the 225 daily raingages having more than 30 years of continuous data.

4. Implementation of the IDF-scaling model

4.1. IDF simple-scaling exponent

The scale invariance of annual-maxima over a large range of durations (1h to 100 h) is first tested thanks to the longest rainfall series available in the study region: Montpellier with 52 years of hourly observations. The box-plot of Figure 3 summarizes the annual maxima distributions for different durations. In the range 4-100 h, the annual-maxima distributions scale linearly in a log-log diagram. In agreement with Equations (3) and (4), the mean and the median of the distributions exhibit log-linearity and the distribution scattering (variance, inter-quantile distance) is constant. For durations below 4 hours, the observed rainfall maxima are lower than the log-linear model prediction. This could be due to sampling issues: stochastic simulations [Molinié et al., 2010] show that maxima are underestimated by up to 40% for durations lower than 4 hours when estimated using clock hourly data instead of moving windows designed to capture maxima from the continuous signal.

To verify if log-linearity of rainfall maxima can be assumed in the study domain, hourly rainfall series have been aggregated from 1 to 100 h in 169 stations with recording raingages. Annual maxima are extracted and the scaling behavior of their first sample moment (i.e. the average) is assessed by computing its correlation (in log-log plot) with the accumulation duration (like in an analogous study by De Michele et al. [2001]). In most of the cases the correlation coefficient R^2 statistics is close to one (Figure 4). To further analyze the deviation from simple scaling in the region, let us consider the moment scaling function defined in Equations (6) and (8): Simple scaling is verified, at least for the first two moments, when the ratio $K(2)/(2K(1))$ is equal to 1. In Figure 5-a the histogram of $K(2)/(2K(1))$ shows that most rain gages display ratios close to one (between 1 and 1.15)

and that slight deviations from simple scaling are present in the region. Mapping the ratio $K(2)/(2K(1))$ (Figure 5-b), we identify that the higher deviations from simple-scaling are found close to the mountain ridge and in the Massif Central plateau (NW from the ridge). The simple-scaling model can therefore be safely adopted at least in the flat area close to the Mediterranean Sea.

By means of the hourly rain gage database we have shown that annual-maxima scale on the range 4h to 4 days and we have, in general, small deviations from simple-scaling of maxima. From a practical point of view, the estimation of the scaling exponent can be performed with the daily database, featuring longer series compared to the hourly database. The first two L-moments of annual-rainfall maxima series in the range 1-4 days are computed and, for each rain gage, a straight line is fitted in the double logarithmic plot as reported in Figure 6. The slope of the straight line provides an estimation of the scaling exponent n , first parameter of the GEV simple-scaling model, according with Equation (21).

4.2. Hybrid extreme modeling at $D = D_{ref}$

The identification of the maxima distribution at a reference scale is the first application step of the GEV simple scaling model of Equation (21). The distribution parameters are estimated from the daily database (see Section 3) characterized by long series (30 to 50 years) and by a dense rain gage network (225 gages).

As stated in Koutsoyiannis [2003], the inference of GEV parameters requires very long series of maxima. This is due to the GEV flexibility to accommodate heavy, exponential as well as bounded tails. To overcome this issue, Koutsoyiannis [2003] proposes to gather rainfall maxima between adjacent gages, neglecting their mutual dependency.

An alternative solution to the classical selection of annual maxima is the “Peaks-Over-Threshold” (POT) approach [Reiss and Thomas, 1997]. It consists in selecting a given number of independent observations above a threshold in the whole rainfall series. The over-threshold observations are then modeled with a Generalized Pareto Distribution (GPD). Numerical studies made by the authors show that, in synthetic series, about one hundred observations are needed to obtain stability of the estimation. The drawback of the GPD approach is that, defining an observation threshold rather than a fixed-window period, the frequency is not directly expressed in years.

Coles [2001] presents an elegant solution for the modeling of extremes using the “Point-Process” analysis. This theory defines a common background to both block-maxima and “Peaks-Over-Threshold” analyses. It has been successfully used in fields other than rainfall, such as the modeling of snowfall accumulation extremes [Blanchet et al., 2009]. The method assumes that the number of observations exceeding a threshold within a block is Poisson-distributed. It allows to consider more than one maximum per year (similarly to the POT method) with the advantage of estimating the same parameters of GEV distribution.

In the context of scale-invariance, the GEV parameters estimation should be made by means of the statistical moments (classic moments or L-moments). The need of expressing the extreme behavior in terms of the statistical moments justifies the use of an hybrid method combining the ease of GEV formulation with the higher number of observations provided by the Peaks-Over-Threshold method. This practical method will be described in the next section.

4.3. Estimation of GEV parameters at $D = D_{ref}$

According with Equation (21) the estimation of the L-moments L_1 and L_2 and of the shape parameter ξ is necessary to describe the rainfall regime at $D = D_{ref}$.

The main problem in the estimation of the GEV model is the ξ parameter estimation. Taking the annual maxima, the sample size does not allow a correct estimation of the third statistical moment, which in turn is at the base of the estimation of ξ . The ξ parameter has the larger importance in the definition of the extremes behavior at large return period. The extreme value theory [Reiss and Thomas, 1997; Salvadori and De Michele, 2001; Coles, 2001] highlights that the shape parameter ξ has the same meaning in the block-maxima and POT approaches. We can therefore inject the shape parameter determined by POT as a fixed value into the GEV estimation.

The process begins with the selection of the Peaks-Over-Threshold. To ensure independence between two extremes, we fix the time lag for which the average auto-correlation function of non-null rainfall is lower than a given value. In our case the lag of 4 days seems adequate (Figure 7). The declustering procedure consists in taking all the independent observations exceeding a threshold. According with empirical test not shown here, a correct estimation of the ξ parameter can be done with no less than 90 observations. For this reason, a threshold of 3 average events per year is chosen at all stations, so that the shorter series had at least 90 observations and the longest over 150 observations.

We point out that our arbitrary NPY determination, constrained by the series length, yields results similar to those obtainable at each gage with purely statistical threshold-determination methods, such as the mean residual life plot [Davison and Smith, 1990; Coles, 2001] or the GPD estimation for a range of threshold [Coles, 2001].

Figure 8 shows a comparison between the shape parameter ξ derived from the GEV and the GPD estimation. The GPD estimation gives lower dispersion of ξ . The ξ parameter is, in average, slightly positive. A limited number of gages have slightly negative ξ values, indicating bounded extreme distributions. Many gages present ξ around zero, which means Gumbel-distributed maxima. A significant number of gages, mainly located in the flat area, have significantly positive ξ values, indicating heavy-tailed extreme distributions.

Estimations of large return periods are very sensitive to the value of ξ . An additional validation has been performed, checking the ξ sensitivity as a function of NPY. We do not note any significant difference in the ξ estimation for NPY equal to 2 or 4. Differences occur for NPY lower than 2 or equal or higher than 5. The first effect is due to limited sample size, while the second, leading to shape-parameter estimations biased towards zero, is due to the inclusion of a consistent number of common observations into the extreme sample set.

4.4. Mapping the model parameters

In order to infer the value of the model parameters over a regular grid, an interpolation process is necessary. The application of geostatistical interpolation require second-order stationarity and gaussian distribution of each parameter within the region. Kriging interpolation allows to infer the values at ungedged stations by linear interpolation of the values at each station; the weight of each station is defined by the covariance function and is chosen such to minimize the error committed.

Figure 9-a and 9-b report the first two L-moments of the annual-maxima distribution at $D = D_{ref}$ (the model parameters of Equation (21)). L_1 and L_2 represent the annual average maximum-rainfall and its rainfall variability (related to the standard deviation), respectively. The two first distribution moments are approximately linearly related, and thus they display similar patterns, with a common maximum along the mountain ridge around the station of "Serre the la Croix de Bauzon".

In Figure 9-c, the shape parameter map shows that along the mountain ridge, where the highest maximum rainfall

depths are recorded, no heavy-tails are found. This means that extreme events in this zone have similar magnitudes. In the highland zone, in the N-W half of the study domain, the shape parameter is weakly positive. In the flat area close to Alès and Nîmes, on the contrary, the shape parameter is significantly positive indicating that extreme events have different magnitude (i.e. heavy-tails). In the zone between Mount Aigoual and the city of Millau, the positive value of the shape parameter is probably an interpolation artifact due to the relative sparseness of the network affecting the parameter interpolation (see Figure 1).

We show in the next section that our GEV simple-scaling model can be used to derive, from the daily-maxima, infra-daily maxima with very different spatial patterns. The scaling exponent plays the important role of relating the maxima distribution moments at different temporal scales.

Figure 9-d shows the scaling exponent n for the GEV simple-scaling model. Differently from the other parameters, this parameter is independently derived from the observations at different scales and therefore it can be directly interpreted. This parameter describes the scaling of the statistical moments from a time scale to another one, and it is a good indicator of the influence of orography at different accumulation periods. Along the crest line of the ridge, the scaling exponent is minimal, meaning that the absolute difference between extremes at different temporal scales is lower.

5. Application example and meteorological interpretation

A first application of the GEV simple scaling model is the comparison of the Intensity- Duration- Frequency curves at two locations characterized by considerably different behavior of extremes: the city of Nîmes, close to the Mediterranean Sea and located on the flat area, and the Mount Lozère, where the orographic signature of rainfall is significant. Figure 10 shows the IDF curves for the return period of 10, 20, 50 and 100 years computed at the two locations and superposed. It is easy to see that the precipitation intensity for $D < 2h$ is lower for the mountainous location, probably because of the low probability of deep convective events. On the other hand, for large accumulation periods, the stability of meso-scale convective system leads to higher rain depths (and consequently higher rainfall intensities) over the mountain. The main parameter determining such a difference between locations is the scaling exponent n , considerably lower in mountainous region (i.e. the extreme rainfall intensity does not change too much with duration).

A second application of the GEV simple-scaling model is reported in Figure 11. This figure shows the rainfall intensity corresponding to a return period of 100 years computed for durations of 1, 4, 8 and 24 hours from the IDF model described in the previous sections. As already noted before, the large uncertainties due to the shortness of infra-daily rainfall series prevent the direct quantitative estimation of hourly-rainfall extreme behavior.

However, the spatial patterns of the rainfall intensities for $D < 4h$ are in line with the few statistical indicators directly derivable from the hourly database, such as the hourly rainfall maxima [Bois et al., 1997; Ceresetti et al., 2010; Molinié et al., 2010].

The main achievement obtained from the implementation of the GEV simple-scaling model is a continuous description of the rainfall intensity as a function of the duration at each point of the region. This is not an obvious result if the intensity for a given return period is computed for different durations in an independent manner. In addition, the

incoherence between durations can be amplified by the rainfall sampling measurement error (Habib et al. [2001]; Ciach [2003] among others).

Figure 11 illustrates that, varying the accumulation period, the localization of maximum rainfall intensities moves in consequence of the combination of various storm typologies. Several case studies of heavy rainfall events producing flash floods in the Cévennes region [Sénési et al., 1996; Ducrocq et al., 2002, 2003, 2008; Nuissier et al., 2008] have shown that extreme rainfall events are due to meso-scale convective systems blocked over the foothill alternatively by the terrain elevation and by the formation of cold pools. The contribution of the associated deep convection to the extreme rainfall regime is visible at short durations (Figure 11-a). Over 100 mm h^{-1} patterns are clearly visible to the North of Montpellier and Alès. As the duration increases, the patterns related to deep convection progressively fades out, replaced by a more elongated pattern along the mountain ridge (Figure 11-b, 11-c, 11-d). These daily rainfall extremes are generated by shallow convective clouds triggered by relief shoulders and organized in orographic rainbands [Miniscloux et al., 2001; Anquetin et al., 2003]. These bands are maintained stationary by specific meteorological conditions [Godart et al., 2009].

More can be said regarding the possibility of GEV model, compared to the Gumbel one, to leave the shape parameter free to vary. The comparison between GEV and Gumbel in terms of extreme rainfall are shown in Figure 12 for the return periods of 50 and 100 yrs for durations of 2 and 24 h. The maximum difference in terms of rainfall amount for $T_R = 100 \text{ yrs}$ and $D = 2 \text{ h}$ is 50 mm , which is about the 15% of the total rainfall amount predicted by these models. For the return period $T_R = 50 \text{ yrs}$ the difference in rainfall depth is 35 mm corresponding to less than 10%. This example shows that the influence of the shape parameter increases with the return period.

Consequently, in every situations where the rainfall extremes in the region of interest features exponential as well as hyperbolic tails, the GEV framework should be recommended for the regionalization of Intensity-Duration-Frequency curves in a scale-invariance context.

6. Conclusion

The paper presents a scale-invariant model for dealing with the estimation of rainfall maxima in a region of Southern France. The model features Intensity-Duration-Frequency curves obtained coupling scale-invariance and extreme-value distribution of rainfall maxima (GEV distribution). The aim of the work is twofold: i) to improve the assessment of the frequency of extreme rainfall events (return period equal or higher to 100 years), incorporating the heavy-tail behavior of extremes (GEV); ii) to derive infra daily estimation from long daily rainfall series (that are commoner) by applying scaling.

The simple-scaling hypothesis applicability has been tested in the region, highlighting that the largest deviations occur in the mountainous regions to the right of Rhône River and on the Massif Central highlands. Even if simple scaling is convenient for dealing with rainfall extremes, it can lead to biases in those areas. Unfortunately, multiple scaling can not be correctly implemented using GEV distributions, due to the incompatibility of the statistical-moments scaling.

The choice between Gumbel or GEV models is relevant for the estimation of large return period rainfall, but is subordinated to the availability of large quantity of observations. Gumbel distribution has been widely used [Koutsoyiannis, 2003], but in the region of interest its exponential tail is not adequate in the flat sub-region, submitted to very intense developed convection.

The GEV simple-scaling model is defined as a function of 4 parameters: the scaling exponent n , the GEV parameters

μ , σ and ξ . The formulation using L-moments proposed in Equation (21) is easier to infer with high variability of series and outliers. A coupled method to estimate the shape parameter ξ is presented with the purpose to decrease the estimation variance without complicating the formulation.

From daily accumulation data, the model is able to predict the behavior of extreme rainfall for short durations in coherence with previous more empirical studies analyzing hourly intensities Bois et al. [1997].

This model is a new tool able to estimate the frequency of an event at every temporal scale within the range 4h - 4 days. It can help to solve problems of design engineering. Extending this model to extreme spatial rainfall in the region would allow to classify the severity of meteorological events at all the temporal and spatial scales of hydrological interest.

Acknowledgments. This work has been performed in the framework of ANR-MEDUP French project on the "Forecast and projection in climate scenario of Mediterranean intense events : Uncertainties and Propagation on environment". We are grateful to Météo-France and to OHM-CV observatory for providing the meteorological data.

References

- Anquetin, S., F. Miniscloux, and J. D. Creutin, Numerical simulation of orographic rainbands, *Journal of Geophysical Research*, 108(D8), doi:10.1029/2002JD001593, 2003.
- Bendjoudi, H., P. Hubert, and D. Schertzer, Interpretation multifractale des courbes intensité-durée-fréquence des précipitations, *C. R. Acad. Sci. Paris*, 2, 323-326, 1997.
- Bernard, M., Formulas for rainfall intensities of long durations, *Transactions of the American Society of Civil Engineers*, 96, 592-624, 1932.
- Blanchet, J., C. Marty, and M. Lehning, Extreme value statistics of snowfall in the swiss alpine region, *Water Resources Research*, 45(5), 2009.
- Bois, P., H. Mailloux, C. Obled, and F. De Saintignon, Atlas expérimental des risques de pluies intenses dans la région Cévennes-Vivarais, *Tech. rep.*, LAMA, Pôle Grenoblois des Risques Naturels, 1997.
- Borga, M., C. Vezzani, and G. Dalla Fontana, Regional rainfall depth-duration-frequency equations for an alpine region, *Natural Hazards*, 36, 221-235, 2005.
- Burlando, P., and R. Rosso, Scaling and multiscaling models of depth-duration-frequency curves for storm precipitation, *Journal of Hydrology*, 187, 45-64, 1996.
- Ceresetti, D., G. Molinié, and J. Creutin, Scaling properties of heavy rainfall at short duration: a regional analysis, *Water Resources Research*, in press, 2010.
- Ciach, G., Local random errors in tipping-bucket rain gauge measurements, *Journal of Atmospheric and Oceanic Technology*, 20, 752-759, 2003.
- Coles, S., *An Introduction to Statistical Modeling of Extreme Values*, Springer-Verlag London Limited 2001, 2001.
- Davison, A. C., and R. L. Smith, Models for exceedances over high thresholds (with discussion), *Journal of the Royal Statistical Society*, 52, 1990.
- De Michele, C., N. Kottegoda, and R. Rosso, The derivation of areal reduction factor of storm rainfall from its scaling properties, *Water Resources Research*, 37(12), 3247-3252, 2001.
- Deidda, R., R. Benzi, and F. Siccardi, Multifractal modeling of anomalous scaling laws in rainfall, *Water Resources Research*, 35, 1853-1867, 1999.
- Ducrocq, V., D. Ricard, J. Lafore, and F. Orain, Storm-scale numerical rainfall prediction for five precipitating events over france: On the importance of the initial humidity field, *Weather and Forecasting*, 17, 1236-1256, 2002.
- Ducrocq, V., G. Aullo, and P. Santurette, Précipitations intenses et les inondations des 12 et 13 novembre 1999 sur le sud de la France, *La Météorologie*, (42), 18-27, 2003.

- Ducrocq, V., O. Nuissier, D. Ricard, C. Lebeaupin, and T. Thouvenin, A numerical study of three catastrophic precipitating events over southern France. ii: Mesoscale triggering and stationarity factors, *Quarterly Journal of the Royal Meteorological Society*, (134), 131–145, 2008.
- Godart, A., S. Anquetin, and E. Leblois, Rainfall regimes associated with banded convection in the cévennes-vivarais area, *Meteorology and Atmospheric Physics*, (103), 25–34, 2009.
- Gupta, V., and E. Waymire, Multiscaling properties of spatial rainfall and river flow distributions, *Journal of Geophysical Research*, 95(D3), 1999–2009, 1990.
- Habib, E., W. Krajewski, and A. Kruger, Sampling errors of tipping-bucket rain gauge measurements, *Journal of Hydrologic Engineering*, pp. 159–166, 2001.
- Harris, D., E. Foufoula-Georgiou, K. Droegemeier, and J. Levit, Multiscale properties of a high-resolution precipitation forecast, *Journal of Hydrometeorology*, 2, 406–418, 2001.
- Hosking, J., L-moments: analysis and estimation of distributions using linear combination of order statistics, *Journal of the Royal Statistical Society*, 52, 105–124, 1990.
- Hubert, P., and J.-P. Carbonnel, Dimensions fractales de l'occurrence de pluie en climat soudano-sahlien, *Hydrologie continentale*, 4, 3–10, 1989.
- Jacq, V., Inventaire des situations à précipitations diluviennes sur les régions Languedoc-Roussillon, PACA et Corse, période 1958-1994 (inventory of extreme precipitation events in the Languedoc-Roussillon, PACA and Corse regions in the period 1958-2004). phénomènes remarquables., *Tech. Rep. 3*, Météo-France (SCEM), 1994.
- Kotz, S., and S. Nadarajah, *Extreme value distributions. Theory and applications*, Imperial College Press, 2000.
- Koutsoyiannis, On the appropriateness of the gumbel distribution for modelling extreme rainfall, in *ESF LESC Exploratory Workshop*, Bologna, Italy, 2003.
- Koutsoyiannis, D., D. Kozonis, and A. Manetas, A mathematical framework for studying rainfall intensity - duration - frequency relationships, *Journal of Hydrology*, 206, 118–135, 1998.
- Ladoy, P., F. Schmitt, D. Schertzer, and S. Lovejoy, Variabilité temporelle multifractale des observations pluviométriques à Nîmes, *C.R. Acad. Sci. Paris*, 317(2), 775–782, 1993.
- Marsan, D., D. Schertzer, and S. Lovejoy, Causal space-time multifractal processes: Predictability and forecasting of rain fields, *Journal of Geophysical Research*, 101, 26,333–26,346, 1996.
- Menabde, M., A. Seed, and G. Pegram, A simple scaling model for extreme rainfall, *Water Resources Research*, 35(1), 335–339, 1999.
- Miniscloux, F., J. D. Creutin, and S. Anquetin, Geostatistical analysis of orographic rainbands, *Journal of Applied Meteorology*, 40, 1835–1854, 2001.
- Molinié, G., D. Ceresetti, E. Yates, S. Anquetin, B. Boudevillain, and J. D. Creutin, Rainfall regimes in a Mediterranean mountainous region, *submitted to Quarterly Journal of the Royal Meteorological Society*, 2010.
- Nuissier, O., V. Ducrocq, D. Ricard, C. Lebeaupin, and S. Anquetin, A numerical study of three catastrophic precipitating events over southern france. i: Numerical framework and synoptic ingredients, *Quarterly Journal of the Royal Meteorological Society*, 134, 111–130, 2008.
- Reiss, R., and M. Thomas, *Statistical Analysis of Extreme Values*, Birkhäuser Verlag, Basel-Boston-Berlin, 1997.
- Salvadori, G., and C. De Michele, From generalized pareto to extreme values law: Scaling properties and derived features, *Journal of Geophysical Research*, 106(D20), 24,063–24,070, 2001.
- Schertzer, D., and S. Lovejoy, Physically based rain and cloud modeling by anisotropic, multiplicative turbulent cascades, *Journal of Geophysical Research*, 92, 9692–9714, 1987.
- Sénési, S., P. Bougeault, J. Chèze, P. Cosentino, and R. Thepenier, The Vaison-la-Romaine flash flood: Mesoscale analysis and predictability issues, *Weather and Forecasting*, 11(4), 417–442, 1996.
- Smithers, J., and R. Schulze, The estimation of design rainfalls for south africa using a regional scale invariant approach, in *Hydrology: Science & Practice for the 21st Century*, edited by U. British Hydrological Society, London, pp. 221–228, 2004.
- Stedinger, J., R. Vogel, and E. Foufoula-Georgiou, *Frequency analysis of extreme events*, pp. 18.1–18.66, McGraw-Hill, New York, 1993.
- Veneziano, D., and P. Furcolo, Multifractality of rainfall and scaling of intensity-duration-frequency curves, *Water Resources Research*, 38(12), 1306, 2002.
- Venugopal, V., E. Foufoula-Georgiou, and V. Sapozhnikov, Evidence of dynamic scaling in space-time rainfall, *Journal of Geophysical research*, 104(D24), 31,599–31,610, 1999.
- Vogel, R., and N. Fennessey, L-moment diagrams should replace product moment diagrams, *Water Resources Research*, 29(6), 1745–1752, 1993.

G. Molinié, LTHE, Laboratoire d'Etudes des Transferts en Hydrologie et Environnement - (CNRS,UJF,IRD,INPG), Grenoble, France (gilles.molinie@ujf-grenoble.fr)

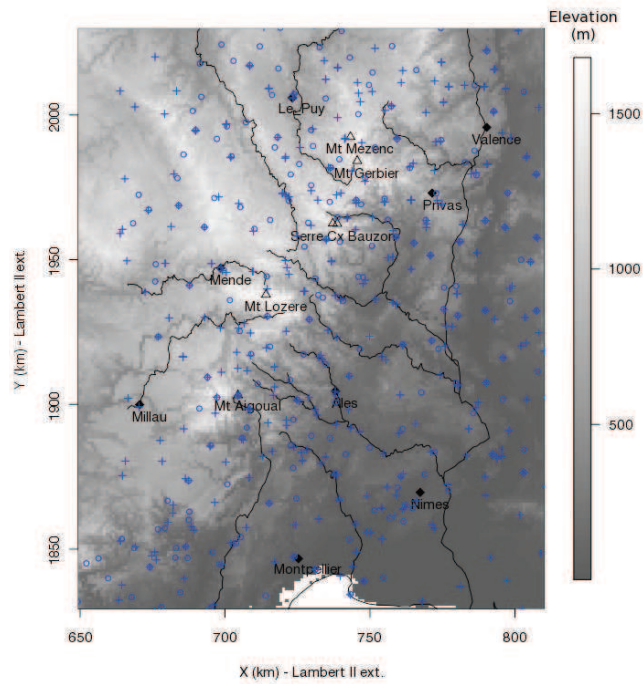


Figure 1. Map of the region of analysis: elevation above sea level (shaded surface), hydrographic network (solid line), the five highest mountain peaks (triangles), main cities (diamonds), the daily (circles) and the hourly (crosses) rain gauge network.

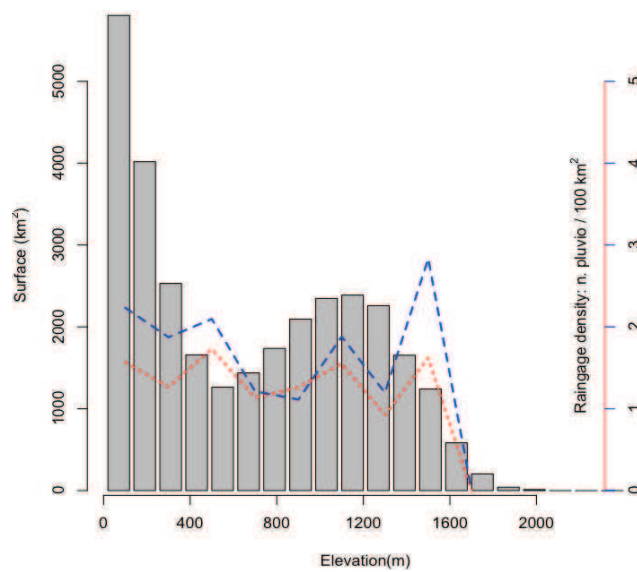


Figure 2. Histogram of topography elevation of the study area (grey boxes, reading on the vertical axis) and rain gauge density as function of the elevation. The rain gauge density is expressed as the ratio of the rain gauge number in a given altitude range to the ground surface of this altitude. The long dashed lines represents the daily rain gauge network and the short-dashed line for the recording rain gauge network (reading on the right axis).

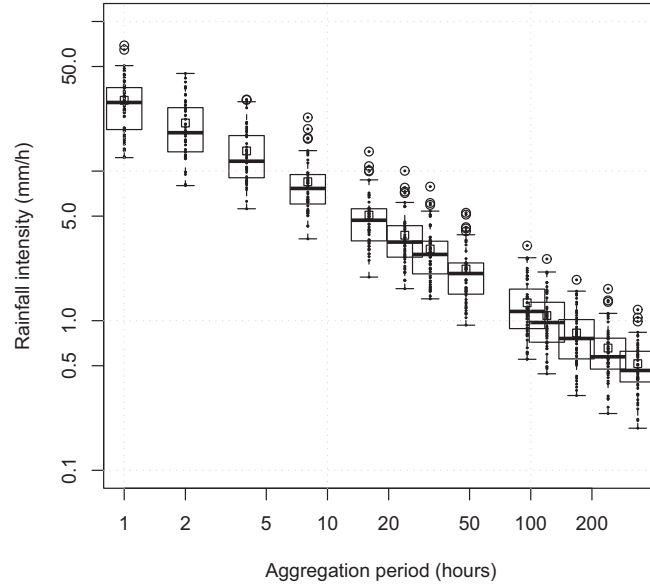


Figure 3. Analysis of the rain gage station of Montpellier (1920-1972): the annual rainfall maxima for different accumulation durations is expressed in $mm h^{-1}$. The small dots represent the 53 individual values, the white square is their mean value. The boxplot defines the lower hinge (q_{25}), the median (q_{50}), the upper hinge (q_{75}). Possible outliers are marked as white circles. A straight line gathers the mean values; its slope represents the scaling exponent of the mean.

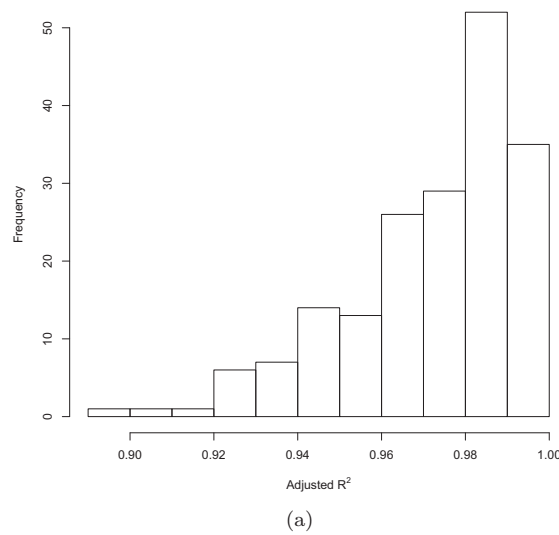


Figure 4. Analysis of 169 stations with recording rain-gages spread over the study area. Histogram of the squared correlation coefficient R^2 relating the mean annual maxima with the accumulation duration, evaluated in the range 1-100 hours, on 169 recording rain gage stations having more than 4 continuous years of hourly observations.

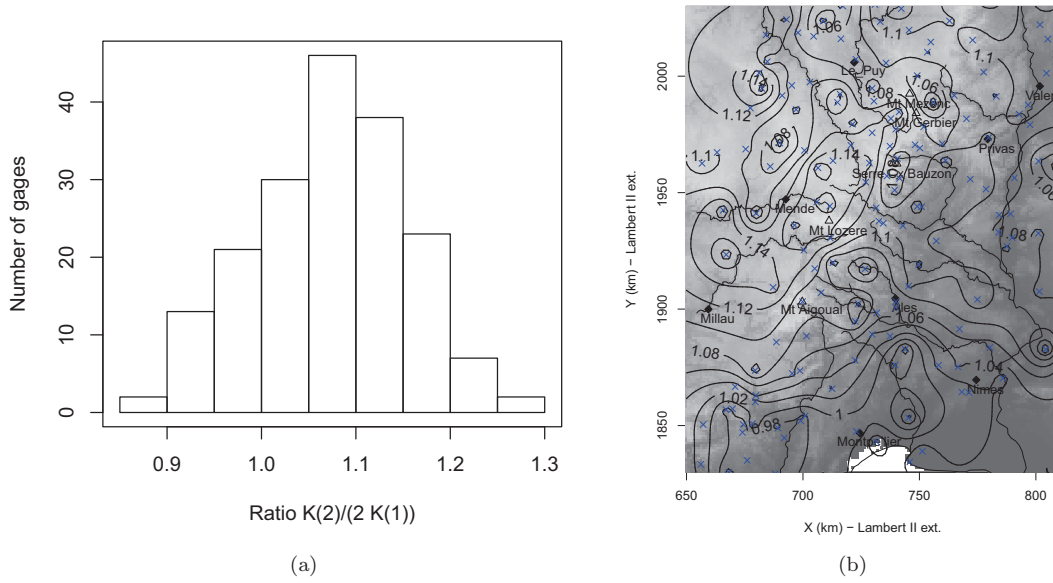


Figure 5. Deviation from simple scaling. a: Histogram of ratios between the scaling exponent of the standard deviation $K(2)/2$ and the scaling exponent of the mean $K(1)$. b: Map of the same index over the study region. Ratio values close to the unit indicate simple scaling of maxima.

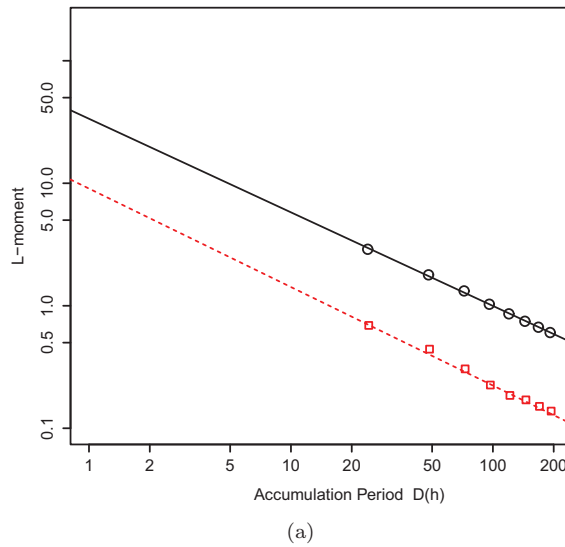


Figure 6. Computed L-moments for the station St. Pierreville. Circles: empirical 1st L-moment L_1 ; Solid line: log-log fit of L_1 versus accumulation time; Squares: empirical 2nd L-moment L_2 ; Dashed line: log-log fit of L_2 versus accumulation period. The slope of L_1 and L_2 versus time provide the values of $K(1)$ and $K(2)$ of Equation (8), respectively. Notice that $K(1) = K(2)/2 = n$ in simple scaling.

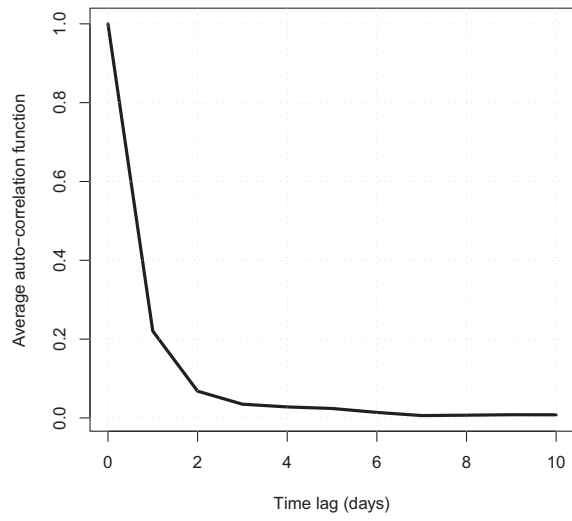


Figure 7. Average temporal auto-correlation function for daily rainfall series for the 225 stations of the daily rainfall database.

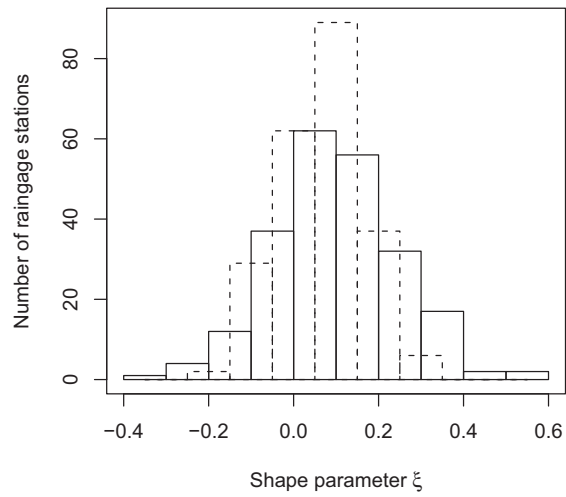
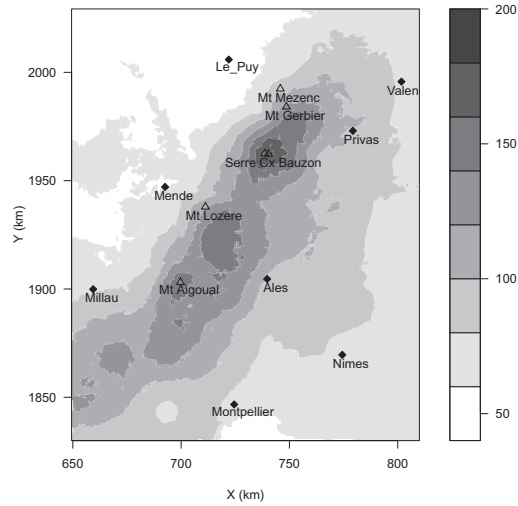
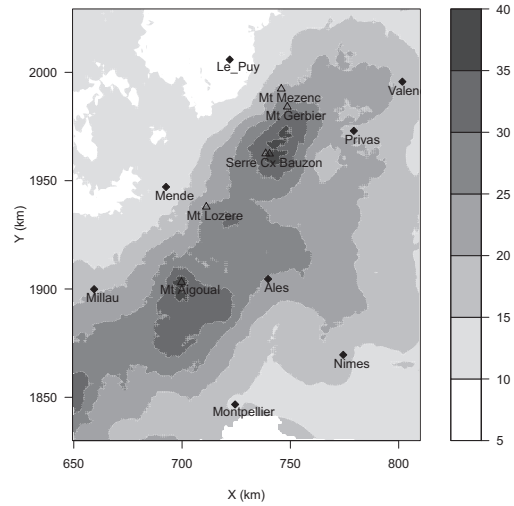


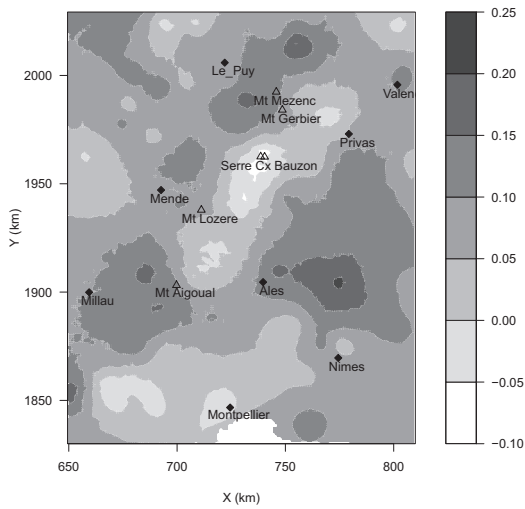
Figure 8. Comparison between the GEV (solid bars) and the GPD (dashed bars) estimation of the shape parameter ξ .



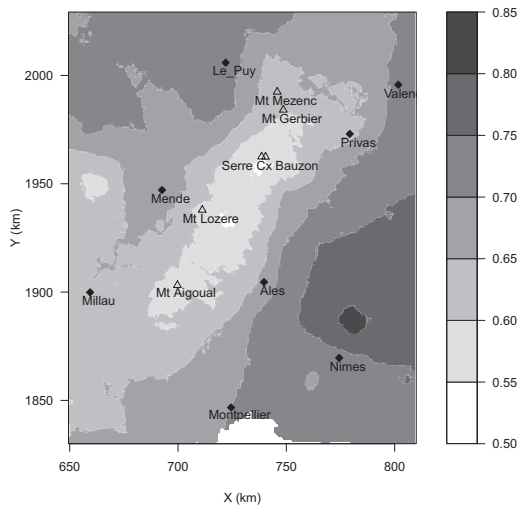
(a)



(b)

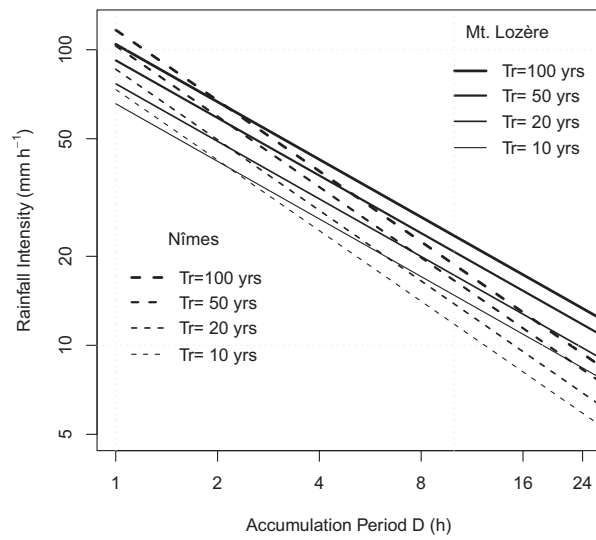


(c)



(d)

Figure 9. Maps representing the GEV simple scaling model parameters over the study regions. a: Average annual maximum rainfall (first L-moment L_1). b: Second L-Moment L_2 . c: Shape parameter ξ . d: Scaling exponent n .



(a)

Figure 10. Intensity- Duration- Frequency curves derived with the GEV scale-invariant model for Nîmes and for Mount Lozère.

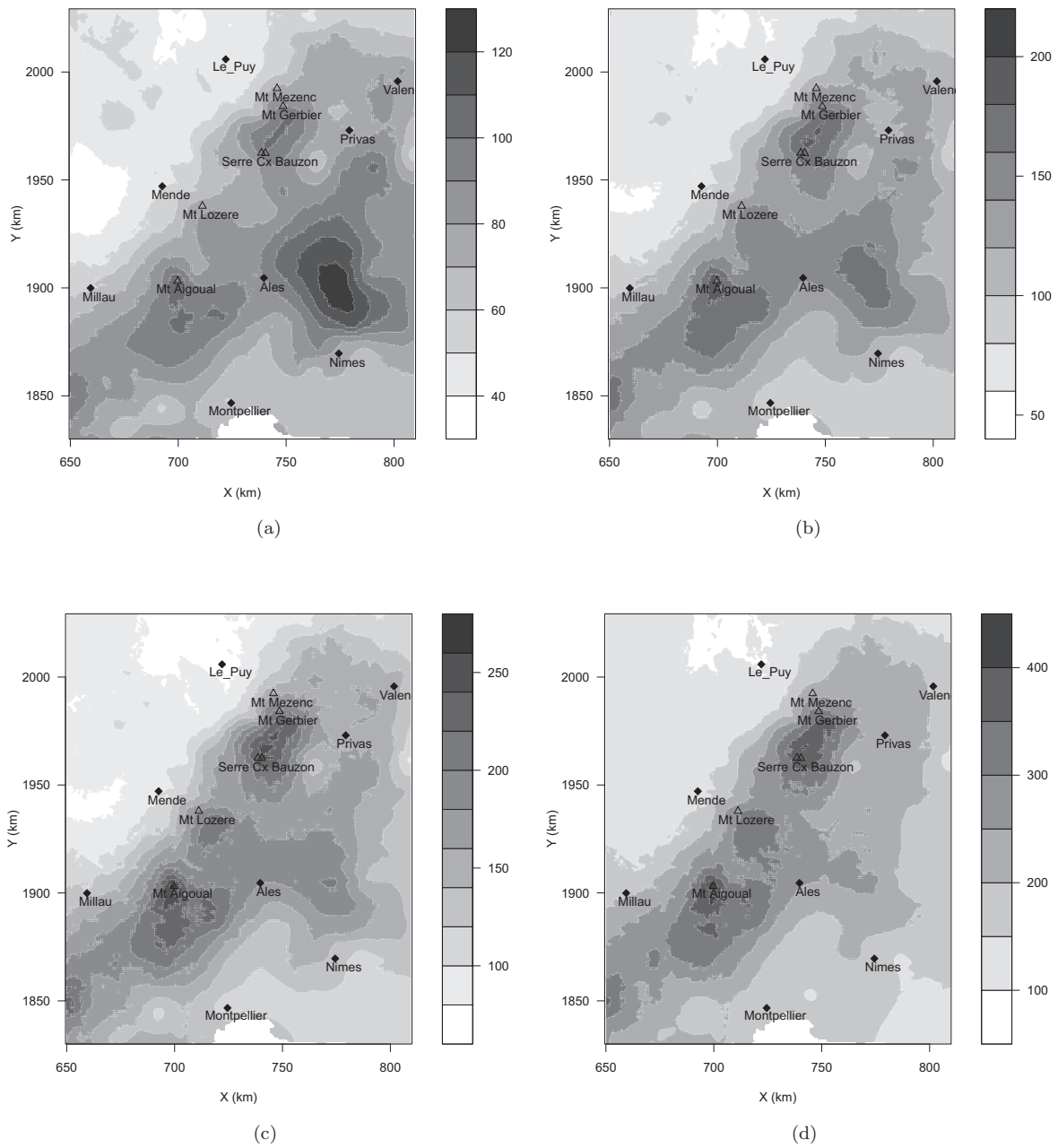


Figure 11. Map of rainfall depth (mm) for a return period of 100 yrs for different durations (a: 1h, b: 4h, c: 8h, d: 24h) according with the GEV simple-scaling model of Equation (21).

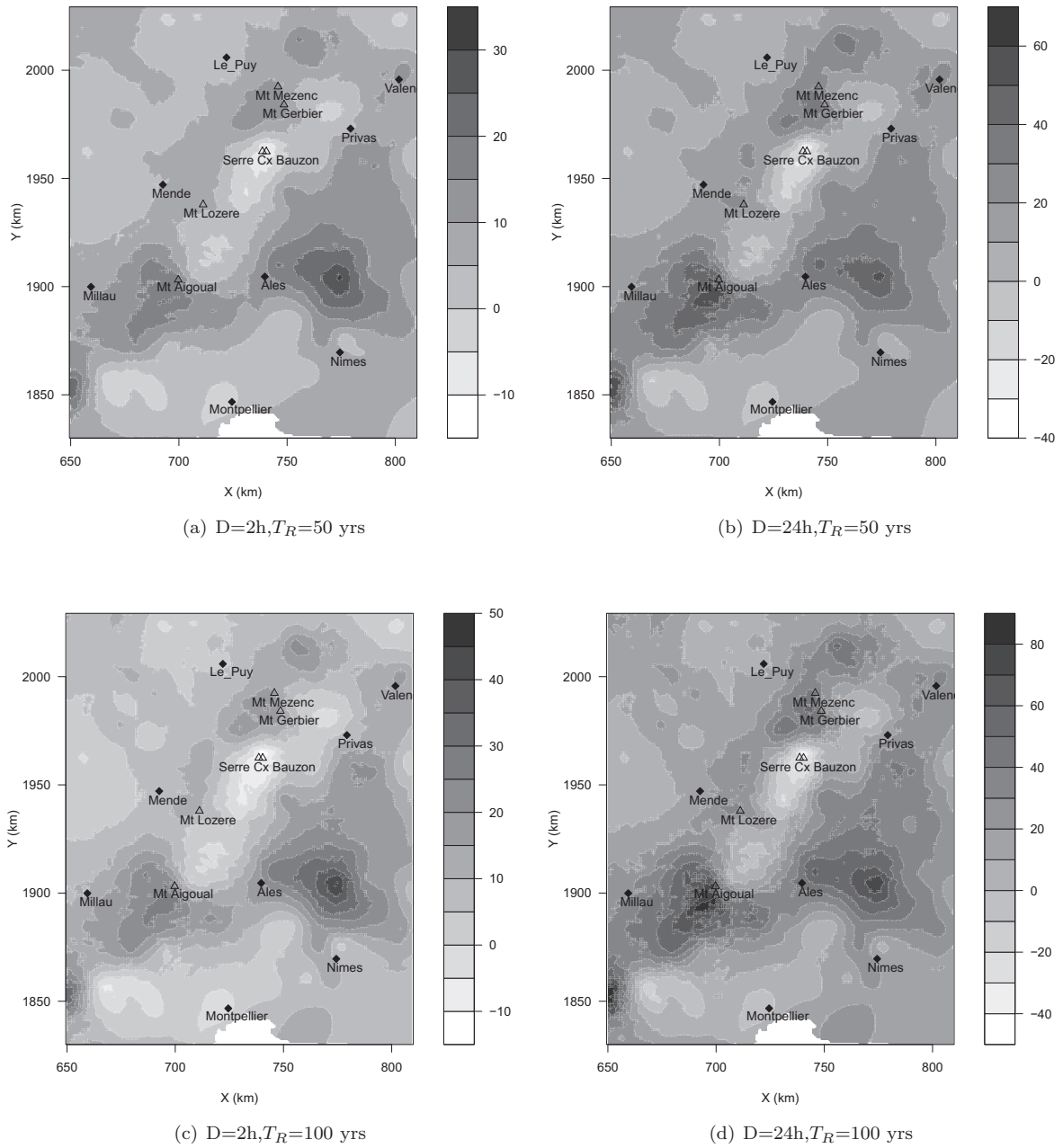


Figure 12. Differences between GEV and Gumbel simple scaling model in terms of rainfall intensities (in mm) for different couples accumulation-return period. a: Return periods 50 yrs and accumulation 2 h; b: Return periods 50 yrs and accumulation 24 h; c: Return periods 100 yrs and accumulation 2 h; d: Return periods 100 yrs and accumulation 24 h.

9 Space-time scaling of rainfall events: the September 2002 storm

Résumé

*Ce chapitre présente les résultats d'une analyse effectuée dans le but de caractériser la **structure spatio-temporelle** d'un événement de précipitation dans la région Cévennes-Vivarais. Nous nous concentrons sur l'événement qui a concerné le département du Gard en Septembre 2002. Pour cet événement, les images radar des précipitations instantanées ont été corrigées des erreurs systématiques et accidentelles, et ensuite calibrées grâce au réseau pluviométrique au sol et aux mesures de terrain issues du retour d'expérience effectué en 2003 (Delrieu et al., 2005) afin d'estimer quantitativement les précipitations.*

*Nous analysons les données radar afin de détecter l'invariance d'échelle spatio-temporelle. La Fonction de Structure du second ordre, liée la structure de corrélation du champ, est utilisé pour indiquer la gamme d'échelles dont l'invariance d'échelle des champs de pluie mesurés par radar est vérifiée. Nous montrons que cette invariance d'échelle est limité à des surfaces inférieures à 100-400 km² en fonction de la période d'accumulation considérée. Ensuite, nous analysons l'événement suivant le chemin proposé par Venugopal et al. (1999) qui conduit à la détermination d'un paramètre d'échelle **dynamique** définissant la relation entre la résolution spatiale et la corrélation temporelle. En perspective, l'adoption systématique de cette analyse sur un certain nombre d'événements pluvieux peut permettre de parvenir à un inventaire des caractéristiques d'invariance d'échelle des différents phénomènes météorologiques extrêmes.*

9.1 Introduction

In this chapter we report the results of an analysis conducted with the aim to characterize the **space-time structure** of rainfall events in the Cévennes-Vivarais region. We focus on the event occurred on September 2002; the radar images of instantaneous rainfall for this event have been corrected from systematic and accidental errors and subsequently calibrated (Delrieu et al., 2009) with respect to the ground rainfall measurements in order to quantitatively estimate the precipitation.

We analyze the radar-estimated fields by means of scale-invariant methods. The second-order Generalized Structure Function or variogram, related to the correlation structure of the field, is used to show scale-invariance of the radar fields. We show that the spatial scale-invariance is limited to surfaces up to 100-400 km^2 depending on the considered accumulation period. Afterwards, we analyze the event following the path proposed by *Venugopal et al.* (1999), leading to the determination of a **dynamic scaling** parameter defining the relation between spatial scale and temporal correlation. In perspective, the systematic development of this analysis on a number of storms can lead to build an inventory of the scale-invariance features of different types of storm.

9.2 Outline of the study

The meteorological radar detects the reflectivity for a given spatial volume in a particular instant. Once removed the numerous systematic and accidental errors affecting the measure (presence of buildings, trees and mountains; accidental sources such as flying bodies, natural and anthropic; rainfall-related effects such as the bright band phenomena), the reflectivity gives information about the water content of the scanned volume. Often, a transformation is required to pass from polar multi-layer scans to a regularly gridded field in 2D. Merging rain gauge data and radar information provides an efficient evaluation of the rainfall amount fallen during an event.

Thanks to the reliability of this database, it is possible to study particular properties of the rainfall process. This analysis can provide some worthy insights into the comprehension of the space-time structure of rainfall.

The first analysis we carry out on the data is the study of the 2nd order Generalized Structure Function (Section 5.5.1). If the GSF for a range of statistical moments is linear as a function of the spatial lag in double logarithmic plot, the variable is scale-invariant. In this context we analyze the analogies between the scale-invariance and the geostatistical analysis. We show that the scale-invariance of GSF implies power-law variogram (Section 4.4); this property may lead to improvements of the geostatistical modeling of rainfall fields at small spatial scales.

The second analysis we perform is the verification of space-time scaling of rainfall, according to the approach followed by *Venugopal et al.* (1999). The underlying assumption is the validity of the Taylor's "frozen field" hypothesis (Section 5.10.1). In spatial-temporal ranges where the advection effect is negligible, a space-time relation establishes, leading to a "dynamic scaling" space-time ratio. Consequently, the temporal dimension can be rescaled and considered as a third spatial dimension. *Venugopal et al.* (1999) focused on the spatio-temporal organization of rainfall at spatial scales of 2 to 20 km in space and 10 min to several hours in time, for a storm in Darwin, Australia. In Section 9.5 we will examine the underlying hypotheses and the meaning of the computation performed in *Venugopal et al.* (1999). A series of analogies with concurrent theories can be drawn, and these can have consequences in understanding the theoretical and empirical findings of multi-scale analysis and geostatistical analysis.

9.3 Data

The rainfall event analyzed occurred in 2002, September the 8-9th, in Southern France. It lasted approximately 36 hours, with a cumulative rainfall amount higher than 600 mm over a 150 km² area, the largest among the events ever recorded in the region (*Delrieu et al.*, 2005). The impressive rainfall amount can be perceived considering that the mean annual rainfall in the region ranges from 700 to 2500 mm. The event has been detected both by rain gauge and radar so that a merging has been possible providing reliable estimations of the rainfall depth and of the spatial structure of the event.

The radar scan, composed by 8 layers characterized by a different incidence angle and by a sampling interval of 5 min, has been treated following the classical approaches: the radar reflectivity has been cleaned from the bright band effects and from the soil echo effects, and after it has been converted into rainfall following empirical Z-R relations. Finally, the 3D polar scan has been converted to a 2D rainfall field, obtaining a 1x1 km² grid. The scanned window is a square of size 100 km.

For the analysis we selected the 8-hour time window characterized by the highest intensity. As suggested by *Venugopal et al.* (1999), the assumption of stationarity is necessary to perform any of these analysis. We should choose the time window such as during the period the statistical moments do not vary significantly. In absence of pure stationarity on data in the maximum rainfall period, we have chosen a period including the highest rainfall intensity and where both average and the dispersion of rainfall values were limited. Figure 9.1-a shows the entire storm behavior in terms of average nonzero rainfall, while Figure 9.1-b shows the standard deviation of the nonzero rainfall field as a function of time. A horizontal line indicates the time window considered for the analysis.

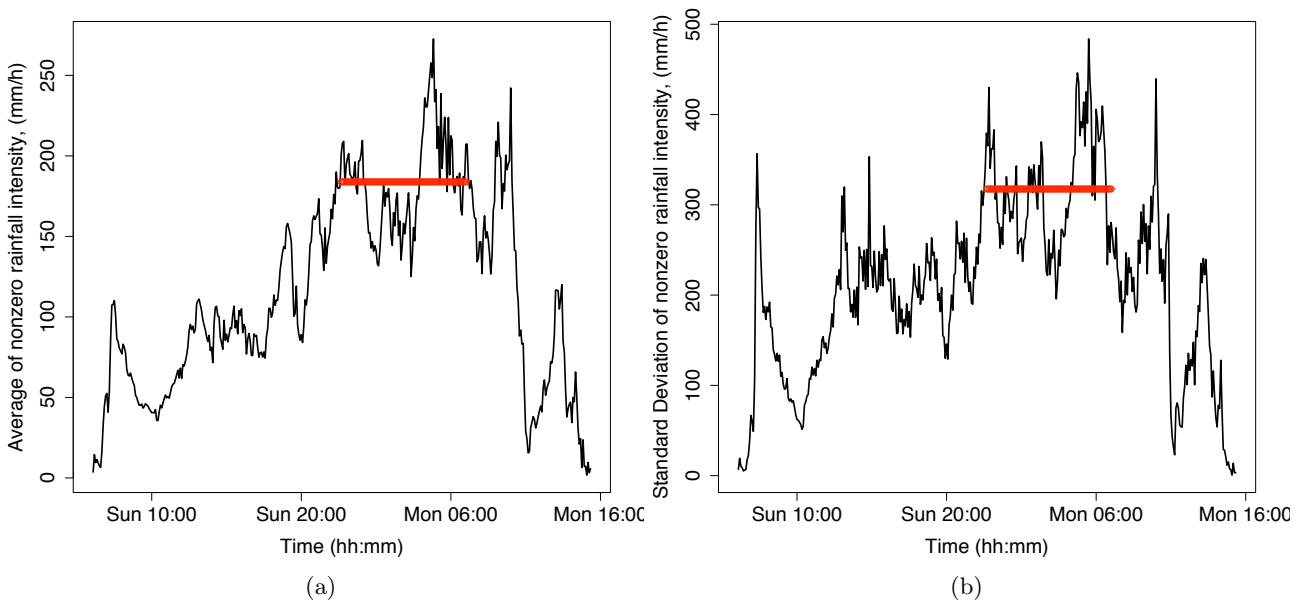


Figure 9.1: Rainfall event of 2002, 08-09 of September on the Cévennes Region. a: Average nonzero radar-estimated rainfall as a function of time. b: 2nd raw moment of the radar-estimated rainfall as a function of time. The horizontal line identifies the time window chosen for the analysis.

9.4 Scale Invariance: Generalized structure Function

The analysis of the average power spectrum is needed to determine which scale-invariant analysis can be safely performed. We analyzed each of the 2D spatial fields and averaged the results. In the present case, the average power density spectrum of the considered event (Figure 9.2) is linear in double logarithmic plot, with slope higher than 2, meaning that the spatial multi-fractal analysis (Section 5.5.2) can not be carried out on the observed field, but only on a field properly transformed to obtain conservativity. In alternative, one can characterize the fluctuations of the fields through the use of Generalized Structure Functions (Section 5.5.1), following the approach proposed by *Harris et al.* (2001).

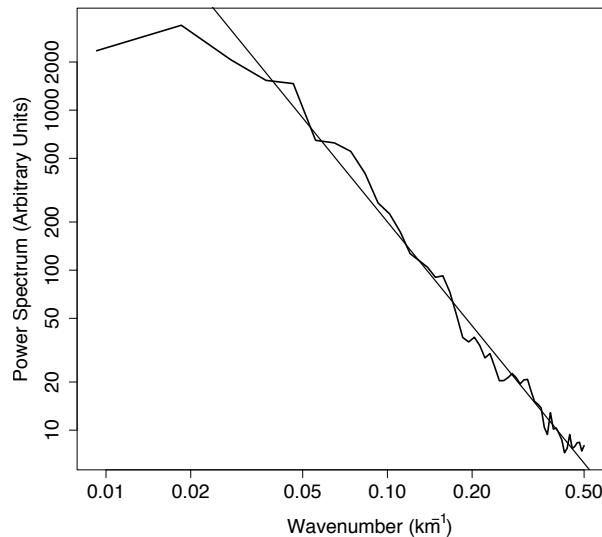


Figure 9.2: Average spectral density power of the event of 2002, September the 8-9th over the Cévennes region. The power spectrum slope is approximately 2.26.

In Figure 9.3 the average variogram (equivalent to the 2^{nd} order SF) for the event of 2002 is plotted, in linear axes, for different time lags, ranging from 5 to 60 min, and for spatial resolutions up to $400km^2$. In the plot, it is easy to see that all the variograms reach the range (i.e. decorrelation distance) at distances in the range 15-30 km. We can compare Figure 9.3 with Figure 9.4, representing the same variograms but this time in double logarithmic axes. The variogram range is obviously the same in the two graphs, and goes from about 10 km for the time lag of 5 minutes to about 20 km for the time lag of 1 h. The variogram shape, on the contrary, is sensibly different. In linear axes, for small lags the variograms present the shape of a spherical or exponential variogram for the small accumulation times, while for higher accumulations it rather show Gaussian shape. Looking at the log-log graph we notice that all the variograms are linear. Deviations from log-log linearity appear for the spatial resolution of $1km^2$ and are probably due to sampling resolution issue (the field resolution is 1 km). This finding is the evidence of a scale-invariant behavior of spatial rainfall for sizes lower than 10-20 km. The spatial scale-invariance range (i.e. of power-law Generalized Structure Function) increases with the time lag. For small lags, we observe that the variogram nugget (the intercept) is

higher for small accumulation durations. This is expression of the small scale variability, higher in instantaneous rainfall and smoothed in time-aggregated rainfall.

This analysis highlights that scale-invariance of spatial rainfall exists. However, due to the presence of characteristic scales of rainfall processes (the concept of variogram range itself expresses presence of a characteristic scale), the scale-invariance is limited to 100-400 km^2 as a function of the accumulation time. This may limit the applications of scale-invariant methods for hydrological purposes.

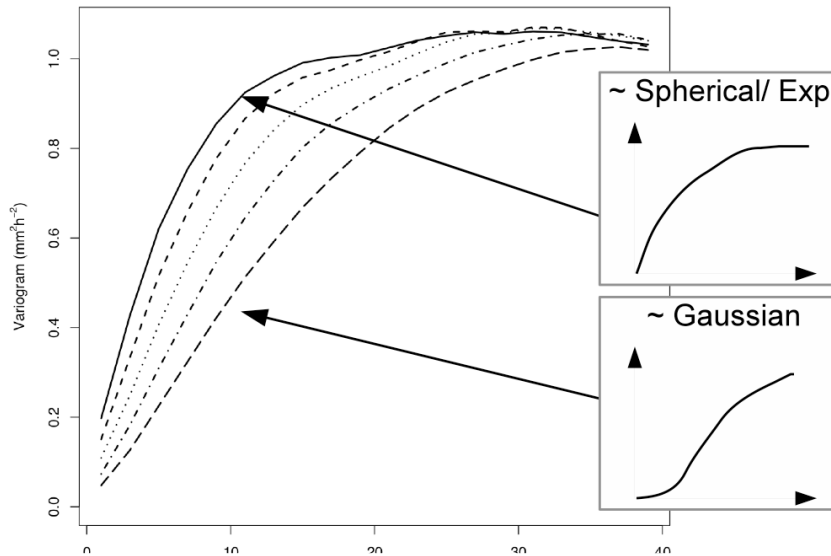


Figure 9.3: Mean variogram of normalized rainfall as a function of the distance between the observed data for different time lags: solid line: 5 min; dashed line: 15 min; dotted line: 30 min; dash-dotted: 45 min; long-dotted: 60 min.

9.5 Scale invariance in Space-Time

The paper presented by *Venugopal et al.* (1999) showed, for the first time, that a direct analogy between the temporal and the spatial domain could be found on high-resolution rainfall data such as radar-estimated rainfall. They start by considering a period within a storm in which the main characteristics of the field (average rainfall intensity, standard deviation) were, as more as possible, stationary. This solution is necessary for the removal of drift effects due to non stationary data, similar to those observed in the geostatistical analysis (Section 4.4.2).

Venugopal et al. (1999) showed that the rainfall fluctuation fields strongly depend on the background intensity of the field. They therefore suggest to deal with a normalized variable, such as the field normalized by the average of the process. A possible solution is to use the variable $\Delta \log I_{i,j,\tau} = \log(I') - \log(I)$. The resulting variable, due to the non-linear transformation and being the expression of fluctuations, is centered at zero and approximately gaussian-diftributed.

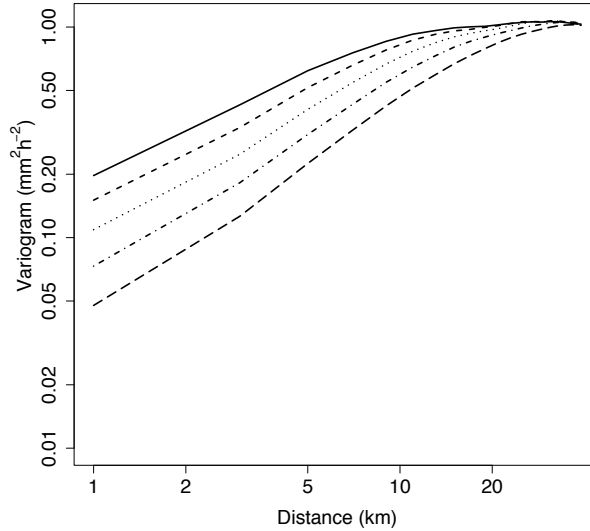


Figure 9.4: Mean variogram of normalized rainfall as a function of the distance between the observed data in double logarithmic axis for different time lags: solid line: 5 min; dashed line: 15 min; dotted line: 30 min; dash-dotted: 45 min; long-dotted: 60 min.

9.5.1 Connections with Geostatistics

In their analysis, *Venugopal et al.* (1999) deal with the statistical characterization of the field

$$\Delta \log I_{i,j,\tau} = \log I_{i,j}^L(t + \tau) - \log I_{i,j}^L(t) \quad (9.1)$$

where $I_{i,j}^L(\tau)$ is the nonzero rainfall intensity at location (i, j) at the spatial scale L . $\Delta \log I_{i,j,\tau}$ represent fluctuations and therefore has null average. Therefore, the analysis performed by *Venugopal et al.* (1999), consisting in studying the variance of $\Delta \log I_{i,j,\tau}^L$ for different aggregation sizes L and time lag τ . This exactly corresponds to the temporal variogram (Section 4.4) of a log-transformed spatial field varying the spatial scale L and the time lag τ ;

$$\text{Var}[\Delta \log I_{i,j,\tau}^L] = \text{Var} [I_{i,j}^L(t + \tau) - \log I_{i,j}^L(t)] = \gamma_\tau(\log(I^L)) \quad (9.2)$$

where t is the time coordinate. Excepted for the use of logarithm, the $\text{Var}[\Delta \log I_{i,j,\tau}^L]$ computed by *Venugopal et al.* (1999) is the dual of the spatial variogram for various accumulation periods that we computed in Section 9.4.

For each spatial aggregation size L and for each time lag τ , we may compute the standard deviation σ of the log-fluctuations $\Delta \log I_{i,j,\tau}^L$ for the 2002, 8-9th September storm. This allows to draw iso- σ maps, corresponding to iso- γ_τ lines.

9.5.2 Space-time scale-invariance

Following the approach of *Venugopal et al.* (1999), we focus our attention on the fluctuation of rainfall fields in the same location (i, j) from the instant t to the instant $t + \tau$. In order to analyze the evolution

of the rainfall fluctuations at different scales, we averaged the rainfall evaluated by radar over a box of size $L \times L$ centered around the location (i, j) . Performing aggregation in space as well as in time, we can observe the changes in the rainfall variability with the scale. We then compare the histograms of these fluctuations (computed only on the positive observations) at different spatio-temporal scales. If space-time scaling holds, we should expect straight lines in the $\log L$ versus $\log \tau$ plot for all the couples (L, τ) respecting a given ratio τ/L^z . If a unique z is found for all couples (L, τ) , the “dynamic scaling” ratio z could be computed as the average slope of these straight lines. The distribution of $\Delta \log I_{i,j,\tau}^L$ is evaluated at different temporal and spatial scales and is shown in Figure 9.5. The figure shows the histogram for 4 couples of surface and aggregation periods; it can be seen that the histogram of $\Delta \log I_{i,j,\tau}^L$ for the surface of 10 min and 4 km² has approximately the same standard deviation as that for 40 min and 36 km². The same reasoning can be performed for any space-time combination respecting the $\tau \sim L^z$ expression.

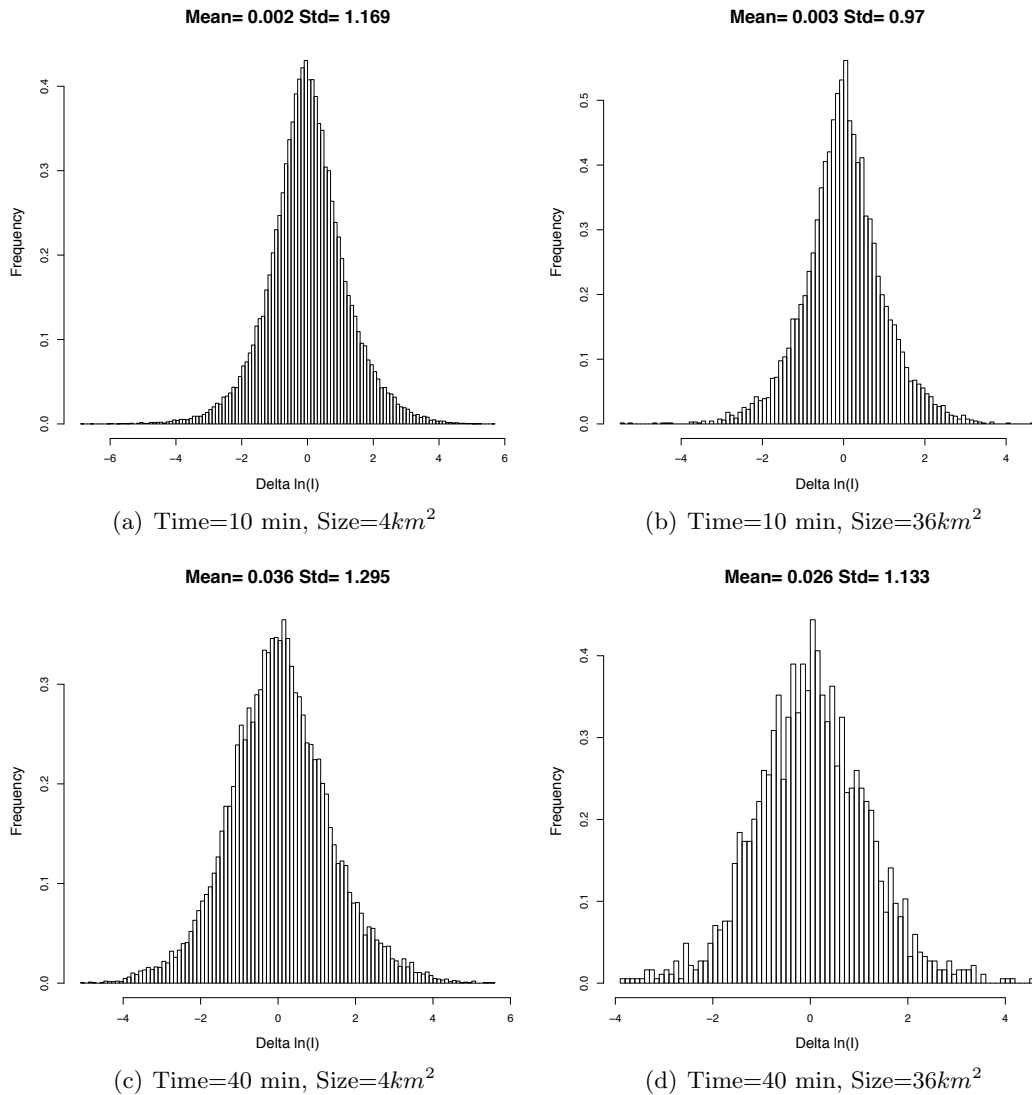


Figure 9.5: Histograms of $\Delta \log I_{i,j,\tau}^L$ for different (τ, L) couples.

9.5.3 Evaluation of dynamic scaling

Evaluating the standard deviation of $\Delta \log I_{i,j,\tau}^L$ for any combination of aggregation time and surfaces, we can draw iso-standard deviation maps as a function of space and time. The diagram of iso-standard deviation curves for the considered event is plotted in Figure 9.6. As we can easily see from the graph, the iso-standard deviation curves are linear in log-log plot and have similar slope. It means that a dynamic scaling relation, linking spatial and temporal rainfall, can be defined. From Figure 9.6, we can define the temporal range of validity of the dynamic scaling assumption, limited to approximately 45 minutes. Evaluating the slope of these iso-standard deviation curves we can estimate the dynamic scaling exponent. Figure 9.7 shows the evaluation, for different iso-standard deviation lines, of the dynamic scaling exponent by least-squares linear fitting. The slope of each iso-standard deviation line is comparable; the dynamic scaling exponents evaluated for different iso-standard deviation lines are reported in Table 9.1.

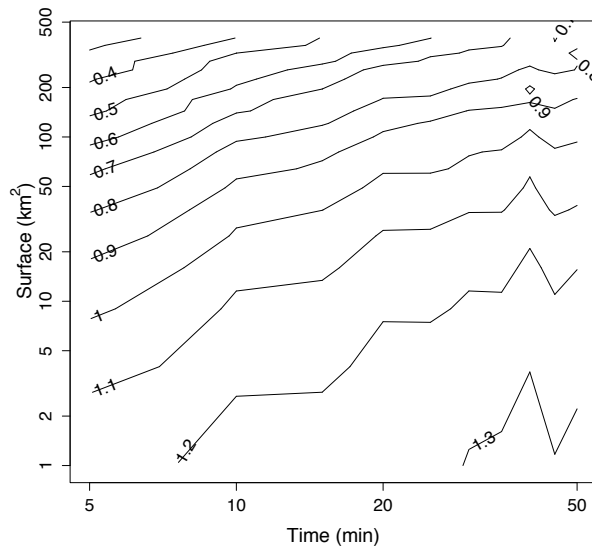


Figure 9.6: Iso-contour diagram of $\sigma[\Delta \log I_{i,j,\tau}^L]$ for each couple (τ, L) .

9.6 Final remarks

The evaluation of the scale-invariant properties of radar rainfall fields for an extreme event occurred in Southern France has been performed through the methods presented by *Harris et al.* (2001) and *Venugopal et al.* (1999). The average spectral slope of the process indicates non-conservative fields and prevents the use of universal multi-fractal unless a field transformation is performed. The Generalized Structure Function analysis show that the spatial scale invariance holds up to the size of 20 km for aggregation periods up to 45 min.

Examining the fluctuations of the log-transformed rainfall fields a dynamic scaling exponent z , relating spatial aggregation size and time lag, has been found: the iso-standard deviation curves of the field fluctuations are linear in log-log plots if the scale ratio $\tau \sim L^z$ is respected (Figure 9.7). The

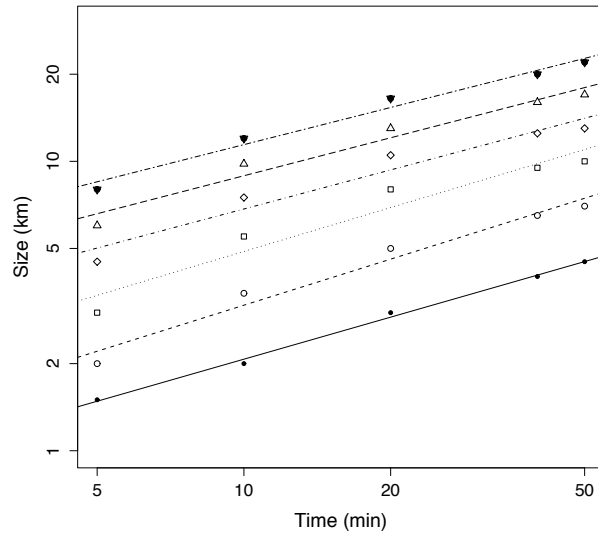


Figure 9.7: Evaluation of the dynamic scaling exponent for several couples (τ, L) .

Table 9.1: Dynamic scaling exponent evaluation. For different iso- σ lines, the spatial size corresponding to a given time (columns) is reported. The values marked with a star are outside the graph of Figure 9.7. In the bottom, the estimation of the dynamic scaling exponent, with the value of the standard deviation.

Iso- σ lines	Time (min)					z
	5	10	20	40	50	
1.2	1.5	2	3	4	4.5	0.48
1.1	1.5	3.5	5	6.5	7	0.63
1.0	3	5.5	8	9.5	10	0.50
0.9	4.5	7.5	10.5	12.5	13	0.45
0.8	6	9.8	13	16	17	0.44
0.7	8	12	16.5	20	22	0.43
0.6	9.5	13	18	24*	26*	0.44
0.5	11.5	18	-	-	-	-
0.4	15	20	-	-	-	-
	\hat{z}	0.48		$\sigma(\hat{z})$	0.07	

log-transformation of the fields is necessary in order to obtain approximately gaussian fields, with null average, for which the standard deviation completely describes their probability distributions. Following this approach, we obtain a dynamic scaling ratio ($z = 0.48 \pm 0.07$) that can efficiently describe the relation between spatial and temporal scales in spatial ranges up to 20 km and temporal ranges up to 45 min.

9.7 Conclusion and perspectives

The merging of ground measurement with radar scans is a solution for generating a reliable radar-rainfall fields. Radar imagery can be efficiently used to determine structural features of the rainfall fields, such as the scale invariance. The aim of the chapter is to define the validity ranges of scale-invariance in space and in space-time.

In this study, we have examined the 2^{nd} order Structure Function, or variogram, of the rainfall field, determining the correlation structure of the fields at various accumulation periods. We show that the variogram of positive rainfall observations can be modeled with a power-law function, expression of the scale invariance of rainfall in space.

Reproducing the study of *Venugopal et al.* (1999), focusing on the ratio between spatial and temporal scales, we aimed to characterize the “dynamic scaling” ratio. We found evidences of a constant dynamic scaling exponent throughout a single storm, based on the study of the fluctuations of non-linearly transformed rainfall at different spatial and temporal aggregations. These findings have been analyzed from the innovative point of view of the comparison with the geostatistical and the multi-scale analysis.

The limitation of the work is that the event is not representative of the totality of rainfall events in the region, and therefore the results cannot be generalized.

Even though the presence of a dynamic scaling ratio is promising, the results of this study are not directly exploitable for the disaggregation of spatial rainfall data, due to the limited representativity of the analyzed event with respect to the ensemble of the storm types in the region. In perspective, repeating the same analysis for a number of storms in the region could give the elements for a proper statistical downscaling of spatial rainfall fields such as radar or satellite imagery.

10

Qualification of Meso-scale meteorological simulations

Résumé

Une partie consistante de la thèse a été consacrée à la description des propriétés des pluies extrêmes face aux changements d'échelle temporelle et spatiale. Un modèle continu Intensité-Durée-Fréquence a été établi (Section 5.11) permettant la détermination de la probabilité d'occurrence des événements de précipitation ponctuels se produisant dans la région Cévennes-Vivarais. Le couplage des courbes IDF avec un modèle du facteur de réduction surfacique (ARF, Section 6) permet d'obtenir la fréquence d'occurrence des pluies spatiales (modèle IDAF). Cette évaluation est possible dans la gamme d'échelles où l'invariance d'échelle est confirmée: plage temporelle de 1-100 h et spatiale 0-3000 km². Une application dans le contexte du projet MEDUP ("Prévisions et projection des événements Méditerranéens intenses dans le scénario climatique: incertitudes et propagation sur l'environnement) est la construction des diagrammes de sévérité (Ramos et al., 2005) dans la région Cévennes-Vivarais. Cet outil, récemment développé, a été adopté pour détecter et comparer la sévérité des épisodes de pluie forte dans des bassins versants en milieu urbain, son utilisation peut être étendue à une plus grande région caractérisée par une densité plus faible et par une hétérogénéité du régime pluviométrique.

En particulier, nous proposons d'utiliser les diagrammes de sévérité comme une approche complémentaire pour évaluer la capacité des modèles météorologiques à reproduire la structure de la pluie extrême à toutes les échelles spatio-temporelles. Dans la section suivante, la construction de diagrammes de sévérité et leur utilisation et qualités sont détaillées.

10.1 Introduction

A consistent part of the thesis has been devoted to the description of the modifications that heavy rainfall intensities undergo as a result of temporal and spatial scale changes. A continuous IDF model has been established (Section 5.11), and the frequency of any point rainfall event occurring in the region may be determined, for temporal ranges of 1-100 h. A continuous Areal Reduction Factor (ARF, Section 6) model will be proposed, in order to estimate the occurrence frequency of spatial

rainfall in the above cited temporal ranges and in the spatial range 0 - 3000 km^2 . An application in the context of the MEDUP ("Forecast and projection in climate scenario of Mediterranean intense events: Uncertainties and Propagation on environment") project is the construction of Severity Diagrams (*Ramos et al.*, 2005) in the Cévennes-Vivarais region. This tool, recently developed, has been adopted to detect and compare the severity of heavy rainfall events occurred in urban catchments, and now its use can be extended to a larger region characterized by coarser rain gauge density and heterogeneous climatic behavior.

In particular, we propose to use severity diagrams as a complementary tool to evaluate the performance of meteorological models in reproducing the rainfall space-time structure. In the next section, the construction of Severity Diagrams and their utility are detailed.

10.2 Severity Diagrams

The Severity Diagram is a multi-scale tool for the description of heavy meteorological events. *Ramos et al.* (2005) proposed its use for discriminating between three storms occurred in the urban catchment of Marseille, France. The three storms did not show significant differences in terms of maximum rainfall intensity, total rainfall depth and duration to explain the marked difference in terms of damages that they produced (Figure 10.1). Actually, these single-scale descriptors of the storms are not sufficient to give all the necessary details, and in particular they do not focus on the critical scales of the event. Storms can be of very different kinds in nature, from deep convective events lasting few minutes to long stable convective systems characterized by weak instantaneous intensity but considerable rainfall amounts. Each storm has a critical temporal and spatial scale.

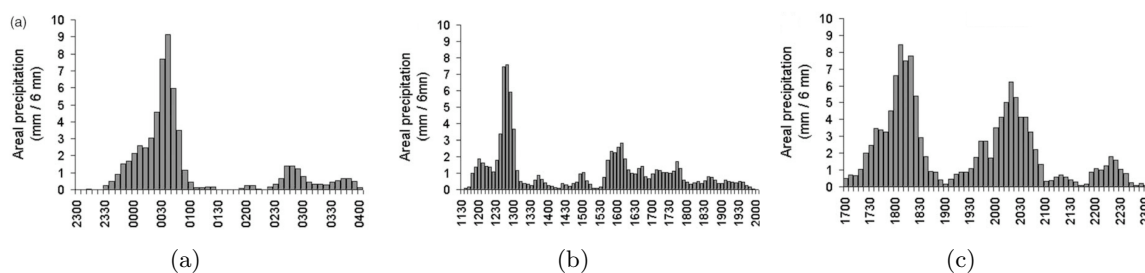


Figure 10.1: Time distribution of mean areal precipitations for the storm events observed in Marseille. (a) 22-23 September 1993, (b) 7 September 1998, and (c) 19 September 2000. From *Ramos et al.* (2005)

Adopting severity diagrams (Figure 10.2), *Ramos et al.* (2005) characterized the magnitude of the three events and give elements for the evaluation of the related damages. It consists in representing, for each temporal and spatial scales (x and y axes), the maximum return period of the event (i.e. the occurrence frequency of the event). The chronological order and the spatial location of the maxima are lost, and the x and y axes represent the accumulation duration and the aggregation surface (and not the time scale and the position in space), respectively. From the diagrams in Figure 10.2 it appears that the event occurred in 2000 had the highest magnitude, interesting an extended zone and leading to return periods higher than 40 years for extended surfaces (up to 200 km^2). The urban catchment was severely damaged by this event, leading to two fatalities and about 60 million euros damages. This gives credit to severity diagrams as descriptors of the space-time structure of an extreme event.

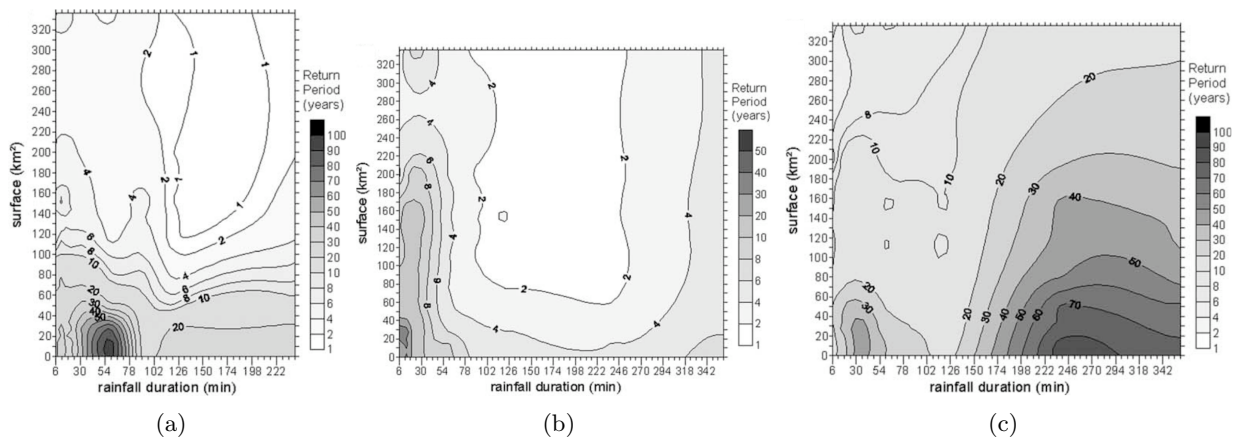


Figure 10.2: Severity diagrams for the storm events observed in Marseilles. (a) 22-23 September 1993, (b) 7 September 1998, and (c) 19 September 2000. From (Ramos et al., 2005)

The construction of severity diagrams is complex. It consists in the coupling of two widespread engineering design tools, Intensity-Duration-Frequency curves and Areal Reduction Factors, Ramos et al. (2005). IDF curves are generally used for determining the rainfall intensity that can fall in a given interval of time with a given recurrence interval. ARF curves report the rainfall intensity decay of rainfall extremes with the increase of the integration surface.

The process to create severity diagrams can be summarized in three steps:

- i. the region of study is scanned with a moving window of given size, the spatial rainfall is derived by arithmetically averaging the rainfall measurement included in the window (Figure 10.3); for each aggregation surface and duration, the maximum rainfall intensity is recorded;

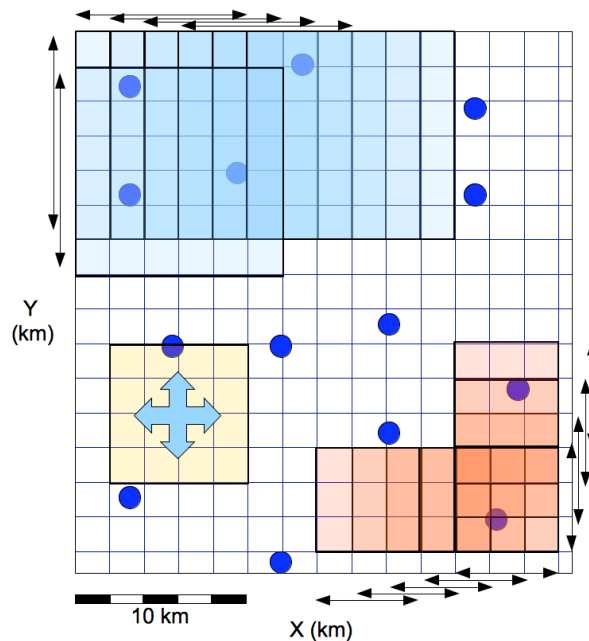


Figure 10.3: Moving window technique for detecting maximum rainfall. Different window sizes are shown: 36 km^2 (orange), 64 km^2 (yellow), 144 km^2 (blue). The dots represent the ground rainfall measurement points.

- ii. using the Areal Reduction Factors, it is possible to estimate the “equivalent point rainfall” to be associated with the areal rainfall value; the equivalent point rainfall and the areal rainfall are characterized by the same probability of occurrence;
- iii. using Intensity-Duration-Frequency curves with the equivalent point rainfall, the frequency of the event is estimated and reported into the severity diagram.

In addition to the comparison between events, the visualization of storm severity can be an useful indicator in an operational context for the real-time survey of urban catchments. The advantage is that this tool provides an objective characterization of the storm, but its use is limited by the complex implementation and by the relative difficulty to be understood by technical operators, due to the loss of the chronological sequence and of the spatial location of the storm.

In our context, we propose the use of severity diagrams to qualify the performances of meteorological models in reproducing the actual space-time structure of extreme events.

10.3 Rainfall intensity diagrams: an indicator of the true model resolution

Taking as a reference the event of September 2002 in the Cévennes Region (*Anquetin et al.*, 2005), *Labalette* (2009) performed a series of preliminary analyses on a series of meteorological models. The aim was to verify the space-time structure of the rainfall fields predicted by the model.

As a preliminary analysis, we constructed the maximum intensity diagrams. They represent, at each time accumulation and spatial scale, the maximum rainfall intensity recorded for a selected storm. It corresponds to the step 1 indicated in Section 10.2 of the construction of Severity Diagrams.

The compared model are:

- i. BOLAM: BOlogna Limited Area Model, hydrostatic model working in a nested framework and forced by the ECMWF analysis, operating at 0.05 ° resolution;
- ii. MOLOCH: non-hydrostatic high-resolution model that integrates the fully compressible set of equations, nested into the higher resolution BOLAM simulation;
- iii. RAMS: Regional Atmospheric Modeling System of LaMMA Meteorology laboratory at Florence (Italy), whose output are provided on a 2 km grid.
- iv. MM5: research non hydrostatic meso-scale model proposed by UIB, Spain, use a three two ways nested grids at 54, 18 and 6 km resolution; the large-scale forcing is provided by NCEP analyses.
- v. MesoNH: Atmospheric Simulated System (*Lafore et al.*, 1998), realised by Météo-France with CNRM (Centre National de Recherches Météorologiques) and Laboratoire d’Aerologie de Toulouse, is a research tool for small and meso-scale atmospheric process, non-hydrostatic. In the specific case, it runs over two two-way nested grids with resolution of 9.5 and 2.4 km, respectively. The large-scale forcing is provided by the ARPEGE analysis.

The construction of maximum intensity diagrams helps in detecting the true spatial resolution of the model. In Figure 10.4 the diagrams are shown for the model Meso-NH, RAMS and BOLAM, respectively. In Figure 10.4 it is possible to notice that RAMS works at a coarser spatial resolution compared to MesoNH, because no difference in terms of rainfall intensity is seen for area lower than 500 km^2 .

Concerning the intensities, MesoNH predict a significant rainfall amount, resulting in a maximum underestimation of 20-30 % compared to the model. Even if the result in terms of rainfall depth is comforting, large underestimations in terms of return period are expected. The two other models, whose diagrams are not shown here, are unable to reconstruct the maximum intensity pattern of the event.

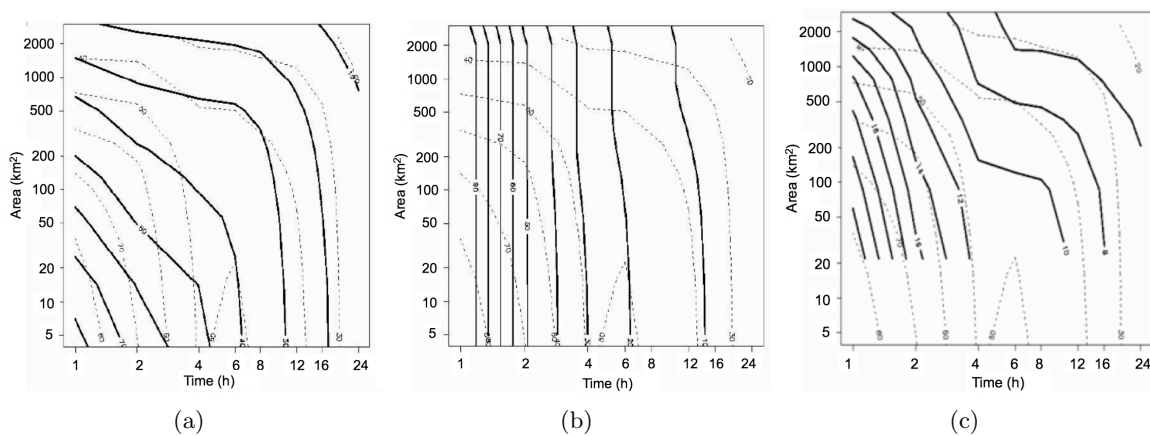


Figure 10.4: Maximum intensity diagrams for the event of 8-9 September 2002 for three research meso-scale models. a: Meso-NH b: RAMS c: BOLAM. The dashed iso-lines represent the observed diagram, derived from ground measurements.

As we have seen, the construction of maximum intensity diagram is a preliminary analysis that can be conducted to characterize the the maximum rainfall observed at each scale. It gives information on the over/under estimation of the rainfall intensity provided by the model and allows to detect the actual model resolution. For instance, we have shown that the model RAMS is unable to reproduce the space-time structure for areas lower than 1000 km^2 .

10.4 Article. Severity diagrams: a new approach for the multi-scale evaluation of extreme rainfall events

Submitted to *Weather and Forecasting*, November 2010

Severity diagrams: a new approach for the multi-scale evaluation of extreme rainfall events

DAVIDE CERESSETTI , SANDRINE ANQUETIN *, GILLES MOLINIÉ

LTHE, Laboratoire d'Etudes des Transferts en Hydrologie et Environnement - (CNRS,UJF,IRD,INPG), Grenoble, France

ETIENNE LEBLOIS

CEMAGREF, Lyon, France

JEAN-DOMINIQUE CREUTIN

LTHE, Laboratoire d'Etudes des Transferts en Hydrologie et Environnement - (CNRS,UJF,IRD,INPG), Grenoble, France

ABSTRACT

The comparison between observations and simulations of relevant rainfall events is usually done by analyzing i) the total rainfall depth produced by the event and ii) the location of the rainfall maximum. We propose in this paper a challenging approach that compares meso-scale simulated rainfall intensities with the ground rainfall observations in a multi-scale framework (i.e. the severity diagram). The severity diagram displays the maximum occurrence frequency of the rainfall intensities as a function of the spatial and temporal aggregation scale, highlighting the space-time scales mainly concerned by the event. For the application in a complex-relief region, a generalization of severity diagrams has been implemented in order to incorporate the regional behavior of heavy rainfall events. To demonstrate the efficiency of this approach, three major storms that occurred in the last decade over a Mountainous Mediterranean region of Southern France are analyzed and compared with the MesoNH simulations outputs. Thanks to severity diagrams, it is possible to detect the critical space-time scales of a rainfall event, and to compare them with those predicted by the simulation. This validation approach can be easily adapted to evaluate the simulation performance of various types of storm, even involving different regions.

1. Introduction

In the recent years, a relevant number of Mediterranean storms caused serious social and economic damages in Southern Europe. These events hit different zones (coasts, foothill or flat areas) and have involved disparate spatial and temporal scales. Recent researches (Ramos et al. 2005; Ruin et al. 2008) showed that fatalities related to extreme meteorological events occur both in small and large basins, characterized by considerably different spatial and temporal characteristic scales. The variability of heavy rainfall events changes with the scale of analysis: two events occurred at two different scales can not be compared based only on the average rainfall depth or rainfall intensity (Bousquet et al. 2006; Yates et al. 2007).

These considerations highlight the need of an objective evaluation of the impact of storms at all spatial and temporal scales for a better understanding of i) storm structure, ii) critical scales of the storm iii) hydrological impact of the storm.

Similar considerations can be drawn in the context of

the evaluation of meteorological models: the performances of meso-scale models is usually done comparing the simulated rainfall fields at a spatial resolution of the order of 1 km^2 to the ground measurements at the rain gauges. A series of scores have been presented with the aim to compare model outputs and observations. In literature (Mason (1989); Ducrocq et al. (2002); Venugopal et al. (2005) among others), scores derived from contingency tables like the probability of detection (POD) or false alarm rate (FAR) mainly qualify the capability of the model to simulate the rainfall depths and storm locations. Some of these indicators focus on the extreme values or quantiles and few of them on the capability to reproduce the whole storm structure at all scales (Zepeda-Arce et al. 2000; Yates et al. 2007).

Another objective evaluation approach allowing multi-scale comparison of rainfall events has been presented by Ramos et al. (2005). The aim of “Severity Diagrams” was to represent the magnitude of a storm over a range of spatial and temporal scales in a normalized framework adapted for inter-event comparisons. Intending the “severity” in

probabilistic terms, severity diagrams represent, for each combination of aggregation time and integration surface, the return period (expressed in years) of the storm.

In this paper we show that such an approach can be efficiently used for the multi-scale comparison between the meso-scale simulated rainfall intensities and point observations. The implementation of severity diagrams requires the knowledge of the extreme-rainfall behavior in the region. We show that a preliminary comparison between observations and simulations could be derived through the analysis of an intermediate product of the severity diagram implementation, the “maximum intensity diagrams”, representing the maximum rainfall intensity for each spatial and temporal scale. Maximum intensity diagrams could be useful to compare events among them, but they are scale-dependent. Severity diagrams, thanks to a transformation of the maximum rainfall intensity at each scale into return period, allows either inter-scale comparisons of single storms or comparison between storms, with an intuitive representation in terms of return period.

In urban area target of the Ramos et al. (2005) study, the dense high-resolution raingage network and the climatic homogeneity are two main factors simplifying the spatial-rainfall frequency estimation. Differently from the application of Ramos et al. (2005), the present paper aims at the characterization of storms occurred over a larger region. The Cévennes-Vivarais region (Southern France) features inhomogeneity of the extreme rainfall behavior; a regional approach is therefore required. Due to its localization, close to the Mediterranean Sea, and its complex topography, this region is particularly prone to heavy rainfall events and flash-floods (Delrieu et al. 2005; Nuissier et al. 2008; Ducrocq et al. 2008).

The construction of severity diagrams is based on the coupled application of the Areal Reduction Factors (ARF) (NERC 1975; Rodriguez-Iturbe and Mejía 1974; Bacchi and Ranzi 1996; Sivapalan and Blöschl 1998; Asquith and Famiglietti 2000) and the Intensity-Duration-Frequency (IDF) curves (Burlando and Rosso 1996; Koutsoyiannis et al. 1998).

The paper is structured as follows: the region and the data set used for this study are presented in Section 2. The methodology is fully developed in Section 3. Section 4 briefly introduces the three cases. In Section 5 the severity diagrams are drawn for the three storms using the observed and simulated rain fields. For a particular rainfall event, the analysis of the severity diagram aims at exploring the hydrological effects of the storm. The discussion is reported in Section 6 and aims at giving some inputs on the advantages and limits of severity diagrams for the evaluation of meteorological models. Then, synthesis and conclusion follow in Section 7.

2. Data description

a. Region of interest

The studied region is located between the Massif Central and the Mediterranean Sea and is bounded by the Rhône River in the Eastern side. The area covers a window of about $160 \times 200 \text{ km}^2$, gathering a coastal zone, a large plain, a mountainous region and a high plateau (Figure 1).

The rainfall regime in the region has been widely studied (Lebel and Laborde 1988; Bois et al. 1997; Molinié et al. 2010). The region is prone to flash floods which caused several social and economic damages in the last decades. As an example, one of the heaviest event occurred in 2002. It caused 50 million euros damages and consistent human losses (Ruin et al. 2008). Relatively long rainfall series have been recorded in the region by several services (the French meteorological agency Météo-France, water management and power supply agencies, Electricité de France). Since 2000, the Mediterranean Hydrometeorological Observatory Cévennes-Vivarais centralizes the data and ensures a quality control. Two kinds of rainfall data are used for this study: i) the observed rain gauge data provided by OHMCV and ii) the rainfall fields simulated by the Meso-NH model by Météo-France.

b. Raingage rainfall data

The data set includes hourly and daily rainfall intensity series. The daily rainfall database (Figure 1) is composed of 225 stations featuring more than 30 years of continuous records. The hourly rainfall database is composed of 150 continuous raingage records over the 1993-2008 period. Figure 2 reports the average density of the hourly and daily raingage networks as a function of the elevation.

c. Simulated rainfall data

The simulated rainfall-intensity fields are available for specific severe rainy events thanks to the research program MedUP. They are the product of the cloud resolving model MesoNH (Meso-scale Non-Hydrostatic, Lafore et al. (1998)). MesoNH is run on two-way nested grids at 9.5-km and 2.4-km resolution, respectively. The coarser MesoNH domain is driven by the limited area model ALADIN. The finer-scale MesoNH domain is centered over the northwestern Mediterranean where the studied rainy events initiated and includes the study region. The rainfall fields are provided at the spatial resolution and temporal resolution of 2.4 km 1 hour, respectively.

3. Methodology

The implementation of the severity diagram requires three steps as schematized in Figure 3 and includes a study on the spatial-rainfall extreme climatology. The first step (Section 3.a) consists in building the spatial rainfall database

by interpolation of the rain gauge observations. Thanks to this database, we derive Areal Reduction Factors (ARF) expressing the rainfall intensity decrease as the integration surface increases (Section 3.b). As severity diagrams require to determine the frequency of occurrence of spatial rainfall intensities at any accumulation period and surface, a continuous ARF model has to be coupled to IDF relationships. The IDF model adopted in this study is described in Section 3.c.

a. Spatial rainfall database

The spatial rainfall database was built from the hourly rainfall intensity database. The rainfall point observations were cumulated over periods of 2, 4, 8, 12 and 24 h. For each accumulation period, the spatial-rainfall database is built following three steps:

- i. Definition of rainy events: evaluate the number of working raingages and the average rainfall depth measured by the network. When these two indicators are lower than a fixed threshold, the field is rejected because the event is probably weak and local.
- ii. Determination of the spatial structure of the rainfall field: if the field is retained, compute the variogram of the positive rainfall values; a spherical variogram model is associated to the field; if the variogram has singular behavior (it is the case when the number of working raingages is low), force a “climatological variogram” to the field (Lebel and Laborde 1988).
- iii. Interpolation: using the variogram model above defined, compute the kriging interpolation over a regular grid. A grid spacing of $2 \times 2 \text{ km}^2$ has been chosen, in agreement with the resolution of the meteorological-model output.

b. Areal Reduction Factors

The maximum rainfall intensity of a storm at a given accumulation duration decreases with the surface. This property can be used to derive a probable rainfall intensity level for a given surface A when a single observation is available, by applying Areal Reduction Factors. The Areal Reduction Factor (ARF) is defined as the average ratio of the areal-rainfall observation over the surface A to the point-rainfall observation corresponding to the same frequency level.

ARFs are equal to 1 for $A \rightarrow 0$ and decrease with the surface of aggregation. ARFs usually increase with the accumulation period, as a consequence of the increase of the characteristic size of the storm. The relation with the return period T_R is also intuitive: the higher the rainfall intensity, the lower the ARFs (i.e. very intense events are generally more localized).

The empirical ARF curves can be derived from the spatial rainfall database. Omolayo (1993) showed that statistically-significant ARF curves can be derived in a fixed-area framework, while ARF based on a storm-centred approach are significantly underestimated.

In practice, storm-centred ARFs evaluate the rainfall intensity decay of selected events over concentric windows of increasing size, while fixed-area ARF curves rely on the maxima over moving windows of increasing size. In the former approach the concomitance between the maxima observations is not required. The loss of physical significance of fixed-area ARF curves is compensated by the gain in statistical significance.

According to empirical evidences (Bacchi and Ranzi 1996), ARF depends on the frequency of the events considered for its computation. The ARF computed heavy rainfall events is steeper than the ARF computed gathering heavy and regular rainfall events. Weak rainfall events can be extended over large regions, leading to smoother ARF curves, while for instance deep convective events are rather localized, leading to steeper ARF curves. For this reason, ARF is computed on a series of independent and identically distributed (i.i.d.) events characterized by return period equal or higher to a fixed value. For this study, an average of 2 events per year (return period = 0.5 years), corresponding to 32 rainfall values for each considered surface, have been selected.

The ARF curves are then computed as the ratio between areal rainfall and point rainfall for a given return period T_R :

$$ARF(A, D, T_R) = \frac{I_A(A, D, T_R)}{I_0(0, D, T_R)} \quad (1)$$

where I_A is the areal rainfall over the area A , for the duration D , and I_0 is the point rainfall for the same duration D . By definition, the fixed-area ARF curves are continuously decreasing with the integration surface, and their maximum is 1 (corresponding to the storm center $A = 0$, where areal and point rainfall maxima are equal).

Due to the limited sample set size, the ARF dependence on T_R can not be assessed for large return period. In agreement with previous studies (NERC (1975), for instance), we assume that ARF curves are independent of the return period T_R :

$$ARF(A, D) = \frac{I_A(A, D)}{I_0(0, D)} \quad (2)$$

Up to this point, the ARF curves are empirically computed for a discrete number of surfaces and accumulation durations. Since we are interested on the relation between point and spatial rainfall at any accumulation periods and surfaces, we adopt the ARF model proposed by De Michele et al. (2001).

De Michele et al. (2001) worked on a space-time self-similar model of annual maxima. The model introduced the use of a “dynamic scaling” coefficient, expressing the relation between spatial and temporal scale in the definition of the rainfall intensity. The concept of dynamic scaling, originally introduced by Venugopal et al. (1999) in agreement with the Taylor hypothesis of “frozen turbulence” (Taylor 1938), is physically consistent at small space-time scales (size up to 100 km^2 and accumulation period lower than 1 hour). This hypothesis causes that the temporal variation at fixed locations can be reinterpreted as spatial variation (Deidda 2000). At larger scales, the dynamic scaling is not related to the physics of the phenomenon, conserving only its statistical significance.

The ARF formulation proposed by De Michele et al. (2001) is:

$$ARF(D, A) = \left[1 + \omega \left(\frac{A^a}{D^b} \right) \right]^{-v/b} \quad (3)$$

where v is the scaling exponent of point-rainfall with time, ω a homogenization parameter, a and b express the power-law decay of ARF curves with the integration surface and with the duration; the dynamic scaling z is related to this expression by the relation $z = a/b$.

The comparison of the application of Equation 3 with the empirical results presented by NERC (1975) showed a substantial agreement between the two studies. The fitting of this model requires to take into account the possible undersampling due to the raingage-network density. For these reasons, considering that the average raingage density is of $1/50 \text{ km}^{-2}$, the fitting will not take into account surfaces lower than 50 km^2 and durations smaller than 2 h.

The regional heterogeneity of the extreme rainfall behavior is the main factor limiting the interpretation of spatial rainfall occurrences for large surfaces. The orography, in particular, forces the anisotropy and increases the temporal persistence of the rainfall fields (Prudhomme and Reed 1999; Haberlandt 2007; Berne et al. 2009; Godart et al. 2009). To accurately compute ARF curves, it is therefore proposed to split the region in two domains supposed to be quasi-homogeneous in terms of extreme rainfall behavior (Figure 1), according with the results of previous studies (Ceresetti et al. 2010a): a flat sub-region (Region 1) located in the South-East of the domain of interest extended up to the foothills of the Cévennes massif, and a mountainous sub-region (Region 2, composed by the mountain ridge and the Massif Central highlands located North-West). The anisotropy of ARF curves has not been taken into account due to the limited sample set size. The results of the fitting are shown in Table 1, and plotted in Figure 4.

Figure 4 shows, as expected in both sub-regions, a regular decrease of Areal Reduction Factors with the involved surface, and a corresponding increase of ARFs with time

(corresponding to a similar parameter a in the two regions). The accumulation duration of 1 hour has not been used for the fitting due to the above described undersampling issues. Except for the 1h accumulation duration, the model fits well with the experimental data. The main difference between the mountainous sub-region and the flat area is the effect of the accumulation duration on the trend of ARF curves. The impact of the accumulation period on ARF curves is significantly smaller in the mountainous sub-region (the model shows differences in the value of the b parameter in the two regions). This phenomenon could be physically explained by the persistence of meso-scale convective systems over the mountainous ranges (Sénési et al. 1996; Ducrocq et al. 2003; Molinié et al. 2010).

c. Intensity-Duration-Frequency curves

In order to implement the severity diagram construction, a continuous IDF model is required. For the construction of a regional IDF model we use the daily rainfall database. The IDF curves are used in this context to estimate the frequency of given rainfall observations knowing their intensity and accumulation duration. Dealing with three uncommonly heavy events, the blocking point is the capability to estimate large return periods for infra-daily events: the longest daily series featuring 50 years of data, the uncertainty in the estimation of return periods higher than 100 years is too large to provide a reliable value of the return period.

Following the works on the scale invariance of IDF and Depth-Duration-Frequency curves (Burlando and Rosso 1996; Bendjoudi et al. 1997; Menabde et al. 1999; Borga et al. 2005) it has been possible to implement a scale-invariant model for IDF curves. The model is based on the assumption that, for the region of interest, the maxima are distributed according to a Generalized Extreme Value distribution (GEV). Ceresetti et al. (2010b) showed that at least in the range 4 - 100 h the scale-invariance of maximum rainfall intensities holds, and a GEV simple scaling regional model has therefore been proposed by (Ceresetti et al. 2010b).

Gathering the ARF dynamic scaling model fitting of Section b and the simple-scaling IDF model of Ceresetti et al. (2010b), we derive the Intensity-Duration-Area-Frequency (IDAF) curves for the two sub-region of the domain in agreement with De Michele et al. (2001). Figure 5 shows the IDAF obtained from the 32 heaviest observations over 16 years of data, corresponding to a return period slightly higher than 1 year. It is easy to see that the rainfall intensity decreases with the area and with the accumulation period, because of the smoothing introduced by the spatial and temporal integration. The IDAF model satisfactorily reproduces the empirical behavior of extremes in space and time.

By means of the study presented in this section, we now

have a continuous IDAF model able to provide an estimation of the occurrence probability of any spatial rainfall observation within the study region. We can now turn to the description of the events analyzed in this study.

4. Description of the events

Three storms are studied: Event 1 occurred on December 3, 2003, Event 2 on September 6, 2005 and Event 3 on September, 8-9, 2002. These events differ in their structure, extension and location of the rainfall intensity maxima. They therefore represent an assorted selection of the extreme meteorological situations observed in the region. For the three events, observed and simulated rainfalls are analyzed.

The simulated rain fields are provided by Météo-France and are the outputs of the MesoNH (Lafore et al. 1998) meso-scale model simulations. The model configuration, successfully tested for simulations of Mediterranean intense rainfall events (Ducrocq et al. 2002), is the same for the 3 events. The simulated meteorological fields have been already deeply analyzed in Lebeaupin et al. (2006) for Event 1, Yan et al. (2009) for Event 2 and Nuissier et al. (2008) for Event 3.

The rainfall output is provided on a regular grid in polar coordinates (size of about $2.4 \times 2.4 \text{ km}^2$). It is regularized in a Cartesian grid ($2 \times 2 \text{ km}^2$) by nearest neighbour interpolation.

In the following, the main synoptic features of each event are given.

As explained in Lebeaupin et al. (2006), Event 1 lasted in total 4 days, starting November, 30, 2003 and ending December, 4, 2003. On December 1st, an upper level low-pressure area centered over Spain favored an intense southerly flow over Southern France. A cold surface front established from Northern to Southeastern France. The frontal perturbation formed a MCS that remained until December 3. The 30-hours rainfall accumulation reached 300 mm over the Rhône Valley. The highest daily rainfall amount was reached on December 3. This study focuses on this specific day (from 0000 UTC to 2400 UTC). During this day, the upper level trough area associated with the surface cold front began to turn slowly to a NW-SE axis. After few hours over the Gulf of Lion, it progressed westward in the late evening.

Event 1 mainly involved the Massif Central foothill and mountain ridge (Figure 6-a). The total accumulation did not exceed 200 mm, but the event was extremely extended in space. Approximately 25000 km^2 received more than 150 mm of precipitation, causing a major flood of the Rhône river ($13000 \text{ m}^3 \text{ s}^{-1}$). The maximum rainfall is observed on December, 3, and reached 150 mm. The low-level winds intensified during this day, with easterly winds rushing up to 150 km h^{-1} , causing sea waves of almost 10 m.

The flood caused 7 fatalities in the region.

The large-scale pattern of Event 2 is detailed in Yan et al. (2009). The authors showed that an upper-level cold flow located over the near Atlantic generated a rapid cyclonic upper-level flow over Western France on September, 5, 2005. On September 7, the low-pressure system moved towards the South-East. At lower levels, a low-pressure over Eastern Spain deepened and generated a low-level southerly flow over the Mediterranean. A frontal system with embedded convection over Southern France was responsible for the heavy precipitation amount recorded on 5-6 September. This study focuses on an 18-hours window (starting the 6th of September at 0000UTC up to 1800 UTC).

Between 5 and 9 September 2005, several precipitating systems affected the South-East of France leading to an accumulated rainfall depth higher than 300 mm in most of the region of interest. The night between September, 5 and 6, heavy precipitations fell over the west of Gard department, reaching over 300 mm in the city of Nîmes (Figure 6-c). Despite this important rainfall amount, the runoff process was limited by the dry initial soil-moisture conditions. Weaker precipitations were observed on September, 7, followed by a precipitation event coming from the Mediterranean Sea in the morning of September, 8, affecting the Gard department. The intensity of the rainfall event reinforced during the afternoon, reaching a total rainfall amount of 220 mm near Nîmes. The high soil moisture level in this second part of the event caused the runoff to be significantly higher.

As described in Delrieu et al. (2005) and Nuissier et al. (2008), the first convective cells of Event 3 appeared over the Mediterranean Sea around 0400 UTC on September, 8, 2002. Four hours later, the convection formed a meso-scale convective system (MCS) just south of the Gard department and moving northward. The convective system remained approximately 24 hours over the same region (from 1200 UTC September, 8 to 1200 UTC September, 9). This temporal window is the one analyzed in this study. During this period, a high-level cloud shield displayed a V shape with the V tip facing the upper level southerly flow (Nuissier et al. 2008). Beneath the cloud shield, the convective precipitations mainly affected the Gard region, while stratiform precipitations extended further to the north. At this time, the MCS assumed a southwest-northeast orientation according to the prevailing upper-level tropospheric flow. In the late night of September, 8, the convective system oriented north-south and began a northwestward motion. The precipitating system decayed during the late morning of September, 9.

Event 3 was exceptional from many points of view. The intensity of the event was extreme: the maximum precipitation was around 600 mm in 24 hours near Alès (Figure 6-e). The area affected by heavy precipitations was considerable: more than 3000 km^2 , covering the whole Gard De-

partment as well as the Massif Central foothills in Ardèche, received at least 200 mm of rain. The river discharges were exceptional, especially for the Gard and Vidourle rivers, where peak discharges higher than twice the 10-year return period flow were recorded (Delrieu et al. 2005). The event caused damages estimated to 1.2 billion US dollars and led to 25 fatalities.

In the following, we compute the maximum intensity diagrams and the severity diagrams of the 3 events. We use Event 3 for testing the skills of the approach for assessing the hydrological impact of a storm, isolating the basin where the damages were the most important (i.e. the Gard basin, 6-e).

In Figure 6, the observed and the simulated rainfall depths are plotted for the three events. These figures suggest the following general comments:

- the study domain is large enough to include the totality of the event;
- the simulated rainfall depths are underestimated for Event 2 and Event 3, whereas the model overestimates the rainfall located in the mountainous area (Event 1);
- the locations of the rainfall maxima are approximately well estimated by the model for Event 1; less accurate in Event 2 and 3;
- the spatial extent of the rainfall pattern is captured by the model; nevertheless, the model does not correctly reproduce the intensity and the pattern of the rainfall fields (Figure 6-e and 6-f).

In Table 2 the main thresholded scores (POD: Probability of Detection, FAR: False Alarm Ratio, ACC: Forecast Accuracy) of the simulations are reported for the three events. These scores origin from contingency tables (Yates 1984). The threshold levels have been chosen in agreement with Yates et al. (2007) as the 70th, 80th and 90th percentiles. According to these statistics, Event 1 has been modeled with good accuracy, even though a not negligible False Alarm Ratio is found for the 90th percentile. Event 2 has been badly reproduced, with low POD and elevated FAR for the three thresholds. For Event 3, it is clear that the mislocated maximum rainfall depth has caused very low scores for the 90th quantile. Considering that both the POD and the ACC scores are very poor, it seems that the extreme rainfall has been wrongly estimated for this event. We emphasize the difficulty of using these scores in a multi-scale framework and for hydrologic purposes.

5. Multi-scale evaluation of the simulated events

This section aims at bringing to the fore the interest of a complete multi-scale evaluation of the simulated fields.

We first describe the maximum intensity diagrams. They allow a preliminary multi-scale comparison between events based on rainfall intensities.

We remind that the observed field is obtained by interpolation of the ground measurements at the rain gages. By consequence, the Maximum Intensity Diagrams and the Severity Diagrams of observed fields are affected, in some ranges, by undersampling errors of the measurement network. The spatial undersampling affects areas in the range 0-50 km^2 , for which the raingage density is inadequate. For duration lower than 4 hours, rainfall intensity data are affected by uncertainties due to the resolution of the time series (1 hour).

a. Maximum intensity diagrams

The maximum intensity diagrams report the maximum rainfall intensity recorded during the event for each accumulation duration and for each integration surface. The maximum intensity diagrams relative to the rain gauge observations are given in Figure 7-a, 8-a, 9-a for the three events, respectively. The corresponding diagrams relative to the simulations are reported in Figure 7-b, 8-b, 9-b. The diagrams of Figure 7 and 8 indicate that, in Event 1 and 2, the observed and simulated maximum intensities are in good agreement at accumulation periods larger than 4h. On the contrary, the maximum intensities of Event 3 (Figure 9-a) seems to be poorly reproduced by the simulation at any scale.

Smaller spatial and temporal scales reveal interesting features for Events 1 and 3. The maximum intensity diagrams of Figure 7-a, 9-a and 9-b present sharp singularities. The Event 1 singularity for durations in the range 4-6 hours and surfaces between 20 and 200 km^2 is not present in the diagram of the simulated field. Concerning Event 3, both the diagram of simulation (Figure 9-b) and the diagram of simulations (Figure 9-a) show small scale singularities, but they appear at different scales.

b. Storm Severity: a forecast qualification approach

The severity diagrams allow to perform multi-scale comparisons in terms of return period. Figure 10-a displays the severity diagram of observations for Event 1. The observed severity presents a maximum higher than 300 years involving large temporal and spatial scales (time scales ranging from 8 to 16 hours and for surfaces up to 400 km^2). For small surfaces and durations, the event did not provide significant severities. The simulated fields provided by Meso-NH yield a severity diagram similar to the observed one (Figure 10-b). The maximum severity is of the same magnitude order, as well as the spatial extent of the event. In contrast, the critical time scale has been overestimated (14-18 h against 10-12 h for the observations). In general, the main features of the event seems to be well reproduced

by the simulation.

Concerning Event 2, the severity diagram of observations (Figure 11-a) and simulation (Figure 11-b) are very different in terms of magnitude. Figure 11-a shows that the absolute severity maximum is recorded at scales lower than 4 hours and 50 km^2 . A similar pattern is found for the simulation (Figure 11-b). In addition, a secondary maximum is observed for the duration of 8-12 h and for surfaces lower than 500 km^2 , corresponding to a rainfall amount higher than 200 mm. This severity peak has not been detected in the simulated data.

From these two cases, the interest of drawing severity diagrams for atmospheric model evaluation is clearly demonstrated. Differently from the maximum intensity diagrams, severity diagrams highlights the time and space structure of the storm simplifying the multi-scale comparison between observations and simulation outputs.

c. Hydrological aspects of storm severity: the case of 08-09 September 2002

Since the second half of XX^{th} century, the studies on extreme flood events evidenced that river flow data could not provide reliable estimations for large quantiles (Guillot and Duband 1967; Guillot 1993). Because of the limits of flow measurements, it is preferable to study the occurrence of the rainfall process. Even if the rainfall-runoff modifies the impact of storms on a basin, for rainfall events characterized by return periods higher than 10 years, the transfer function between runoff and rainfall can be reasonably considered equal to 1. This motivates the use of severity diagrams for the assessment of the impact of a storm over a basin.

From a practical point of view, in order to compute the severity of an event within a basin, the method does not change significantly: the only added step is to multiply the rainfall-field matrix by a mask containing the value “1” in the pixels within the studied basin, and the value “0” in the external pixels.

In the present case we deal with the 2002 event (Event 3), involving the Gard basin. The severity is estimated with respect to the Gard basin at the Remoulins outlet (indicated in Figure 1), with a maximum surface of 2200 km^2 . The severity analysis will be conducted on spatial scales lower than the basin surface.

Event 3 is far the largest storm in our records. Consequently, the return period associated with the event can not be correctly assessed due to the uncertainties involved in the extrapolation of the extreme behavior for large return periods. The diagram has therefore been limited to maximum severities of 500 years. The severity diagram related to the observations is reported in Figure 12-a. Despite the large uncertainties in the evaluation of the return period, the severity has a sharp increase with the accumulation duration, reaching severities larger than 500 years

already for the 4-hour accumulation period. The critical scales of the event are reached for the accumulation time of 16-24 hours and a surface of 500 km^2 . It means that a small sub-region within the basin (fortunately not corresponding to a catchment basin) received an impressive amount of rainfall.

The severity diagram for the simulation is reported in Figure 12-b. It is clear that the simulation provides lower severities. The critical space-time scales of the event have not been properly detected by simulation: in Figure 12-b, the severities are negligible for surfaces higher than 50 km^2 . This result is likely to be due to the wrong location of the simulated rainfall maxima (as one can see looking at Figure 6-e and 6-f), leading to poor skills from a hydrological point of view.

6. Discussion

The multiscale analysis of maximum rainfall intensities and return periods (severity) reveals that severity diagrams are most sensitive diagnostics than maximum intensity diagrams:

- i. the simulated rainfall fields of events 1 and 2 are of good quality in terms of maximum intensities in a large range of scales (Figures 7 and Figure 8).
- ii. the severity diagrams of these two events (Figure 10 and Figure 11) show significant differences between simulated and observed rainfall fields.

These differences can have a double origin. The first origin is the non linear transformation of rainfall intensity into occurrence frequency. The return period is obtained as $1/(1 - P)$ where P is the cumulative density function of the extreme value law, obtained by a double exponentiation (in the simplest case) of the intensity (see Kotz and Nadarajah (2000) for further information). Therefore a small difference of maximum intensities is highly amplified when transformed into return period.

The second is that the severity is highly dependent of the location. Taking for example daily rainfall, Ceresetti et al. (2010b) shows that the 100-year return level varies from 100 mm (over the Massif Central plateau) to 400 mm (over the mountain ridge and the southeastern foothill). Therefore, the same storm occurred in two different locations can give very different severities. Severity diagrams incorporate this effect and their analysis can be of interests if one needs to investigate the consequences of the use of simulated rainfall fields as input for hydrological models.

This analysis reveals also that the two diagnostics (severity and maxima diagrams) are complementary. The maximum intensity diagrams have highlighted singularities in Event 1 and Event 3. Some of these singularities are not present in the corresponding severity diagrams. On the other hand, severity diagrams are able to delineate the

critical scales of an event giving additional elements to compare the actual and simulated events. In Event 1, for instance, the critical time scale has been slightly overestimated by the model (16 h instead of 12h). The spatial extent of the event has been correctly modeled, since in both panels of Figure 10 it is possible to detect severities higher than 200 years, along a vertical cut, from 0 to 2000 km². In Event 2, even though the observed and simulated maximum intensity are similar (Figure 7), the severity have an order of magnitude of difference which mean the simulation essentially failed to identify the correct location of rainfall cells.

Throughout our discussion, we did not expressly interpreted the diagrams for time scales lower than 2 hours and spatial scales lower than 50 km². The reason is that, at these scales, the observation network is submitted to undersampling problems leading to uncertainties in the measurement of point rainfall and, in larger measure, spatial rainfall. The same problems affect the simulated fields, since the time resolution of the rainfall fields is 1 hour and their spatial resolution is 6.25 km². We must mention that these undersampling problems affect not only the event measurements, but also the estimation of IDF and ARF curves that are needed to define the climatology of extreme rainfall events.

7. Conclusion

The main purpose of the present work is to show the utility of i) the multiscale assessment of simulated rainfall fields and ii) the assessment of the severity (i.e. return period) of rainfall events at multiple scales. This work proposes an extension of severity diagrams (Ramos et al. 2005) to larger surfaces and accumulation durations. We also introduce maxima intensity diagrams, a preliminary diagnostic showing the maximum rainfall intensity of the event at each space-time aggregation scale. The use of severity diagrams in a regional context gives the possibility to illustrate the effects of a wrong positioning of maxima or over/underestimation of the maxima rainfall depth on the severity of an event. It also detects the spatial/temporal scales in which the model has low skills and gives a support for evaluating whether the overall space-time patterns of given storms have been correctly identified.

From these analyses, it is clear that the severity diagram has mainly three merits: i) it is useful for comparing in an objective way different space-time scales of a single event, detecting its critical scales; ii) it is an useful complementary indicator for the comparison between severe events, as shown by Ramos et al. (2005); iii) it is innovative in its ability to evaluate the meso-scale model skills to reproduce the space-time structure of the rainfall events; iv) its use is independent of the available measurement network and can be successfully adopted in case of

radar-estimated fields.

Despite their numerous skills, the use of severity diagrams presents some limits: i) the spatial heterogeneity of the extreme rainfall distribution in the region may prevent to assign an unitary and reliable severity value to spatially extended rainfall observations; ii) the maximum occurrence frequency that can be assigned to an observation strongly depends on the available rainfall database, and should not exceed the observation period; iii) in severity diagrams the space-time coordinates are lost in favor of the multi-scale description.

A first application on the qualification of the model proficiency to predict the hydrological effects of a storm has been performed for Event 3, showing that our approach gives a number of information that were unknown before the analysis.

Summarizing the results obtained by the severity diagram comparison for the three cases analyzed, we can conclude that the meteorological model Meso-NH is able to reproduce many of the features of the events. In some cases, however, the model experiences i) rainfall underestimation in the plain region; ii) rainfall overestimation over the foothill due to orographic effects; iii) difficulties in reproducing some of the small-scale features of the events; iv) mislocation of the maxima leading to the impossibility to apply the simulations in hydrology. We have shown that severity diagrams can have interest for the hydrological modeling: the comparison between the hydrological impact of the storm by severity diagram and the frequency of the related flood may provide information upon the rainfall-runoff transfer function. Working with a probabilistic ensemble, and by means of hydrological simulations, a detailed analysis on the uncertainties propagation involved in the rainfall-runoff process could be performed.

Concerning the short-term perspectives of the work, we are currently carrying out severity analyses on the ensemble forecasts for a specific event occurred in 2008 in the same region. The first aim is to quantify, on an ensemble forecast produced with the AROME operational model (MétéoFrance) and composed by 17 members, the changes in severity induced by modifications of initial and boundary conditions of the simulation. At the light of the preliminary results, it seems that the severity diagrams could take a relevant role in the production of statistically consistent ensembles.

Acknowledgments.

J. Labalette is deeply acknowledged for his valuable contribution to this study. This work has been performed in the framework of ANR-MEDUP French project on the "Forecast and projection in climate scenario of Mediterranean intense events : Uncertainties and Propagation on environment". The climatological data was collected and provided by OHMCV, Hydro-Meteorological Observatory

Cévennes-Vivarais. We are indebted to Dr. O. Nuissier for providing us with the results of the Meso-NH simulations carried out at Météo-France.

REFERENCES

- Asquith, W. and J. Famiglietti, 2000: Precipitation areal-reduction factor estimation using an annual-maxima centered approach. *Journal of Hydrology*, **230**, 55–69.
- Bacchi, B. and R. Ranzi, 1996: On the derivation of the areal reduction factor of storms. *Atmospheric Research*, **42**, 123–135.
- Bendjoudi, H., P. Hubert, and D. Schertzer, 1997: Interpretation multifractale des courbes intensité-durée-fréquence des précipitations. *C. R. Acad. Sci. Paris*, **2**, 323–326.
- Berne, A., G. Delrieu, and B. Boudevillain, 2009: Variability of the spatial structure of intense mediterranean precipitation. *Advances in Water Resources*, 1031–1042.
- Bois, P., H. Mailloux, C. Obled, and F. De Saintignon, 1997: Atlas expérimental des risques de pluies intenses dans la région Cévennes-Vivarais. Tech. rep., LAMA, Pôle Grenoblois des Risques Naturels.
- Borga, M., C. Vezzani, and G. Dalla Fontana, 2005: Regional rainfall depth-duration-frequency equations for an alpine region. *Natural Hazards*, **36**, 221–235.
- Bousquet, O., C. Lin, and I. Zawadski, 2006: Analysis of scale dependence of quantitative precipitation forecast verification: A case study over the mackenzie river basin. *Quarterly Journal of the Royal Meteorological Society*, **132** (620), 2107–2125, doi:10.1256/qj.05.154.
- Burlando, P. and R. Rosso, 1996: Scaling and multiscaling models of depth-duration-frequency curves for storm precipitation. *Journal of Hydrology*, **187**, 45–64.
- Ceresetti, D., G. Molinié, and J. Creutin, 2010a: Scaling properties of heavy rainfall at short duration: a regional analysis. *Water Resources Research*, in press.
- Ceresetti, D., G. Molinié, and J.-D. Creutin, 2010b: Intensity - duration - frequency curves in a gev scale-invariant framework. *submitted to Water Resource Research*.
- De Michele, C., N. Kottegoda, and R. Rosso, 2001: The derivation of areal reduction factor of storm rainfall from its scaling properties. *Water Resources Research*, **37** (12), 3247–3252.
- Deidda, R., 2000: Rainfall downscaling in a space-time multifractal framework. *Water Resources Research*, **36**, 1779–1794.
- Delrieu, G., et al., 2005: The catastrophic flash-flood event of 8-9 september 2002 in the Gard region, France: a first case study for the Cévennes-Vivarais Mediterranean Hydro-meteorological Observatory. *Journal of Hydrology*, **6** (1), 34–52.
- Ducrocq, V., G. Aullo, and P. Santurette, 2003: Précipitations intenses et les inondations des 12 et 13 novembre 1999 sur le sud de la France. *La Météorologie*, **42**, 18–27.
- Ducrocq, V., O. Nuissier, D. Ricard, C. Lebeau-pin, and T. Thouvenin, 2008: A numerical study of three catastrophic precipitating events over southern France. ii: Mesoscale triggering and stationarity factors. *Quarterly Journal of the Royal Meteorological Society*, **134**, 131–145.
- Ducrocq, V., D. Ricard, J. Lafore, and F. Orain, 2002: Storm-scale numerical rainfall prediction for five precipitating events over france: On the importance of the initial humidity field. *Weather and Forecasting*, **17**, 1236–1256.
- Godart, A., S. Anquetin, and E. Leblois, 2009: Rainfall regimes associated with banded convection in the cévennes-vivarais area. *Meteorology and Atmospheric Physics*, **103**, 25–34.
- Guillot, P., 1993: The arguments of the gradex method: a logical support to assess extreme floods. *Proceedings of the Yokohama Symposium, Extreme Hydrological Events: Precipitation, Floods and Droughts*, Vol. IAH-SPubl. no. 213.
- Guillot, P. and D. Duband, 1967: la méthode du GRADEX pour le calcul de la probabilité des crues à partir des pluies. *Journées de la SHF*, Société Hydrologique de France, Paris, Société Hydrologique de France, Vol. Rapport 7, question 1.
- Haberlandt, U., 2007: Geostatistical interpolation of hourly precipitation from rain gauges and radar for a large-scale extreme rainfall event. *Journal of Hydrology*, **332**, 144–157.
- Kotz, S. and S. Nadarajah, 2000: *Extreme value distributions. Theory and applications*. Imperial College Press.
- Koutsoyiannis, D., D. Kozonis, and A. Manetas, 1998: A mathematical framework for studying rainfall intensity - duration - frequency relationships. *Journal of Hydrology*, **206**, 118–135.

- Lafore, J. P., et al., 1998: The Meso-NH atmospheric simulation system. part i: Adiabatic formulation and control simulations. *Annales Geophysicae*, **16**, 90–109.
- Lebeaupin, C., V. Ducrocq, and H. Giordani, 2006: Sensitivity of torrential rain events to the sea surface temperature based on high-resolution numerical forecasts. *Journal of Geophysical Research*.
- Lebel, T. and J. P. Laborde, 1988: A geostatistical approach for areal rainfall statistics assessment. *Stochastic Hydrology and Hydraulics*, **2** (4), 245–261, doi:10.1007/BF01544039.
- Mason, I., 1989: Dependence of the critical success index on sample climate and threshold probability. *Aust. Meteorol. Mag.*, **37**, 75–81.
- Menabde, M., A. Seed, and G. Pegram, 1999: A simple scaling model for extreme rainfall. *Water Resources Research*, **35** (1), 335–339.
- Molinié, G., D. Ceresetti, S. Anquetin, B. Boudevillain, and J. D. Creutin, 2010: Rainfall regimes in a Mediterranean mountainous region. *submitted to Quarterly Journal of the Royal Meteorological Society*.
- NERC, 1975: Flood studies report - meteorological studies. Tech. rep., UK Natural Environmental Research Council.
- Nuissier, O., V. Ducrocq, D. Ricard, C. Lebeaupin, and S. Anquetin, 2008: A numerical study of three catastrophic precipitating events over southern france. i: Numerical framework and synoptic ingredients. *Quarterly Journal of the Royal Meteorological Society*, **134**, 111–130.
- Omolayo, A., 1993: On the transportation of areal reduction factors for rainfall frequency estimation. *Journal of Hydrology*, **145**, 191–205.
- Prudhomme, C. and D. Reed, 1999: Mapping extreme rainfall in a mountainous region using geostatistical techniques: a case study in scotland. *International Journal of Climatology*, **19**, 1337–1356.
- Ramos, M. H., J. D. Creutin, and E. Leblois, 2005: Visualization of storm severity. *Journal of Hydrology*, **315**, 295–307.
- Rodriguez-Iturbe, I. and J. M. Mejía, 1974: On the transformation of point rainfall to areal rainfall. *Water Resources Research*, **10** (4), 729–735.
- Ruin, I., J.-D. Creutin, S. Anquetin, and C. Lutoff, 2008: Human exposure to flash-floods relation between flood parameters and human vulnerability during a storm of september 2002 in southern france. *Journal of Hydrology*, 199–213.
- Sénési, S., P. Bougeault, J. Chèze, P. Cosentino, and R. Thepenier, 1996: The Vaison-la-Romaine flash flood: Mesoscale analysis and predictability issues. *Weather and Forecasting.*, **11** (4), 417–442.
- Sivapalan, M. and G. Blöschl, 1998: Transformation of point rainfall to areal rainfall: Intensity - duration - frequency curves. *Journal of Hydrology*, **204**, 150–167.
- Taylor, G., 1938: The spectrum of turbulence. *Proceedings Royal Society London A*, **164** (919).
- Venugopal, V., S. Basu, and E. Foufoula-Georgiou, 2005: A new metric for comparing precipitation patterns with an application to ensemble forecasts. *Journal of Geophysical Research*, **110** (D8), D08 111.1–D08 111.11.
- Venugopal, V., E. Foufoula-Georgiou, and V. Sapozhnikov, 1999: Evidence of dynamic scaling in space-time rainfall. *Journal of Geophysical research*, **104** (D24), 31 599–31 610.
- Yan, X., V. Ducrocq, P. Poli, M. Hakam, G. Jaubert, and A. Walpersdorf, 2009: Impact of gps zenith delay assimilation on convective-scale prediction of mediterranean heavy rainfall. *Journal of Geophysical Research*, **114**.
- Yates, E., J. Creutin, S. Anquetin, and J. Rivoirard, 2007: A scale dependant quality index of areal rainfall prediction. *Journal of Hydrometeorology*, **8**, 160–170.
- Yates, F., 1984: Tests of significance for 2x2 contingency tables. *Journal of the Royal Statistical Society - Series A*, **147** (3), 426–463.
- Zepeda-Arce, J., E. Foufoula-Georgiou, and K. Droegemeier, 2000: Space-time rainfall organization and its role in validating quantitative precipitation forecasts. *Journal of Geophysical Research*, **105**, 10 129–10 146.

TABLE 1. Scale-invariant Areal Reduction Factor model parameters ((De Michele et al. 2001) model) for region 1 (Flat land) and region 2 (Mountainous region).

Region	ω	a	b	v
1 : Flat area	0.00632	0.55	0.34	0.84
2: Mountainous region	0.00234	0.52	0.14	0.64

TABLE 2. Contingency tables and definition of the thresholded statistics; thresholded statistics for the three events. POD: Probability of Detection; FAR: False Alarm Ratio; ACC: Forecast Accuracy.

	Forecast < Threshold	Forecast \geq Threshold
Ref. < Threshold	a	b
Ref. \geq Threshold	c	d

$$\mathbf{POD} = \frac{\mathbf{d}}{\mathbf{c+d}} ; \mathbf{FAR} = \frac{\mathbf{b}}{\mathbf{b+d}} ; \mathbf{ACC} = \frac{\mathbf{a+d}}{\mathbf{a+b+c+d}} ;$$

SCORE Quantile	Event 1			Event 2			Event 3		
	70%	80%	90%	70%	80%	90%	70%	80%	90%
POD	0.79	0.78	0.84	0.58	0.54	0.21	0.93	0.80	0.29
FAR	0.32	0.43	0.62	0.54	0.60	0.64	0.01	0.02	0.14
ACC	0.82	0.84	0.85	0.67	0.74	0.88	0.92	0.79	0.31

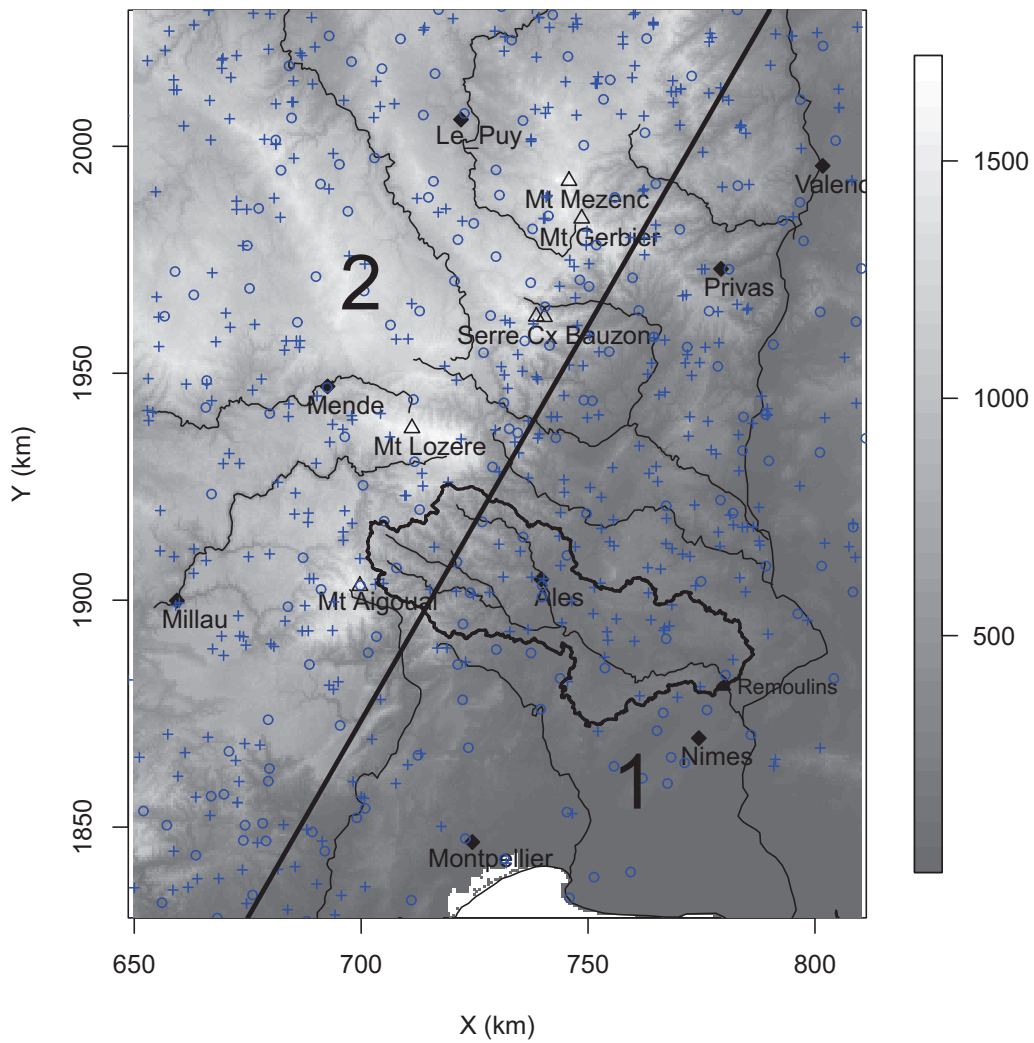


FIG. 1. Main features of the region of analysis: elevation above sea level, main rivers (solid line), main mountain peaks (triangles), main cities (diamonds). Ground measurement network: daily network (circles) and hourly network (crosses). The solid thick line indicates the boundaries of Gard basin, the outlet of Remoulins is indicated. A line separates mountainous (1) and flat (2) sub-regions. The rectangle identifies the area where the analysis has been carried out.

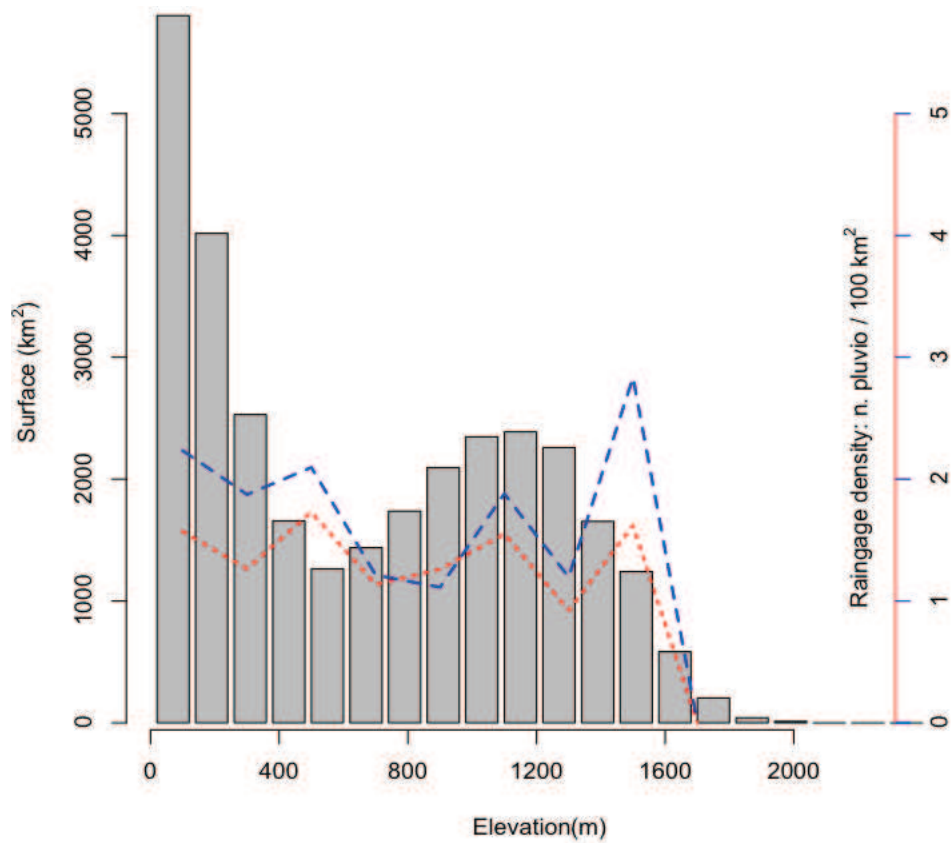


FIG. 2. Rainage distribution in the Cévennes-Vivarais region. The histogram represents the surface associated to each elevation band (left axis). The rainage density is shown as function of the elevation, and its value can be read in the right axis. Daily rainage network: long-dashed line; Hourly network: dashed-line.

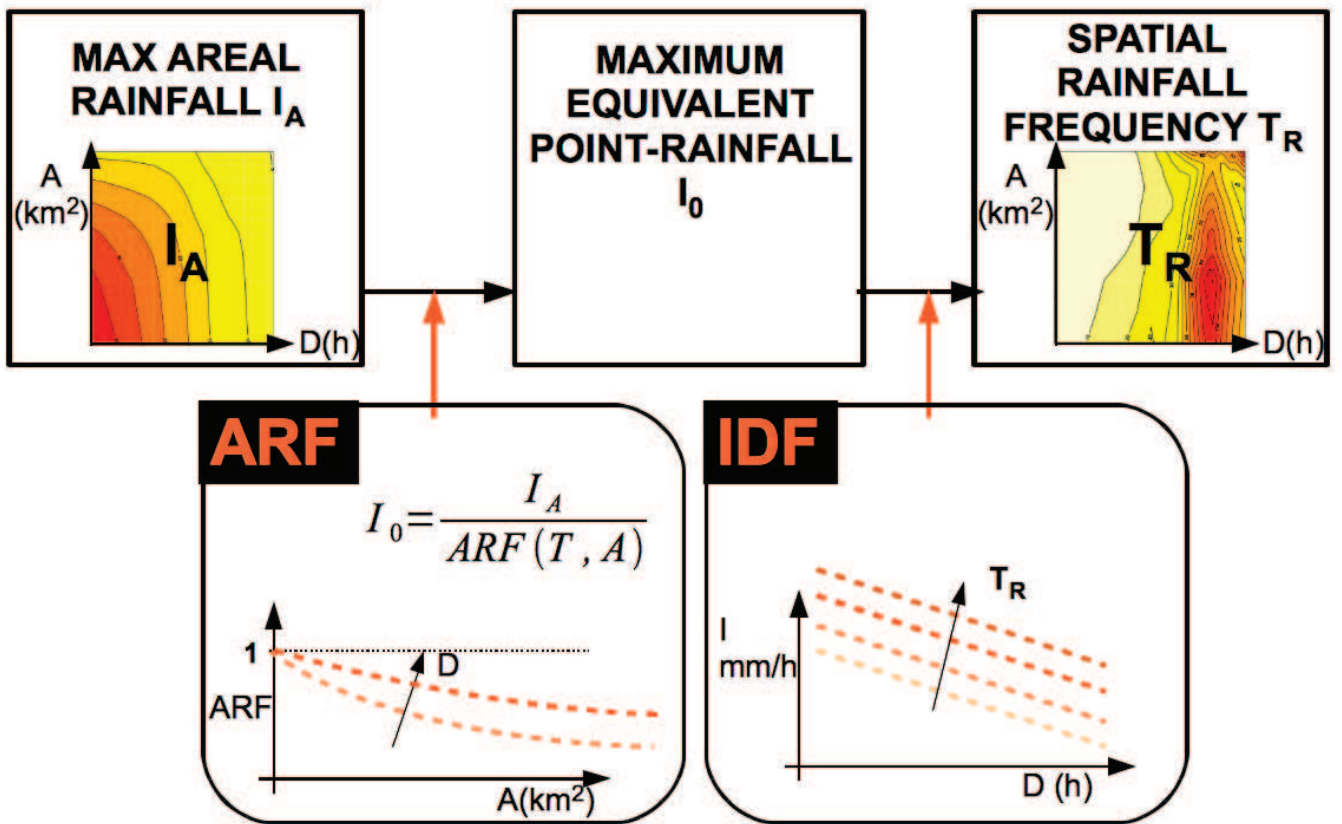


FIG. 3. Diagram of the steps required to compute the severity of spatial rainfall.

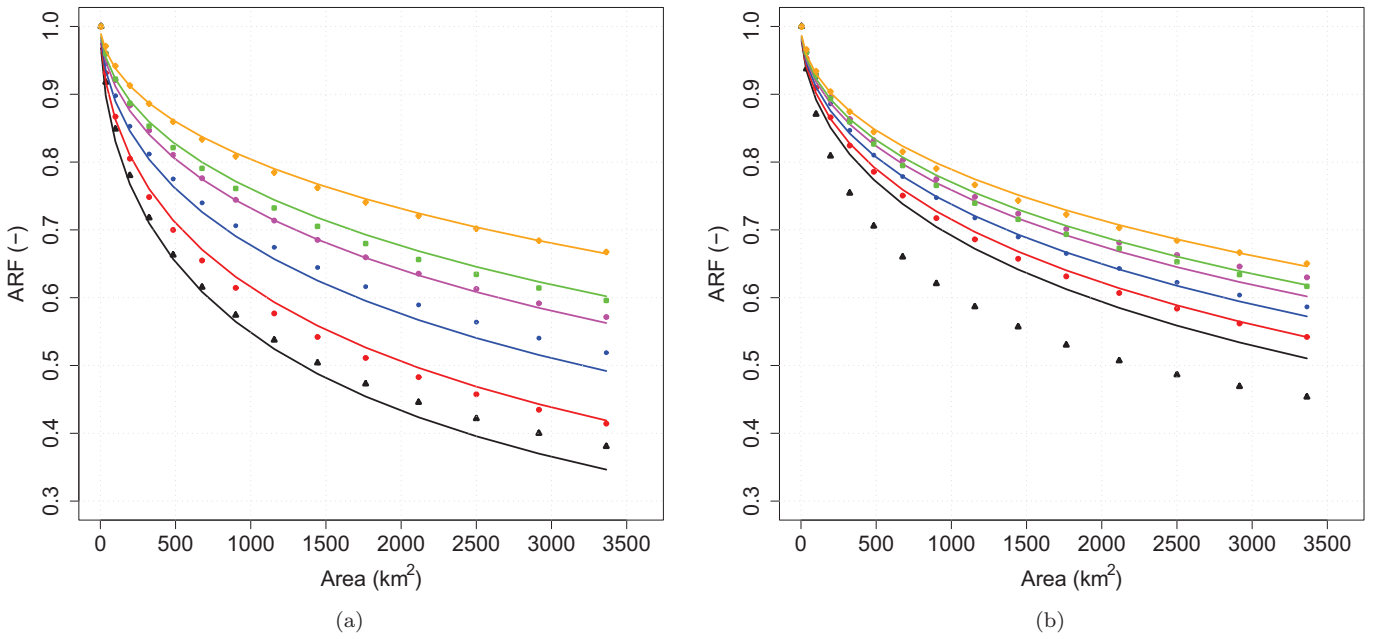


FIG. 4. Modeled ARF curves (lines) and empirical ARFs (dots) for different durations as a function of the area. a: Region 1- flat land. b: Region 2 - Mountainous region. In both diagrams, the duration of 1 hour is not used for ARF model fitting, but the empirical as well as modeled curve are reported. Legend: Diamonds: 24 h, Squares: 12h, Circles: 8h, Small circles: 4h, Circles: 2h, Triangles: 1h. Durations lower than 2 h and surfaces lower than 50 km^2 have not been taken into account for the fitting due to undersampling volume.

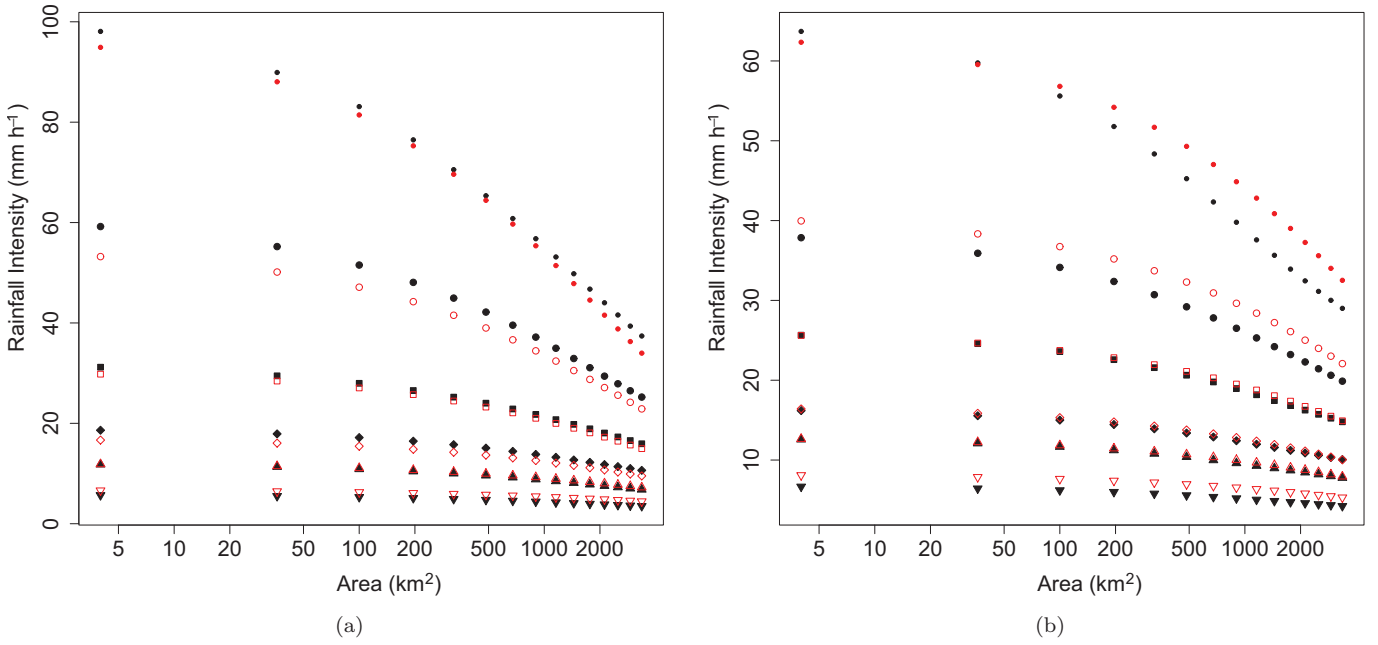


FIG. 5. Intensity-Duration-Area-Frequency model obtained gathering IDF and ARF models: comparison between observed events and modeled events following De Michele et al. (2001); from the top to the bottom the plotted accumulation duration is 1, 2, 4, 8, 24 h. a: Region 1- flat land. b: Region 2 - Mountainous region. Legend: Solid symbols=empirical; Empty symbols=model. Small circles = 1h, large circles =2 h, squares=4h, diamonds=8h, upward triangles=12 h, downward triangles=24h.

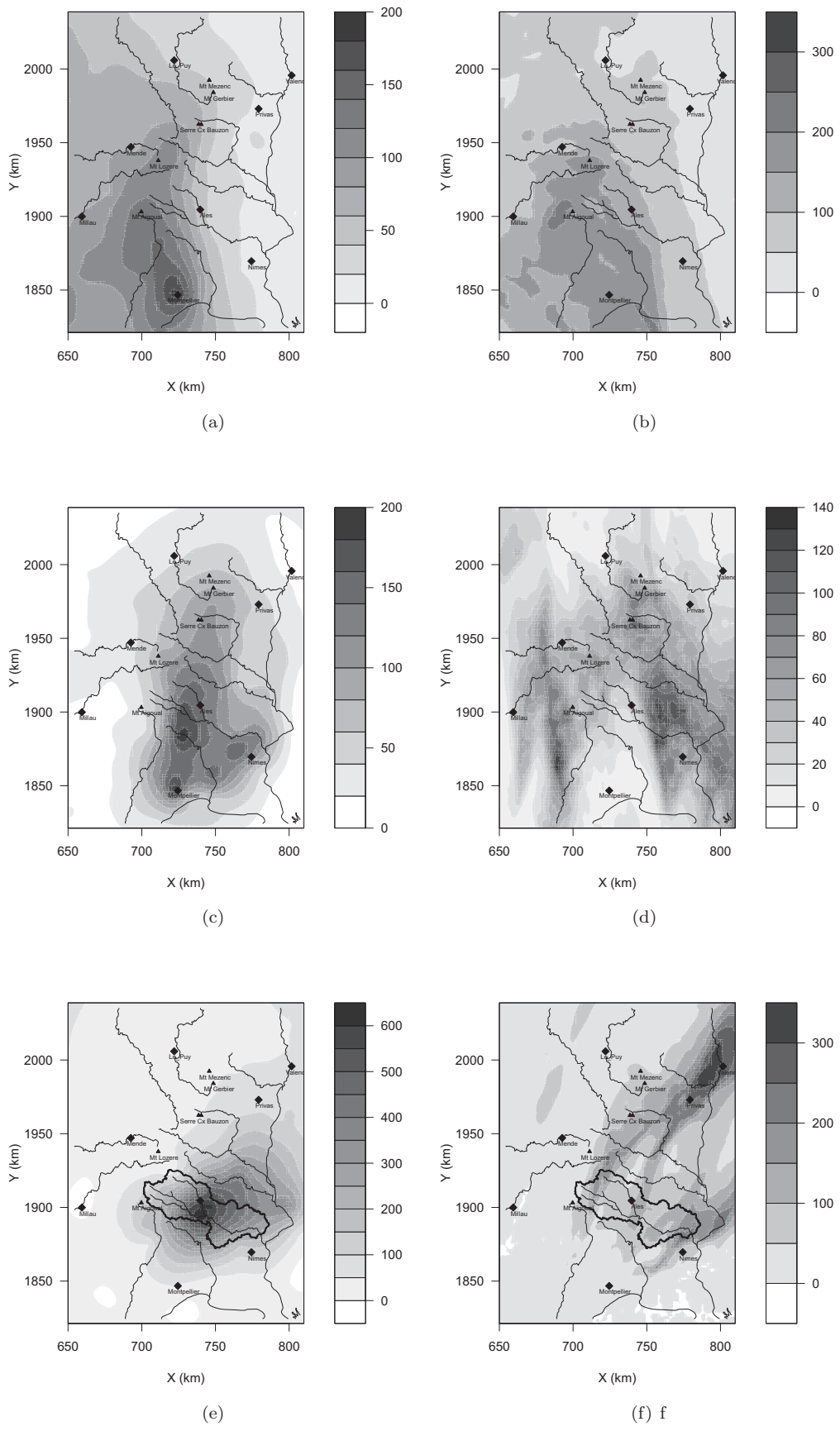


FIG. 6. Comparison between observed and simulated rainfall depth for the three considered events. 2003.12.03 0000UTC - 2003.12.04 0000UTC, a: Ground Rainfall, b: Meso-NH Simulation. 2005.09.05 1200UTC - 2005.09.05 0600UTC, c: Ground rainfall, d: Meso-NH Simulation. 2002.09.08 1200 UTC - 2002.09.09 1200UTC, e: Ground rainfall; f: Meso-NH Simulation.

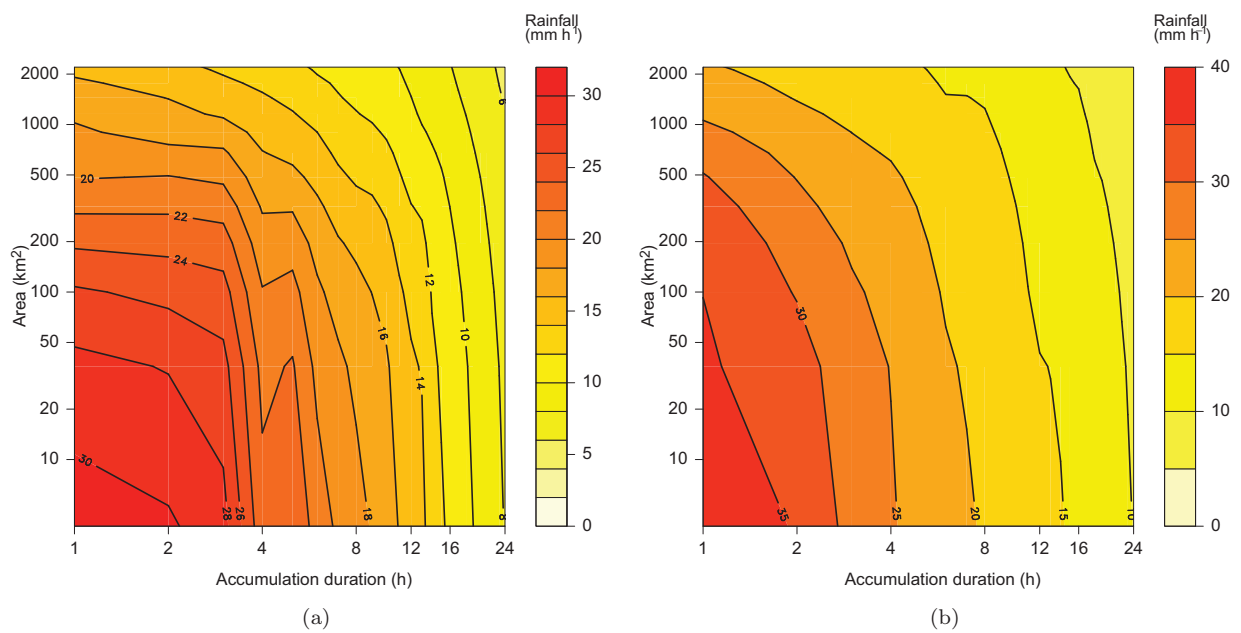


FIG. 7. Maximum intensity diagram for Event 1, from 03.12.2003 00h UTC to 04.12.2003 00 h UTC. a) Observed Hourly Rainfall. b) Simulated rain fields.

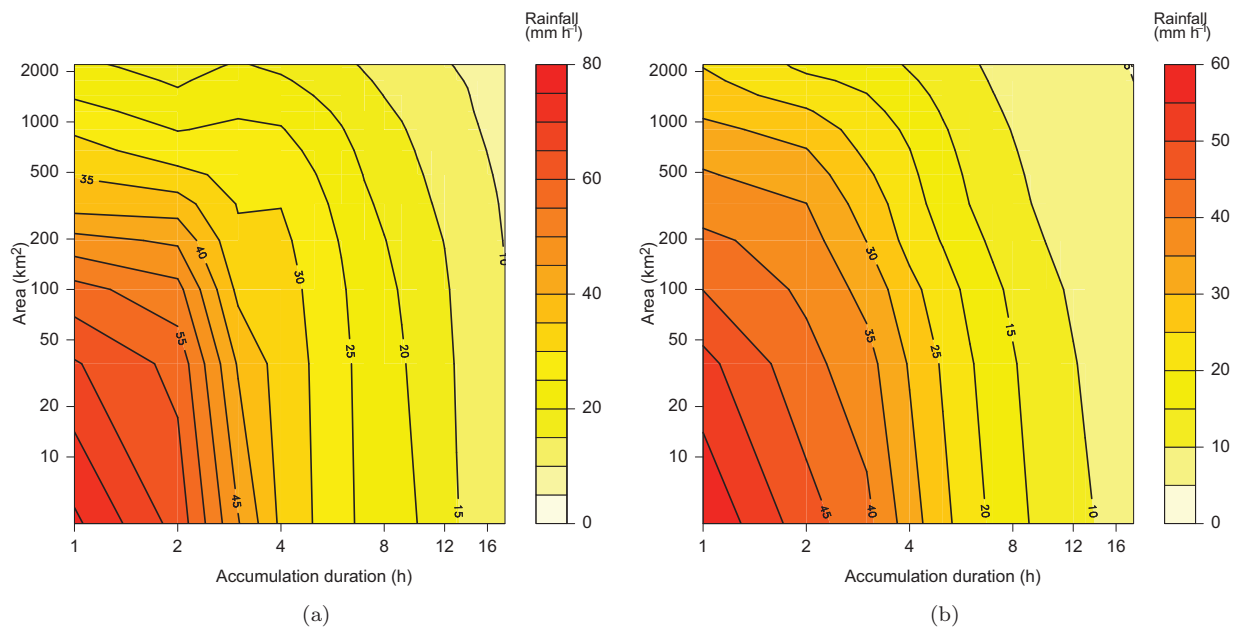


FIG. 8. Maximum intensity diagram for Event 2, from 06.09.2005 00h UTC to 06.09.2005 18 h UTC. a) Observed Hourly Rainfall. b) Simulated rain fields.

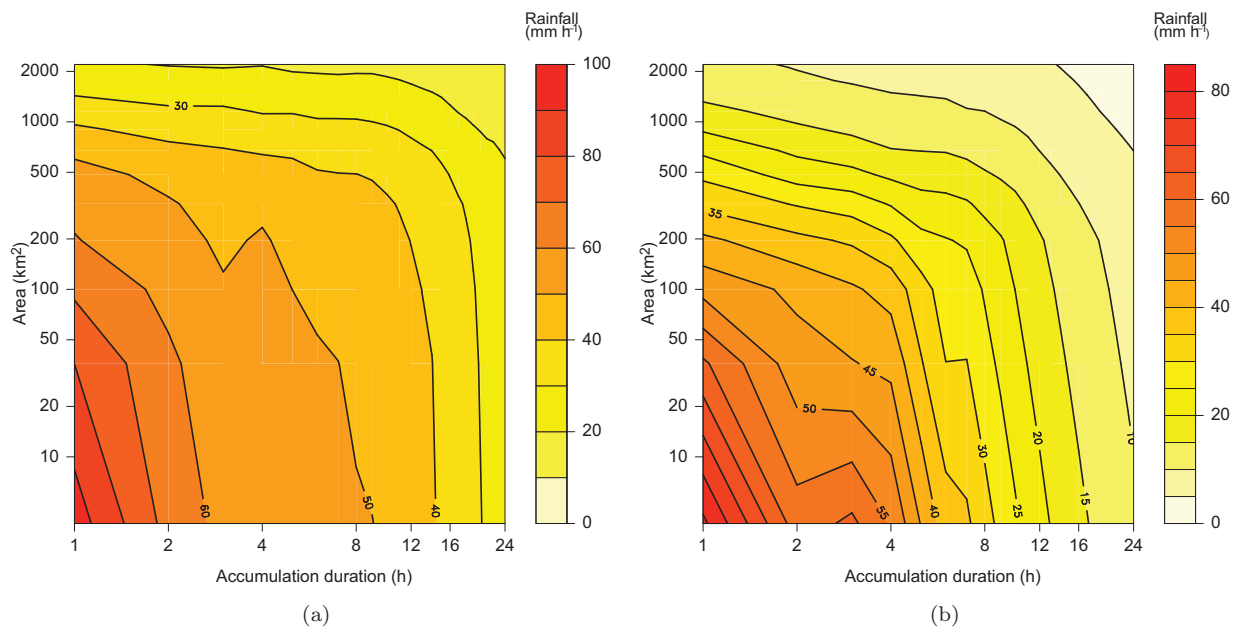


FIG. 9. Maximum intensity diagram over the Gard basin for Event 3, from 08.09.2002 12h UTC to 09.09.2002 12 h UTC. a) Hourly Rainfall Network. a) Observed Hourly Rainfall. b) Simulated rain fields.

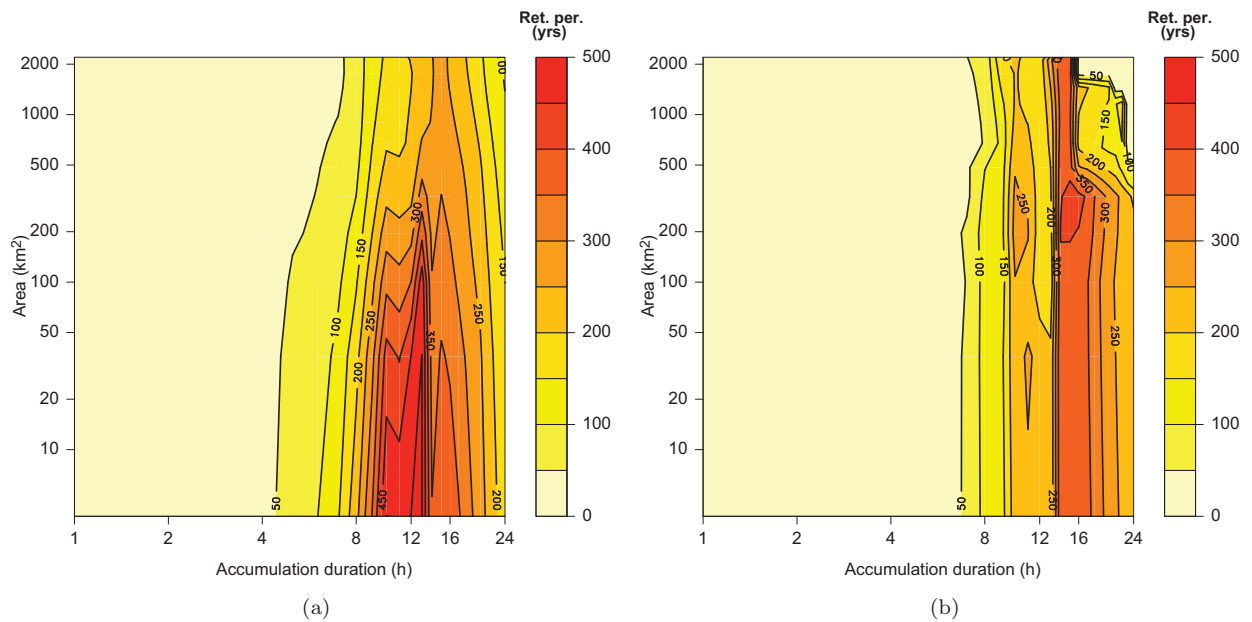


FIG. 10. Severity diagram for Event 1, from 03.12.2003 00h UTC to 04.12.2003 00 h UTC. a) Observed Hourly Rainfall. b) Simulated rain fields.

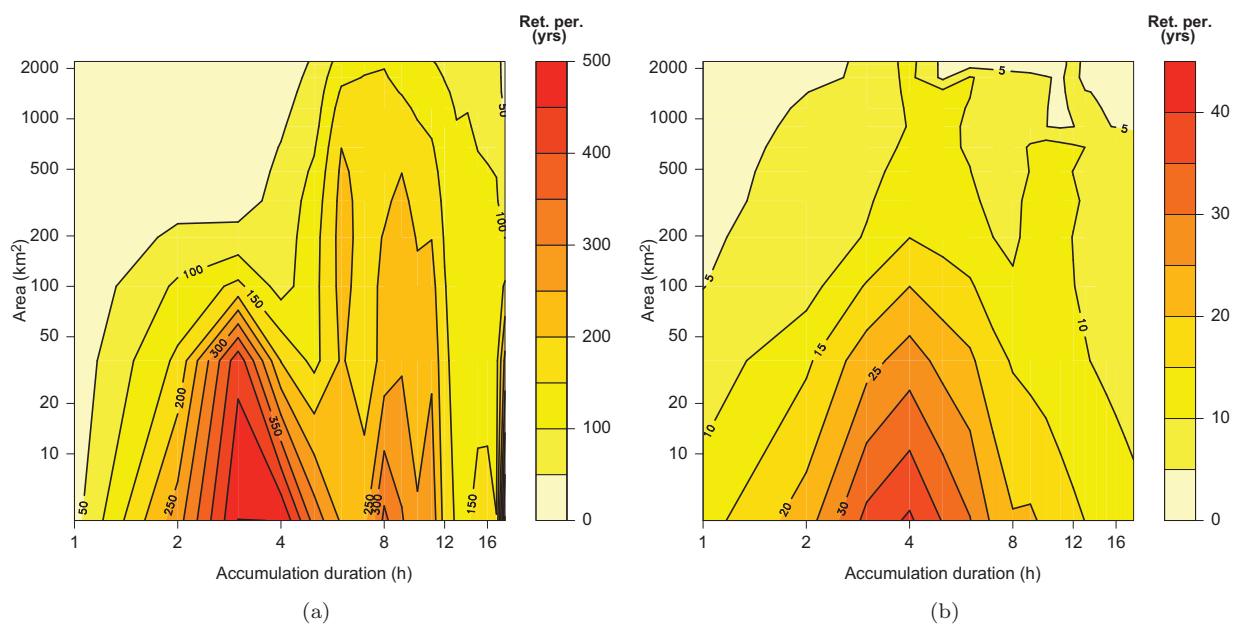


FIG. 11. Severity diagram for Event 2, from 06.09.2005 00h UTC to 06.09.2005 18 h UTC. a) Observed Hourly Rainfall. b) Simulated rain fields.

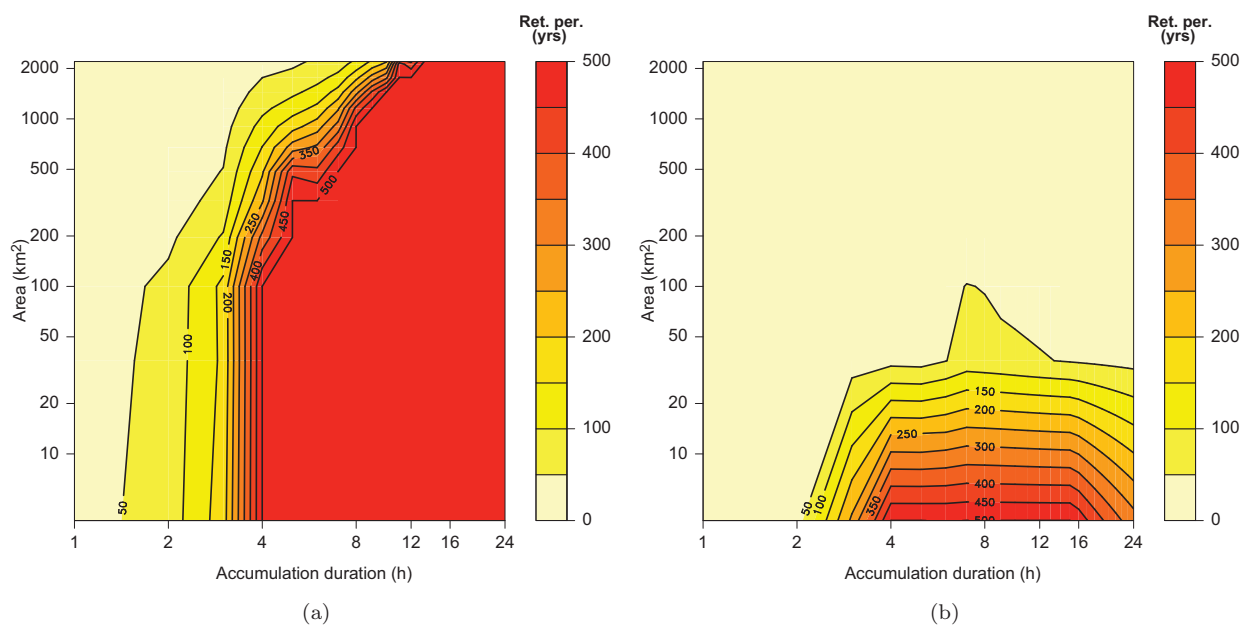


FIG. 12. Severity diagram within the Gard basin for Event 3, from 08.09.2002 12h UTC to 09.09.2002 12 h UTC. a) Observed Hourly Rainfall. b) Simulated rain fields. Due to the large uncertainties in the return period estimation, the absolute severity value for return periods higher than 500 years is not reported.

10.5 Severity diagrams and ensemble simulations

The article in Section 10.4 reports the use of severity diagrams to evaluate the capability of meso-scale models to reproduce the actual space-time structure of extreme events. It is clear that this tool is able to highlight the model drawbacks, but it can't help in improve the meteorological modeling.

A secondary application concerns the ensemble simulations. Due to extreme variability of the results even when the input is the same (butterfly effect, *Lorenz (1963)*), the probabilistic approaches in meteorological modeling are preferable to the deterministic ones.

Ensemble simulation is a group of simulations that differ among each other by one or more properties. Each simulation is initialized with a series of initial and boundary conditions deriving from direct measures. They are affected to sampling problems (weak resolution and measure uncertainties). The ensemble is supposed to represent a database of situations likely to happen, obtained perturbing i) the boundary conditions; ii) the initial conditions.

The event of 2008, November the 1st has been simulated by Météo-France using the operational meteorological model AROME. The event lead to a total rainfall depth of over 350 mm, mainly located over the foothill of Cévennes. Figure 10.5-a reports the total rainfall depth fallen in 24 h starting from 2008, November the 1st at 12h UTC.

Two different ensemble are created in order to separately determine the influence of boundary conditions and initial conditions.

The severity diagrams are computed on the hourly rainfall extracted from the AROME simulation, beginning the 01 Nov 2008 at 12 UTC and lasting 24 hours. The severity diagrams derived from the model output are compared with the severity diagram relative to the ground observations, that is shown in the next section.

10.5.1 Severity Diagram of observed fields

The reference diagram is the severity diagram computed based on the observations coming from the recording rain gauge network of Météo-France. Figure 10.5-b reports the severity diagram computed on this event. The maximum recorded severity is 600 years in a particular rain gauge, reaching about 200 years for surfaces of 100 km^2 . The severity peak is reached for 16-18 hours.

10.5.2 AROME ensemble 1: Variability of Boundary Conditions

In the first ensemble (PEARP-AROME), the model AROME is initialized with the boundary conditions derived from the large-scale Arpège PEARP model. Apart from the unperturbed case (Figure 10.6-a), 10 simulations are generated, and are reported in Figure 10.6-b to 10.6-k.

Comparing each member of the ensemble with the observed severity diagram in Figure 10.5-b, we find that only the members in Figure 10.6-a,b,c,i,k well represent the critical scales of the event. In addition, the magnitude of the event 10.6-a,b,c,d is respected. The other members show different behavior, with extremely underestimated magnitude like in Figure 10.6-f,j. The member whose severity

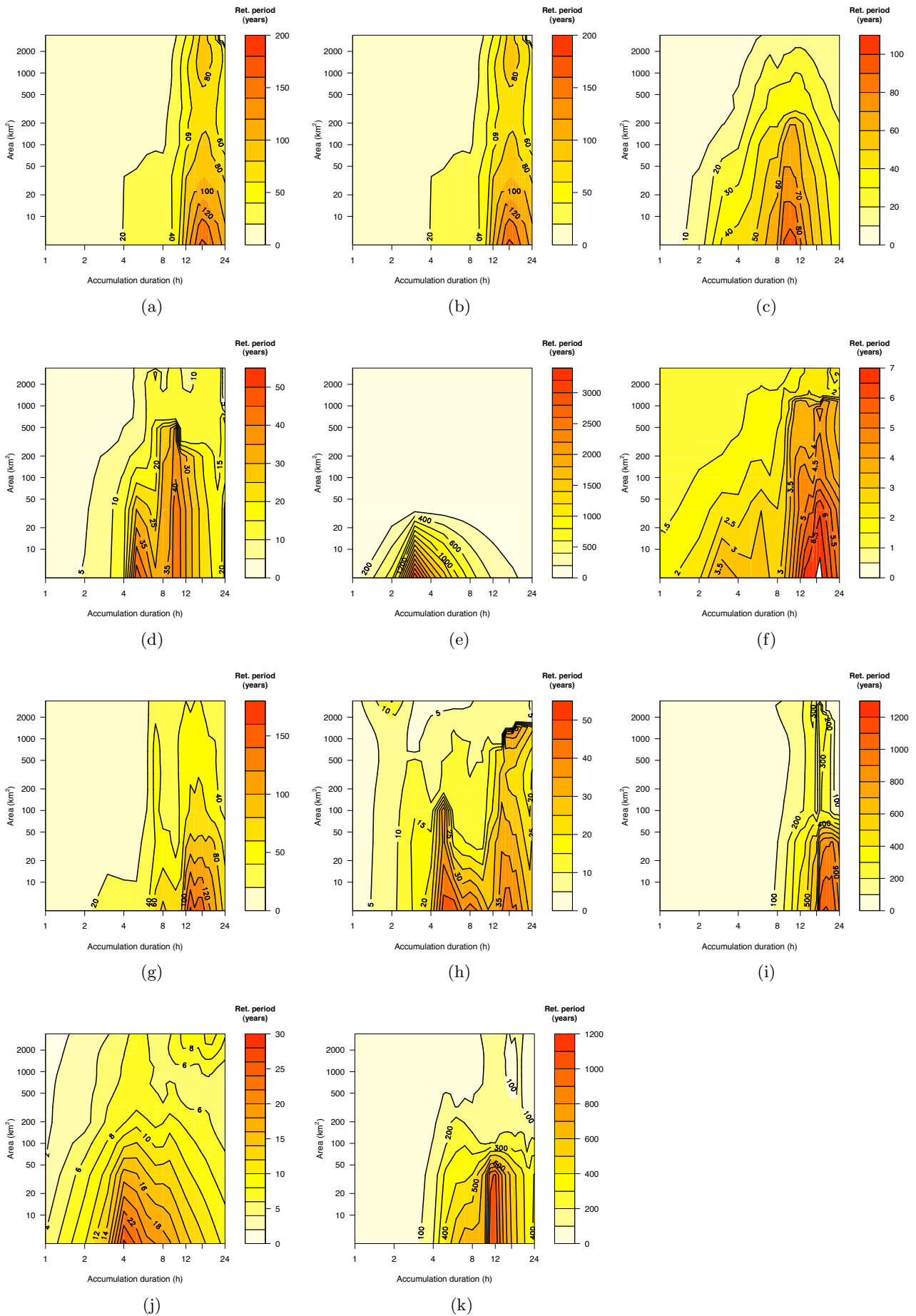


Figure 10.6: Severity diagrams of each member of the ensemble PEARP-AROME, unperturbed simulation (0) and perturbed simulations (a:j).

euros damages and 2 fatalities. The article show that, despite the good agreement between model and observations in terms of maximum rainfall intensity, the critical scales of the events are not always caught by the model. In addition, the mislocation of the maximum rainfall depth, such as in the 2002 case, bring into question the reliability of mesoscale simulations at the scale of hydrological basins.

The second application concerns the use of severity diagrams to interpret the output of ensemble simulations. With this tool it is possible to detect if the noise introduced in the initial/boundary conditions is appropriated to represent the small scale variability of the physical variables. The tool allows to critically evaluate the result of each simulation (critical space-time scales, event magnitude, mislocated maxima) in a qualitative way.

In perspective, we think that this use of severity diagrams can contribute to a better selection of the ensemble members providing a more realistic probabilistic representation of the storms.

Part IV

Conclusion and Perspectives

1 1 Conclusion

Conclusion

L'objectif de la thèse est d'étudier l'estimation des précipitations extrêmes dans une large gamme d'échelles spatio-temporelles, en combinant invariance d'échelle et analyse des valeurs extrêmes. Avant de pouvoir rechercher des propriétés d'invariance d'échelle dans des séries réelles de pluie, il est nécessaire d'évaluer la qualité de la mesure des précipitations ponctuelles ou intégrées dans l'espace. Nous évaluons d'abord les incertitudes liées à la résolution temporelle de la série des précipitation. Il est vérifié que, si le rapport entre fréquence d'échantillonnage et échelle d'analyse est proche de 1, les pluies extrêmes peuvent être sous-estimées jusqu'au 15%. Cette valeur s'ajoute à l'erreur maximale de sous-estimation due à la structure mécanique du pluviomètre qui est de 5%. Pour ce qui concerne la pluie intégrée dans l'espace, une ultérieure sous-estimation est causée par la densité insuffisante des réseaux pluviométriques. Cette estimation peut être quantifiée entre 15 et 40% pour les résolutions temporelles de 1 jour à 1 heure. A partir de ces données, nous pouvons conclure que si l'on examine les séries de précipitations ponctuelles, l'erreur maximale est limitée à 5%, à condition que l'échelle d'analyse soit plus grande que la résolution de mesure. Pour ce qui concerne la pluie intégrée dans l'espace, une série de simulations stochastiques appliquées à un réseau pluviométrique aléatoire montrent que, dans des réseaux pluviométriques peu denses, la sous-estimation des extrêmes est systématique et peut atteindre 50% en fonction de la taille caractéristique des phénomènes pluvieux et de la densité du réseau de mesure. Cela limite considérablement la capacité d'extraire des données statistiques fiables en particulier pour des durées inférieures à 4 heures. Ceci est dû au faible rapport entre la taille caractéristique des phénomènes pluvieux et la densité moyenne du réseau pluviométrique.

La troisième partie de la thèse présente les résultats obtenus à l'aide de la base de données OHM-CV. Le premier article analyse les séries pluviométriques infra-journalières de la région, et vise à déterminer le comportement des queues de distributions. L'étude a mis en évidence qu'un bon nombre de pluviomètres de la région montrent des queues de distribution hyperboliques. Un résultat surprenant est que le comportement des queues n'est pas influencé par la résolution temporelle. Les distributions de pluie enregistrées dans la plaine du Rhône montrent un comportement hyperbolique évident. D'autre part, les distributions de pluie relatives à la région montagneuse montrent plutôt des

queues exponentielles, signe d'une moindre variabilité des précipitations (bien que les cumuls quotidiens soient plus forts). Cette analyse est le point de départ pour la conception d'un modèle des intensités de pluie extrêmes qui puisse représenter correctement le comportement des queues faibles ainsi que des queues fortes, indépendamment de la durée. La deuxième étude a donc consisté à concevoir un modèle Intensité-Durée-Fréquence (IDF) pour précipitations ponctuelles qui répond aux deux conditions mentionnées dans les études précédentes: invariance d'échelle et distribution GEV des maxima, pour modéliser à la fois le comportement des pluie extrêmes en montagne et en plaine. L'invariance d'échelle des maxima de pluie a été vérifié pour le durées allant de 4 à 100 h.

La troisième étude utilise ce modèle IDF couplé avec un modèle semi-empirique qui décrit la décroissance de l'intensité de pluie en fonction de la surface (ARF). Ce couplage permet d'obtenir un modèle Intensité-Durée-Surface-Fréquence (IDAF) qui donne la possibilité d'estimer la fréquence des pluies intégrées sur une surface donnée. Ce modèle est employé pour obtenir un diagramme (diagramme de sévérité) qui montre la période de retour de l'événement dans une large gamme d'échelles spatio-temporelles. En représentant la période maximale de retour sur le diagramme de severite, les hauteurs de pluie sont normalisées par rapport à la fréquence d'occurrence, de sorte de pouvoir i) qualifier les échelles les plus critiques pour un événement; ii) comparer des événements co-localisés entre eux; iii) après régionalisation, comparer des événements qui intéressent différentes régions. L'innovation de cette étude est l'utilisation des diagrammes de sévérité pour évaluer la qualité des champs de pluie produits par simulation météorologique, effectuée grâce au modèle Mésos-NH. Les champs de pluie simulés reproduisent généralement assez bien l'événement dans une large gamme d'échelles (1-2000 km² et 1-24 heures). La période de retour ayant une dynamique plus marquée que les intensités de pluie, le diagramme de sévérité s'est révélé un diagnostic intéressant pour la compréhension et la comparaison de la structure spatio-temporelle des événements. Dans un'étude en cours, les diagrammes de sévérité sont appliqués à la simulation d'ensemble.

Conclusion

The purpose of this thesis is to investigate the heavy rainfall events in a “free of scale” framework, combining scale-invariance and extreme value analysis. Before to look at the scale-invariance properties of rainfall series, we assess the quality of rainfall measurement of point and spatial-rainfall. In the first part, we quantify the uncertainties on the measurement of rainfall intensities using rain gauges. The first analysis quantified the uncertainty related to the temporal resolution of the series. It comes out that when the ratio between the sampling frequency and the scale of analysis tends to 1, this may cause up to 15% of underestimation of extreme rainfall. An additional 5% is the maximum underestimation error of extremes related to the mechanical structure of rain gauge. Finally, if one aims to estimate the spatial rainfall, a further underestimation due to insufficient rainfall network density can be quantified in 15-40% for time resolutions ranging between 1 day and 1 hour. From these data, we conclude that if we examine point rainfall series at a scale larger than the resolution of the device (e.g. of one magnitude order), the error of measurement of extremes is limited to 5%.

Concerning the evaluation of areal integrated rainfall intensities, stochastic rainfall simulations applied on a random rain gauge network show that sparse rain gauge networks lead to systematic

under-sampling of maximum spatial rainfall intensities. The total underestimation may reach 50% depending on the characteristic size of rainfall and of the density of rain gauges. For this reason, the possibility to extract reliable statistics by interpolation of rainfall data is significantly limited for durations lower than 4 hours, for which the characteristic size of the rainfall phenomena is lower than the average density of the rain gauge network.

The third part shows the results of the analysis conducted using the database of the OHM-CV survey. The first article analyzes the hourly time series of the region, and aims at determining the behavior of the distribution tails. The study highlighted that many of the rainfall intensity distributions in the region show hyperbolic tails. A surprising result is that the behavior of tails does not change with the time resolution. The rainfall distributions related to rain gauges located in the flat land show evidences of hyperbolic behaviour at various durations. On the other hand, the mountainous intensity distributions show exponential tails, indicating a smaller rainfall variability (though the daily amount is larger in average). This analysis is preliminary to the application of a scale-invariant model of rainfall maxima that accommodates the extreme rainfall behavior in plain as well as in mountainous areas. The second study concerns the design of a model of Intensity-Duration-Frequency relations for point rainfall that meets the two conditions specified in the previous studies: scale-invariance of rainfall intensities in the range 4-100 h, and GEV distributed maxima, modeling in a comprehensive approach both the hyperbolic tails in the plain of the Rhône valley and the exponential tails observed in mountainous areas. The third study uses the model of Intensity-Duration-Frequency curves above described, together with a semi-empirical model that describes the decay of the rainfall intensity with the area (ARF). The coupling of the two models provide an Intensity-Duration-Area-Frequency (IDAF) model that can be used to estimate the frequency of any spatial rainfall event in the region. We have applied the IDAF model to the implementation of a diagram (severity diagram) that shows the magnitude of an event at a large range of space-time scales. Severity diagrams, transforming the (scale-dependent) rainfall depth into (scale-independent) return period, allow comparisons between different scales of a single event, or events occurred over different regions. The innovation of this thesis is the use of severity diagrams to evaluate the performance of the rainfall intensity fields simulated by a meso-scale model, MesoNH. The maximum intensities of the simulated rainfall satisfactorily match the observed ones in a large range of scales (1-2000 km^2 and 1-24 hours). Since return periods have higher dynamics than rainfall intensities, severity diagrams indicate more accurately the differences between simulated and observed fields, highlighting the space and time scales mainly hit by the storm. In a last application, we apply the severity diagrams to ensemble simulations.

12 Perspectives

Perspectives

Les études relatives aux erreurs dans l'estimation des précipitations ponctuelles et spatiales par pluviomètres menées dans cette thèse donnent un aperçu du problème de la mesure de la pluie. Ces résultats ont permis d'établir la meilleure configuration pour un réseau expérimentale de pluviomètres. Ces résultats préliminaires sont à la base de la mise en œuvre d'un nouveau super-site de mesure au Pradel, Ardèche (HyMeX project).

Une deuxième application est liée à la mesure pluviométrique par radar. L'analyse des champs de pluie relatifs à l'événement de Septembre 2002 a permis de vérifier que les champs issus de la mesure radar sont invariants d'échelle dans l'espace. Cette propriété peut être exploitée afin de redéfinir des relations entre la réflectivité radar et l'intensité des précipitations. La proposition est d'appliquer la désagrégation par cascade. La technique devrait permettre d'obtenir des champs de réflectivité radar à fine-échelle (de l'ordre de celle des pluviomètres) caractérisés par une reproduction de la variabilité du phénomène proche de la réalité. La conversion en intensité de précipitation qui s'en suit donne des lames d'eau moyennes supérieures d'un 10-20% (en fonction de la variabilité du phénomène) par rapport à la transformation effectuée à la grande échelle. La désagrégation donnerait un résultat additionnel: il permettrait de se débarrasser des erreurs d'interpolation intrinsèques à la transformation des données polaires en coordonnées cartésiennes. En effet, chaque scrutation du radar à un angle donné (image de réflectivité courante) contient des données de réflectivités correspondant à des intégrations spatiales effectuées sur des dizaines d'échelles différentes.

Les distributions des intensités extrêmes de pluie ont des queues de types soit hyperboliques, soit exponentielles qui ont la particularité de se conserver sur une gamme étendue de durées d'accumulation. Grâce à la modélisation des relations IDF basée sur les fonctions de densités GEV et leur invariance d'échelle, et aux facteurs de réduction empiriques ARF, il est maintenant possible d'attribuer une fréquence d'occurrence à une intensité de pluie intégrée sur une surface donnée.

De plus, l'implémentation des diagrammes de sévérité permet de mettre en évidence les échelles critiques d'un événement dans l'espace espace-temps. Une base de données élaborées contenant les diagrammes de sévérité des événements intenses survenus dans la région dans ces dernières décennies est en projet. Un étude sur l'invariance d'échelle des événements pluvieux va être réalisé dans le contexte du projet AMMA-CATCH en analysant les pluies ponctuelles. Dans cette région, les aspects liés aux

pluies intégrées sur une surface ne peuvent pas être abordés en suivant la même méthodologie à cause de la faible densité du réseau pluviométrique. Une solution à ce problème pourrait être l'utilisation de la désagrégation statistique des champs de pluie satellitaires à faible résolution.

Une autre application des diagrammes de sévérité révélée lors des discussions au sein du projet MEDUP, est la mise en place d'un diagnostic synthétique des simulations d'ensembles. Ces simulations génèrent des dizaines d'échéances pour plusieurs champs (variables) et pour chaque membre (plusieurs dizaines parfois). L'analyse de ces champs, qu'on imagine laborieuse, doit permettre de sélectionner les membres dont les produits tel que les champs d'intensité de pluie sont les plus pertinents. Comme ils donnent un diagnostic spatio-temporel prenant aussi en compte la localisation, les diagrammes de sévérité pourraient être utilisés pour la sélection ou au moins le tri des membres dont les champs de sortie sont les plus réalistes.

L'analyse combinée de ces diagrammes de sévérité et des variables issues de avec la modélisation hydrologique de ces mêmes événements, pourrait permettre de mieux comprendre les relations entre les incertitudes de la simulation météorologique et les incertitudes de la réponse hydrologique des bassins.

Perspectives

The studies conducted in this thesis on the estimation errors of point and spatial rainfall by ground measurement devices give insights into the identification of the best configuration of an experimental rain gauge network. This is at the base of the new super-site that is under construction in the framework of the HyMeX project in Le Pradel, Ardèche.

Regarding the radar, the scale-invariant properties of spatial rainfall allow a possible redefinition of the relationships relating the radar reflectivity to the rainfall intensity. The proposition is to reduce systematic radar underestimation by applying cascade disaggregation. This method is complementary to the geostatistical studies currently developed at this purpose. A downscaling technique can be adopted to obtain fine-scale radar reflectivity fields characterized by realistic variability. Once the radar reflectivity Z and the raingauge rainfall rate R are at coherent spatio-temporal scales, Z - R calibrations can be implemented. This will result in higher rainfall amounts.

The disaggregation may give another benefit: get rid of the interpolation errors related to the transformation of polar data into a cartesian grid. In each grid mesh of a polar coordinate system, the reflectivity value is integrated over a given volume. This volume is different from one mesh to the next one. Therefore, a raw radar scan merges reflectivities integrated over dozens of spatial scales. If the polar data is downscaled to a scale of interest (say 10 m) and then averaged to reach the map grid resolution, the radar reflectivity maps will have homogeneous resolution.

The evidences that the extreme rainfall show either hyperbolic or exponential tails allow to elaborate a scale-invariant GEV model valid for the entire region. Similar analysis on the scale-invariance of rainfall events are going to be performed in the framework of the AMMA-CATCH survey, on point-rainfall series. Concerning the spatial rainfall, the rain gauge density is too coarse to have reliable estimations of the actual spatial rainfall depth in the AMMA-CATCH region. A possible solution to this problem can be the use of statistical disaggregation on the coarse-resolution satellite data.

Thanks to the scale-invariant IDF model and to the estimation of the Areal Reduction Factor ARF, it is now possible to assign a frequency to a rainfall event. This study lead to the possibility to determine the frequency of a spatial rainfall event, and to detect, through analysis of the severity diagrams, its critical scales in space and time. A database containing the severity diagrams of the intense events occurred in the region Cévennes-Vivarais is in project.

The utility of severity diagrams for the qualification of the meso-scale models performance has been demonstrated. The comparison of different meso-scale models from the point of view of Severity Diagrams allowed us to identify models that provide a good representation of the fine-scale variability of the rainfall phenomenon. Great potential may result from the application of severity diagrams to probabilistic meteorological simulations. In a small example, we have seen that severity diagrams are able to characterize each member of the ensemble prediction under the point of view of severity. By means of severity diagrams, we plan to investigate the initial/boundary conditions that affect the simulations leading to unlikely rainfall fields. In addition, downscaling techniques could be used to obtain fine-scale boundary conditions in which the variability of the variates is correctly reproduced. These two propositions may help to reduce the range of variability of the forecast ensemble.

A last application concerns the project MEDUP: the application of severity diagrams to the meteorological ensemble simulations gave an idea of the uncertainties of the mesoscale simulation. Associating these simulations with the hydrological modeling may allow to envision the relations between uncertainty in the simulation of the rainfall events and uncertainty of the hydrological response of the basin.

Bibliography

- Allamano, P., P. Claps, and F. Laio, Global warming increases flood risk in mountainous areas, *Geophysical Research Letters*, 2009.
- Anquetin, S., F. Miniscloux, and J. D. Creutin, Numerical simulation of orographic rainbands, *Journal of Geophysical Research*, 108(D8), doi:10.1029/2002JD001593, 2003.
- Anquetin, S., et al., The 8 and 9 september 2002 flash flood event in france: a model intercomparison, *Natural Hazards and Earth System Sciences*, 5, 741–754, 2005.
- Arnaud, M., and X. Emery, *Estimation et interpolation spatiale*, Hermes Science Publication, Paris, 2000.
- Arnaud, P., C. Bouvier, L. Cisneros, and R. Dominguez, Influence of rainfall spatial variability on flood prediction, *Journal of Hydrology*, 260, 216–230, 2002.
- Asquith, W., and J. Famiglietti, Precipitation areal-reduction factor estimation using an annual-maxima centered approach, *Journal of Hydrology*, 230, 55–69, 2000.
- Austin, P., and R. Houze, Analysis of the structure of precipitation patterns in New England, *Journal of Applied Meteorology*, 11, 926–935, 1972.
- Bacchi, B., and R. Ranzi, On the derivation of the areal reduction factor of storms, *Atmospheric Research*, 42, 123–135, 1996.
- Badas, M., R. Deidda, and E. Piga, Modulation of homogeneous space-time rainfall cascades to account for orographic influences, *Nat. Hazards Earth Sys. Sci.*, 6, 427–437, 2006.
- Barancourt, C., J.-D. Creutin, and J. Rivoirard, A method for delineating and estimating rainfall fields, *Water Resources Research*, 28(4), 1133–1143, 1992.
- Barrera, A., M. Barriendos, and M. Llasat, Extreme flash floods in barcelona county, *Advances in Geosciences*, 2, 111–116, 2005.
- Bastin, G., B. Lorent, and M. Gevers, Optimal estimation of the average rainfall and optimal selection of rain gauge locations, *Water Resources Research*, 20, 463–470, 1984.
- Bechtold, P., E. Bazile, F. Guichard, P. Mascart, and E. Richard, A mass-flux convection scheme for regional and global models, *Quarterly Journal of the Royal Meteorological Society*, 127(573), 869–886, 2001.

- Bell, F., The areal reduction factor in rainfall frequency estimation. institute of hydrology report no 35,, *Tech. rep.*, Natural Environment Research Council, UK, 1976.
- Bell, T., A space-time stochastic model of rainfall for satellite remote-sensing studies, *Journal of Geophysical Research*, *92*, 9631–9643, 1987.
- Bell, T., A. Abdullah, R. Martin, and G. North, Sampling errors for satellite-derived tropical rainfall: Monte carlo study using a space-time stochastic model, *Journal of Geophysical Research*, *95*, 2195–2205, 1990.
- Bendjoudi, H., P. Hubert, and D. Schertzer, Interpretation multifractale des courbes intensité-durée-fréquence des précipitations, *C. R. Acad. Sci. Paris*, *2*, 323–326, 1997.
- Bernard, M., Formulas for rainfall intensities of long durations, *Transactions of the American Society of Civil Engineers*, *96*, 592–624, 1932.
- Berne, A., G. Delrieu, and B. Boudevillain, Variability of the spatial structure of intense mediterranean precipitation, *Advances in Water Resources*, pp. 1031–1042, 2009.
- Blanchet, J., C. Marty, and M. Lehning, Extreme value statistics of snowfall in the swiss alpine region, *Water Resources Research*, *45*(5), 2009.
- Bois, P., H. Mailloux, C. Obled, and F. De Saintignon, Atlas expérimental des risques de pluies intenses dans la région Cévennes-Vivarais, *Tech. rep.*, LAMA, Pôle Grenoblois des Risques Naturels, 1997.
- Borga, M., C. Vezzani, and G. Dalla Fontana, Regional rainfall depth-duration-frequency equations for an alpine region, *Natural Hazards*, *36*, 221–235, 2005.
- Bousquet, O., C. Lin, and I. Zawadski, Analysis of scale dependance of quantitative precipitation forecast verification: A case study over the mackenzie river basin, *Quarterly Journal of the Royal Meteorological Society*, *132*(620), 2107–2125, doi:10.1256/qj.05.154, 2006.
- Bras, R., and I. Rodriguez-Iturbe, Rainfall generation' a nonstationary time-varying multidimensional model, *Water Resources Research*, *12*(3), 450–456, 1976.
- Brath, A., A. Castellarin, M. Franchini, and G. Galeati, Estimating the index flood using indirect methods, *Hydrological Sciences Journal*, *46*(3), 399–418, 2001.
- Buishand, T., Some remarks on the use of daily rainfall models, *Journal of Hydrology*, *36*, 295–308, 1978.
- Burlando, P., and R. Rosso, Scaling and multiscaling models of depht-duration-frequency curves for storm precipitation, *Journal of Hydrology*, *187*, 45–64, 1996.
- Caniaux, G., J.-L. Redelsperger, and J. Lafore, A numerical study of the stratiform region of a fast-moving squall line. part i: General description and water and heat budgets, *Journal of Atmospheric Sciences*, *51*(14), 2046–2074, 1994.

- Carsteanu, A., and E. Foufoula-Georgiou, Assessing dependence of weights in a multiplicative cascade model of temporal rainfall, *Journal of Geophysical Research*, 101, 26,363–26,370, 1996.
- Castro, J., A. Carsteanu, and C. Flores, Intensity-duration-area-frequency functions for precipitation in a multifractal framework, *Physica A*, 338, 206–210, 2004.
- Ceresetti, D., G. Molinié, and J. Creutin, Scaling properties of heavy rainfall at short duration: a regional analysis, *Water Resources Research*, in press, 2010.
- Ceresetti, D., E. Ursu, J. Carreau, L. Gardes, S. Girard, G. Molinié, S. Anquetin, and J. Creutin, Evaluation of three extreme rainfall regionalization schemes in Cévennes-vivarais, france, *Nat. Hazards Earth Syst. Sci.*, in preparation, 2011.
- Cernesson, F., J. Lavabre, and J. Masson, Stochastic model for generating hourly hyetographs, *Atmospheric Research*, 42, 149–161, 1996.
- Chan, G., and A. Wood, An algorithm for simulating stationary gaussian random elds, *Journal of Royal Statistical Society, C(46)*, 171–181, 1997.
- Chapman, T., Stochastic modelling of daily rainfall:the impact of adjoining wet days on the distribution of rainfall amounts, *Environmental Modelling and Software*, 13, 317–324, 1998.
- Chiles, J., and P. Delfiner, *Geostatistics modeling spatial uncertainty*, New York: Wiley, 1999.
- Ciach, G., Local random errors in tipping-bucket rain gauge measurements, *Journal of Atmospheric and Oceanic Technology*, 20, 752–759, 2003.
- Clauset, A., C. Shalizi, and M. Newman, Power-law distributions in empirical data, *SIAM Review*, 1(51), 661–703, 2009.
- Coles, S., *An Introduction to Statistical Modeling of Extreme Values*, Springer-Verlag London Limited 2001, 2001.
- Cooley, D., P. Naveau, and P. Poncet, *Lecture Notes in Statistics -Dependence in Probability and Statistics*, vol. 187, chap. Part III, pp. 373–390, Springer New York, 2006.
- Cosma, S., E. Richard, and F. Miniscloux, The role of small-scale orographic features in the spatial distribution of precipitation, *Quarterly Journal of the Royal Meteorological Society*, 128, 1–18, 2002.
- Costa, A., and A. Soares, Trends in extreme precipitation indices derived from a daily rainfall database for the south of portugal, *International Journal of Climatology*, 29, 1956–1975, 2009.
- Cowpertwait, P., C. Kilsby, and P. O’Connell, A space-time neyman-scott model of rainfall: Empirical analysis of extremes, *Water Resources Research*, 38(8), 1131, 2002.
- Cressie, N., *Statistics for spatial data*, John Wiley and Sons Inc., New York, 1993.
- Creutin, J., and C. Obled, Objective analyses and mapping techniques for rainfall fields: an objective comparison, *Water Resources Research*, 18(2), 413–431, 1982.

- Davison, A. C., and R. L. Smith, Models for exceedances over high thresholds (with discussion), *Journal of the Royal Statistical Society*, 52, 1990.
- De Lima, M., J. De Lima, and M. Coelho, Spectral analysis of scale invariance in the temporal structure of precipitation in mainland Portugal, *Engenharia Civil*, 16, 73–82, 2003.
- De Michele, C., and P. Bernardara, Spectral analysis and modeling of space-time rainfall fields, *Atmospheric Research*, 77, 124–136, 2005.
- De Michele, C., N. Kottegoda, and R. Rosso, The derivation of areal reduction factor of storm rainfall from its scaling properties, *Water Resources Research*, 37(12), 3247–3252, 2001.
- De Michele, C., N. Kottegoda, and R. Rosso, Idaf (intensity-duration-area-frequency) curves of extreme storm rainfall: a scaling approach, *Water Science and Technology*, 45(2), 83–90, 2002.
- de Montera, L., L. Barthés, C. Mallet, and P. Golé, The effect of rain-no rain intermittency on the estimation of the universal multifractals model parameters, *Journal of Hydrometeorology*, 10(2), 493–506, 2008.
- Deidda, R., Rainfall downscaling in a space-time multifractal framework, *Water Resources Research*, 36, 1779–1794, 2000.
- Deidda, R., R. Benzi, and F. Siccaldi, Multifractal modeling of anomalous scaling laws in rainfall, *Water Resources Research*, 35, 1853–1867, 1999.
- Delclaux, F., and V. Thauvin, *Spatial - Variogramme. Algorithmes, Traitements principaux - Manuel Utilisateur*, IRD Hydrologie, 1993.
- Delhomme, J., and P. Delfiner, Application du krigeage à l'optimisation d'une campagne pluviométrique en zone aride., in *Proceedings of the Symposium on Design of Water Resources - Projects with Inadequate Data*, vol. 2, pp. 191–210, UNESCO, Madrid, 1973.
- Delrieu, G., L'Observatoire Hydro-météorologique Méditerranéen Cévennes-Vivarais (the Cévennes-Vivarais Mediterranean Hydro-meteorological Observatory), *La Houille Blanche*, 6-2003, 83–88, 2004.
- Delrieu, G., B. Boudevillain, B. Chapon, P. Kirstetter, J. Nicol, H. Andrieu, and D. Faure, Bollène 2002 experiment: radar rainfall estimation in the Cévennes-Vivarais region, France., *Journal of Applied Meteorology and Climatology*, 48(7), 1422–1447, 2009.
- Delrieu, G., et al., The catastrophic flash-flood event of 8-9 September 2002 in the Gard region, France: a first case study for the Cévennes-Vivarais Mediterranean Hydro-meteorological Observatory, *Journal of Hydrology*, 6(1), 34–52, 2005.
- Dietrich, C., and G. Newsam, A fast and exact method for multidimensional gaussian stochastic simulations, *Water Resources Research*, 29, 2861–2869, 1993.

- Ducrocq, V., D. Ricard, J. Lafore, and F. Orain, Storm-scale numerical rainfall prediction for five precipitating events over france: On the importance of the initial humidity field, *Weather and Forecasting*, *17*, 1236–1256, 2002.
- Ducrocq, V., G. Aullo, and P. Santurette, Précipitations intenses et les inondations des 12 et 13 novembre 1999 sur le sud de la France, *La Météorologie*, *42*, 18–27, 2003.
- Ducrocq, V., O. Nuissier, D. Ricard, C. Lebeaupin, and T. Thouvenin, A numerical study of three catastrophic precipitating events over southern France. ii: Mesoscale triggering and stationarity factors, *Quarterly Journal of the Royal Meteorological Society*, *134*, 131–145, 2008.
- Eagleson, P., N. Fenessey, W. Quiliang, and I. Rodriguez-Iturbe, Application of spatial poisson models to airmass thunferstorm rainfall, *Journal of Geophysical Research*, *92*, 9961–9978, 1987.
- Fabry, F., On the determination of scale ranges for precipitation fields, *Journal of Geophysical Research*, *101*, 12,819–12,826, 1996.
- Feldmann, A., A.-C. Gilbert, and W. Willinger, Data networks as cascades: Investigating the multifractal nature of internet wan traffic, in *ACM/SIGCOMM 98 Proceedings: Applications, Technologies, Architectures, and Protocols for Computer Communication*, 1998.
- Fraedrich, K., and C. Larnder, Scaling regimes of rainfall time series, *Tellus*, *45A*, 289–298, 1993.
- Frisch, U., *Turbulence: the legacy of A.N. Kolmogorov*, Cambridge University Press, 1995.
- Frisch, U., and G. Parisi, Fully developed turbulence and intermittency, *Turbulence and Predictability in Geophysical Fluid Dynamics and Climate Dynamics*, pp. 84–88, 1985.
- Gaume, E., N. Mouhous, and H. Andrieu, Rainfall stochastic disaggregation models: Calibration and validation of a multiplicative cascade model, *Advances in Water Resources*, *30*, 1301–1319, 2007.
- Gerold, L., and D. Watkins, Short duration rainfall frequency analysis in michigan using scale-invariance assumptions, *Journal of Hydrologic Engineering*, 2005.
- Gnedenko, B., Sur la distribution limite du terme maximum d’une serie aleatoire, *Annals of Mathematics*, *44*, 1943.
- Godart, A., Les precipitations orographiques organisees en bandes dans la region cévemes-vivarais: caracterisation et contribution au regime pluviometrique, Ph.D. thesis, Université Joseph Fourier, Grenoble, France, 2009.
- Godart, A., S. Anquetin, and E. Leblois, Rainfall regimes associated with banded convection in the cévennes-vivarais area, *Meteorology and Atmospheric Physics*, *103*, 25–34, 2009.
- Goldstein, M., S. Morris, and G. Yen, Problems with fitting to the power-law distributions, *European Physics Journal*, *B(41)*, 255–258, 2004.
- Goovaerts, P., Comparative performance of indicator algorithms for modeling conditional probability distribution functions, *Mathematical Geology*, 1994.

- Goovaerts, P., *Geostatistics for Natural Resources Evaluation*, Oxford University Press, New York, 1997.
- Goovaerts, P., Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall, *Journal of Hydrology*, 2000.
- Gottardi, F., Estimation statistique et r analyse des pr cipitations en montagne, Ph.D. thesis, Institut Polytechnique de Grenoble, 2009.
- Guillot, G., and T. Lebel, Disaggregation of sahelian mesoscale convective system rain fields: Further developments and validation, *Journal of Geophysical Research*, 1999.
- Guillot, P., The arguments of the gradex method: a logical support to assess extreme floods, in *Proceedings of the Yokohama Symposium, Extreme Hydrological Events: Precipitation, Floods and Droughts*, vol. IAHS Publ. no. 213, 1993.
- Guillot, P., and D. Duband, la m thode du GRADEX pour le calcul de la probabilit  des crues   partir des pluies, in *Journ es de la SHF*, vol. Rapport 7, question 1, Soci t  Hydrologique de France, Soci t  Hydrologique de France, Paris, 1967.
- Gumbel, E., *Statistics of Extremes*, Columbia University Press, New York, 1958.
- G ntner, A., J. Olsson, A. Calver, and B. Gannon, Cascade-based disaggregation of continuous rainfall time series: the influence of climate, *Hydrology and Earth System Sciences*, 5(2), 145–164, 2001.
- Gupta, V., and E. Waymire, A stochastic kinematic study of subsynoptic space-time rainfall, *Water Resources Research*, 15(3), 637–644, 1979.
- Gupta, V., and E. Waymire, Multiscaling properties of spatial rainfall and river flow distributions, *Journal of Geophysical Research*, 95(D3), 1999–2009, 1990.
- Gupta, V., and E. Waymire, A statistical analysis of mesoscale rainfall as a random cascade, *Journal of Applied Meteorology*, 32, 251–267, 1993.
- Gupta, W., and E. Waymire, On Taylor’s hypothesis and dissipation in rainfall, *Journal of Geophysical Research*, 92(D8), 9657–9660, 1987.
- Haberlandt, U., Geostatistical interpolation of hourly precipitation from rain gauges and radar for a large-scale extreme rainfall event, *Journal of Hydrology*, 332, 144–157, 2007.
- Habib, E., W. Krajewski, and A. Kruger, Sampling errors of tipping-bucket rain gauge measurements, *Journal of Hydrologic Engineering*, pp. 159–166, 2001.
- Harris, D., M. Menabde, A. Seed, and G. Austin, Multifractal characterization of rain fields with a strong orographic influence, *Journal of Geophysical Research*, 101(D21), 26,405–26,414, 1996.
- Harris, D., E. Foufoula-Georgiou, K. Droegemeier, and J. Levit, Multiscale properties of a high-resolution precipitation forecast, *Journal of Hydrometeorology*, 2, 406–418, 2001.

- Hastie, T. J., and R. Tibshirani, *Generalized additive models*, Chapman and Hall Ltd., London, 1990.
- Herr, H., and Krzysztofowicz, Generic probability distribution of rainfall in space: the bivariate model, *Journal of Hydrology*, 306, 234–263, 2005.
- Hosking, J., L-moments: analysis and estimation of distributions using linear combination of order statistics, *Journal of the Royal Statistical Society*, 52, 105–124, 1990.
- Hubert, P., and H. Bendjoudi, Introduction à l'étude des longues séries pluviométriques, in *XIIèmes Journées Hydrologiques de l'Orstom*, 1996.
- Hubert, P., and J.-P. Carbonnel, Dimensions fractales de l'occurrence de pluie en climat soudano-sahlien, *Hydrologie continentale*, 4, 3–10, 1989.
- Hubert, P., Y. Tessier, S. Lovejoy, P. Ladoy, J. Carbonner, S. Violette, and I. Desurogne, Multifractals and extreme rainfall events, *Geophysical Research Letters*, 20(10), 931–934, 1993.
- Jacq, V., Inventaire des situations à précipitations diluviennes sur les régions Languedoc-Roussillon, PACA et Corse, période 1958-1994 (inventory of extreme precipitation events in the Languedoc-Roussillon, PACA and Corse regions in the period 1958-2004). phénomènes remarquables., *Tech. Rep. 3*, Météo-France (SCEM), 1994.
- Journal, A., Nonparametric estimation of spatial distributions, *International Association for Mathematical Geology*, 15(3), 445–468, 1983.
- Journal, A., and C. Huijbregts, *Mining geostatistics*, Elsevier, New York, 1978.
- Kahane, J., and J. Peyriere, Sur certaines martingales de benoit mandelbrot, *Advances in Mathematics*, 22(131-145), 1976.
- Kedem, B., and L. Chiu, On the lognormality of rain rate, in *Proc. Natl. Acad. Sci. USA, Applied mathematical Sciences*, vol. 84, pp. 901–905, 1987.
- Kirshbaum, D., G. Bryan, R. Rotunno, and D. Durran, The triggering of orographic rainbands by small scale topography, *Journal of the Atmospheric Sciences*, 64, 4222–4245, 2007.
- Kolmogorov, A., Local structure of turbulence in an incompressible liquid for very large reynolds numbers, *C. R. Acad. Sci. USSR*, 30, 301–305, 1941.
- Kolmogorov, A., A refinement of previous hypothesis concerning the local structure of turbulence in viscous incompressible fluid at high reynolds numbers, *Journal of Fluid Mechanics*, 13, 82–85, 1962.
- Kottegoda, N., and R. Rosso, *Probability and Reliability for Civil and Environmental Engineers*, Blackwell Pub, 1997.
- Kotz, S., and S. Nadarajah, *Extreme value distributions. Theory and applications*, Imperial College Press, 2000.

- Koutsoyiannis, On the appropriateness of the gumbel distribution for modelling extreme rainfall, in *ESF LESC Exploratory Workshop*, Bologna, Italy, 2003.
- Koutsoyiannis, D., *Statistical Hydrology*, National Technical University of Athens, 1997.
- Koutsoyiannis, D., D. Kozonis, and A. Manetas, A mathematical framework for studying rainfall intensity - duration - frequency relationships, *Journal of Hydrology*, 206, 118–135, 1998.
- Krajewski, W. F., et al., Devex-disdrometer evaluation experiment: Basic results and implications for hydrologic studies, *Advances in Water Resources*, 29, 311–325, 2006.
- Krige, D., A statistical approach to some basic mine valuation problems on the witwatersrand, *Journal of the Chemical, Metallurgical and Mining Society*, pp. 119–139, 1951.
- Kumar, P., and E. Foufoula-Georgiou, A new look at rainfall fluctuations and scaling properties of spatial rainfall using orthogonal wavelets, *Journal of Applied Meteorology*, 32, 1993.
- Kumar, P., P. Guttorp, and E. Foufoula-Georgiou, A probability-weighted moment test to assess simple scaling, *Stochastic Hydrology and Hydraulics*, 8, 173–183, 1994.
- Kyriakidis, P., and A. Journel, Geostatistical space-time models: A review, *Mathematical Geology*, pp. 651–684, 1999.
- Labalette, J., Characterisation de la sévérité des orages dans les cévennes, Master's thesis, Université Joseph Fourier Grenoble, 2009.
- Ladoy, P., F. Schmitt, D. Schertzer, and S. Lovejoy, Variabilité temporelle multifractale des observations pluviométriques à Nîmes, *C.R. Acad. Sci. Paris*, 317(2), 775–782, 1993.
- Lafore, J. P., et al., The Meso-NH atmospheric simulation system. part i: Adiabatic formulation and control simulations., *Annales Geophysicae*, 16, 90–109, 1998.
- Langousis, A., The areal reduction factor: A multifractal analysis, Master's thesis, MIT Boston, 2005.
- Lavallée, D., Multifractal analysis and simulation technique and turbulent fields, Ph.D. thesis, McGill University, Montréal, 1991.
- Le Cam, L., A stochastic description of precipitation, in *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, 3, edited by C. J. Neyman, Berkeley, pp. 165–186, 1961.
- Lebeaupin, C., V. Ducrocq, and H. Giordani, Sensitivity of torrential rain events to the sea surface temperature based on high-resolution numerical forecasts, *Journal of Geophysical Research*, 2006.
- Lebel, T., and G. Bastin, Variogram identification by the mean-squared interpolation error method with application of hydrologic fields, *Journal of Hydrology*, 77(1-4), 31–56, 1985.
- Lebel, T., and J. P. Laborde, A geostatistical approach for areal rainfall statistics assessment, *Stochastic Hydrology and Hydraulics*, 2(4), 245–261, doi:10.1007/BF01544039, 1988.

- Lebel, T., G. Bastin, C. Obled, and J.-D. Creutin, On the accuracy of areal rainfall estimation: a case study, *Water Resources Research*, *23*, 2,123–2,138, 1987.
- Lorenz, E., Deterministic non-periodic flow, *Journal of the Atmospheric Sciences*, *20*, 130–141, 1963.
- Lovejoy, S., Area-perimeter relation for rain and cloud areas, *Science*, *216*(4542), 185–187, 1982.
- Lovejoy, S., and B. Mandelbrot, Fractal properties of rain, and a fractal model, *Tellus*, *37*(A), 209–232, 1985.
- Mandelbrot, B., How long is the coast of Britain? statistical self-similarity and fractional dimension, *Science*, *156*, 636–638, 1967.
- Mandelbrot, B., Intermittant turbulence in self-similar cascades: Divergence of high moments and dimension of the carrier, *Journal of Fluid Mechanics*, *62*, 331–358, 1974.
- Mandelbrot, B., *The Fractal Geometry of Nature*, W.H. Freeman, 1982.
- Mandelbrot, B., *Fractals and Scaling in Finance: Discontinuity, Concentration, Risk*, Springer-Verlag, New York, 1997.
- Mandelbrot, B., and J. Wallis, Some long-run properties of geophysical records, *Water Resources Research*, *5*(2), 1969a.
- Mandelbrot, B., and J. Wallis, Robustness of the rescaled range r/s in the measurement of noncyclic long run statistical dependence, *Water Resources Research*, *5*(5), 967–988, 1969b.
- Mantoglou, A., and J. Wilson, The turning bands method for simulation of random elds using line generation by a spectral method., *Water Resources Research*, *18*, 1379–1394, 1982.
- Marani, M., On the correlation structure of continuous and discrete point rainfall, *Water Resources Research*, *39*(5), 2003.
- Marani, M., Non-power-law-scale properties of rainfall in space and time, *Water Resources Research*, *41*(W08413), doi:10.1029/2004WR003822,, 2005.
- Marsalek, J., Calibration of the tipping bucket raingage, *Journal of Hydrology*, pp. 343–354, 1981.
- Marsan, D., D. Schertzer, and S. Lovejoy, Causal space-time multifractal processes: Predictability and forecasting of rain fields, *Journal of Geophysical Research*, *101*, 26,333–26,346, 1996.
- Mason, I., Dependence of the critical success index on sample climate and threshold probability, *Aust. Meteor. Mag.*, *37*, 75–81, 1989.
- Matheron, G., *Traité de géostatistique appliquée, Tome 1*, Mémoires du Bureau de Recherche Géologiques et Minières (n14), 1962.
- Matheron, G., Les variables régionalisées et leur estimation, Thèse de Doctorat d'Etat (in French), Masson, Paris, 1965.

- Matheron, G., *La théorie des variables régionalisées et ses applications.*, vol. Cours fascicule n5, Centre de morphologie mathématique de Fontainebleau, 1970.
- Matheron, G., *Théorie des variables régionalisées*, pp. 306–378, Masson, Paris, 1972.
- Matheron, G., The intrinsic random functions and their applications, *Advances in Applied Probability*, 5, 439–468, 1973.
- Matheron, G., Suffit-il, pour une covariance, d'être de type positif?, *Sciences de la Terre, série informatique géologique*, 26, 51–66, 1987.
- Menabde, M., and M. Sivapalan, Modeling of rainfall time series and extremes using bounded random cascades and levy-stable distributions, *Water Resources Research*, 36(11), 3293–3300, 2000.
- Menabde, M., D. Harris, A. Seed, G. Austin, and D. Stow, Multiscaling properties of rainfall and bounded random cascades, *Water Resources Research*, 33(12), 2823–2830, 1997.
- Menabde, M., A. Seed, and G. Pegram, A simple scaling model for extreme rainfall, *Water Resources Research*, 35(1), 335–339, 1999.
- Meneveau, C., and K. Sreenivasan, Simple multifractal cascade model for fully developed turbulence, *Physical Review Letters*, 59, 1424–1427, 1987.
- Mestre, O., Modles de valeurs extremes et prditeurs, 3mes Rencontres Mto/MathAppli Toulouse, 2008.
- Miniscloux, F., J. D. Creutin, and S. Anquetin, Geostatistical analysis of orographic rainbands, *Journal of Applied Meteorology*, 40, 1835–1854, 2001.
- Molinié, G., D. Ceresetti, S. Anquetin, B. Creutin, and J. D. Creutin, Rainfall regimes in a Mediterranean mountainous region, *submitted to Quarterly Journal of the Royal Meteorological Society*, 2010.
- Muzy, J., E. Bacry, and A. Arneodo, Multifractal formalism for fractal signals: The structure-function approach versus the wavelet-transform modulus-maxima method, *Physical Review*, E(47), 875–884, 1993.
- Naidu, C., K. Durgalakshmi, K. Muni Krishna, R. R. S., G. Satyanarayana, P. Lakshminarayana, and L. Malleswara Rao, Is summer monsoon rainfall decreasing over india in the global warming era?, *Journal of Geophysical Research*, 2009.
- Neppel, L., Analyse de l'épisode pluvieux du 08 et 09 septembre 2002, *Tech. rep.*, Hydrosiences Montpellier - Ministère de l'Écologie et du Développement Durable, 2003.
- NERC, Flood studies report - meteorological studies, *Tech. rep.*, UK Natural Environmental Research Council, 1975.
- Nguyen Thao, T. P. N. G., P. Bois, and J. A. Villasenor, Simulation in order to choose a fitting method for extreme rainfall, *Atmospheric Research*, 30, 13–36, 1993.

- Nolan, J., *Lévy processes: theory and applications*, chap. VI, Birkhauser, Boston, 2001.
- Nolan, J., *Stable Distributions. Models for Heavy Tailed Data*, chap. 1, Math/Stat Department - American University, 2009.
- Northrop, P., A clustered spatial-temporal model of rainfall, in *Proceedings of the Royal Society London*, vol. Ser. A-454, pp. 1875–1888, 1998.
- Nuissier, O., V. Ducrocq, D. Ricard, C. Lebeaupin, and S. Anquetin, A numerical study of three catastrophic precipitating events over western mediterranean region (Southern France): Part I: Numerical framework and synoptic ingredients, *Quarterly Journal of the Royal Meteorological Society*, 134(630), 111–130, 2008a.
- Nuissier, O., V. Ducrocq, D. Ricard, C. Lebeaupin, and S. Anquetin, A numerical study of three catastrophic precipitating events over southern france. i: Numerical framework and synoptic ingredients, *Quarterly Journal of the Royal Meteorological Society*, 134, 111–130, 2008b.
- Obled, C., Cours d'initiation à la géostatistique, Institut National Polytechnique de Grenoble - INPG, 2007.
- Obled, C., J. Wendling, and K. J. Beven, The sensitivity of hydrological models to spatial rainfall patterns: an evaluation using observed data., *Journal of Hydrology*, 159, 305–333, 1994.
- Olsson, J., Limits and characteristics of the multifractal behaviour of a high-resolution rainfall time series, *Nonlinear Processes in Geophysics*, 2, 23–29, 1995.
- Olsson, J., Validity and applicability of a scale-independent, multifractal relationship for rainfall, *Atmospheric Research*, 42, 53–65, 1996.
- Olsson, J., Evaluation of a scaling cascade model for temporal rainfall disaggregation, *Hydrology and Earth System Sciences*, 2(1), 19–30, 1998.
- Olsson, J., and J. Niemczynowicz, Multifractal analysis of daily spatial rainfall distributions, *Journal of Hydrology*, 187, 29–43, 1996.
- Olsson, J., J. Niemczynowicz, and R. Berndtsson, Fractal analysis of high-resolution rainfall time series, *Journal of Geophysical Research*, 1993.
- Olsson, J., V. P. Singh, and K. Jinno, Effect of spatial averaging on temporal statistical and scaling properties of rainfall, *Journal of Geophysical research*, 104(D16), 19,117–19,126, 1999.
- Omelayo, A., On the transportation of areal reduction factors for rainfall frequency estimation, *Journal of Hydrology*, 145, 191–205, 1993.
- Over, T., and V. Gupta, Statistical analysis of mesoscale rainfall: Dependence of a random cascade generator on large-scale forcing, *Journal of Applied Meteorology*, 33, 1526–1542, 1994.
- Over, T., and V. Gupta, A space-time theory of mesoscale rainfall using random cascades, *Journal of Geophysical Research*, 101, 26,319–26,331, 1996.

- Padoan, S., Computational methods for complex problems in extreme value theory, Ph.D. thesis, University of Padova, 2008.
- Palmer, T., and Räisänen, Quantifying the risk of extreme seasonal precipitation events in a changing climate, *Nature*, *415*, 2002.
- Pandey, G., S. Lovejoy, and D. Schertzer, Multifractal analysis of river flows including extremes for basins of five to two million square kilometers, one day to 75 years, *Journal of Hydrology*, *208*, 62–81, 1998.
- Perica, S., and Foufoula-Georgiou, Model for multiscale disaggregation of spatial rainfall based on coupling meteorological and scaling descriptions, *Journal of Geophysical Research*, *101*(D21), 26,347–26,361, 1996.
- Pickands, J., Statistical inference using extreme order statistics, *The Annals of Statistics*, *3*(1), 119–131, 1975.
- Prudhomme, C., and D. Reed, Mapping extreme rainfall in a mountainous region using geostatistical techniques: a case study in scotland, *International Journal of Climatology*, *19*, 1337–1356, 1999.
- Ramos, M. H., J. D. Creutin, and E. Leblois, Visualization of storm severity, *Journal of Hydrology*, *315*, 295–307, 2005.
- Reiss, R., and M. Thomas, *Statistical Analysis of Extreme Values*, Birkhäuser Verlag, Basel-Boston-Berlin, 1997.
- Renard, B., and M. Lang, Use of a gaussian copula for multivariate extreme value analysis: Some case studies in hydrology, *Advances in Water Resources*, *30*, 897–912, 2007.
- Ribatet, M., Pot: Modelling peaks over a threshold, *R News*, *7*, 34–36, 2007.
- Ricard, D., Initialisation et assimilation de données à méso-échelle pour la prévision à haute résolution des pluies intenses de la région cévennes-vivarais (in french), Ph.D. thesis, Université Pierre Sabatier-Toulouse III (France), 2002.
- Roche, M., *Hydrologie de Surface*, Gauthier-Villars, Paris, 1963.
- Rodriguez-Iturbe, I., and J. M. Mejía, On the transformation of point rainfall to areal rainfall, *Water Resources Research*, *10*(4), 729–735, 1974.
- Rodriguez-Iturbe, I., D. Cox, and P. Eagleson, Spatial modelling of total storm rainfall, in *Proceedings of the Royal Society London*, vol. Ser.A-403, pp. 27–50, 1986.
- Rosso, R., and P. Burlando, Scale invariance in temporal and spatial rainfall, in *Proceedings XV General Assembly European Geophysical Society*, 1990.
- Ruin, I., J.-D. Creutin, S. Anquetin, and C. Lutoff, Human exposure to flash-floods relation between flood parameters and human vulnerability during a storm of september 2002 in southern france, *Journal of Hydrology*, pp. 199–213, 2008.

- Rupp, D., R. Keim, M. Ossiander, M. Brugnach, and J. Selker, Time scale and intensity dependency in multiplicative cascades for temporal rainfall disaggregation, *Water resources research*, 45(9), 74,091–74,914, 2009.
- Salvadori, G., and C. De Michele, From generalized pareto to extreme values law: Scaling properties and derived features, *Journal of Geophysical Research*, 106(D20), 24,063–24,070, 2001.
- Samorodnitsky, G., and M. S. Taqqu, *Stable Non-Gaussian Random Processes*, Chapman and Hall, New York, 1994.
- Schertzer, D., and S. Lovejoy, Generalized scale invariance in turbulent phenomena, *Physico Chemical Hydrodynamics Journal*, 6, 623–635, 1985.
- Schertzer, D., and S. Lovejoy, Physically based rain and cloud modeling by anisotropic, multiplicative turbulent cascades, *Journal of Geophysical Research*, 92, 9692–9714, 1987.
- Schertzer, D., and S. Lovejoy, Hard and soft multifractal processes, *Physica A*, 185, 187–194, 1992.
- Schertzer, D., S. Lovejoy, and F. Schmitt, Structures in turbulence and multifractal universality, in *Small Scale Structures in 3D hydro and MHD turbulence*, vol. Lecture notes in physics vol. 462, edited by A. P. M. Meneguzzi and P. Sulem, pp. 137–144, Springer, 1995.
- Schlater, M., Models for stationary max-stable random fields, *Extremes*, 5(1), 33–44, 2002.
- Schmitt, F., S. Vannitsem, and A. Barbosa, Modeling of rainfall time series using two-state renewal processes and multifractals, *Journal of Geophysical Research*, 103, 23,181–23,193, 1998.
- Sénési, S., P. Bougeault, J. Chèze, P. Cosentino, and R. Thepenier, The Vaison-la-Romaine flash flood: Mesoscale analysis and predictability issues, *Weather and Forecasting.*, 11(4), 417–442, 1996.
- She, Z., and E. Waymire, Quantized energy cascade and log-poisson statistics in fully developed turbulence, *Physical Review Letters*, 74(2), 262–265, 1995.
- Sivakumar, B., Fractal analysis of rainfall observed in two different climatic regions, *Hydrological Sciences Journal*, 45(5), 727–738, 2000.
- Sivapalan, M., and G. Blöschl, Transformation of point rainfall to areal rainfall: Intensity - duration - frequency curves, *Journal of Hydrology*, 204, 150–167, 1998.
- Skoien, J., and G. Blöschl, Sampling scale effects in random fields and implications for environmental monitoring, *Environmental Monitoring and Assessment*, 114, 521–552, 2006.
- Slimani, M., and T. Lebel, Comparison of three methods of estimating rainfall frequency parameters according to the duration of accumulation., in *International symposium on flood frequency and risk analyses*, vol. Hydrologic frequency modeling, pp. 277–291, Singh, J.V.; Reidel, D. (eds.) ; D. Reidel Publishing Company, 1986.
- Smith, R., Max-stable processes and spatial extremes, department of Mathematics, University of Surrey, England, 1990.

- Smithers, J., and R. Schulze, The estimation of design rainfalls for south africa using a regional scale invariant approach, in *Hydrology: Science & Practice for the 21st Century*, edited by U. British Hydrological Society, London, pp. 221–228, 2004.
- Sornette, D., *Critical Phenomena in Natural Sciences*, Springer-Verlag Berlin, 2004.
- Stedinger, J., R. Vogel, and E. Foufoula-Georgiou, *Frequency analysis of extreme events*, pp. 18.1–18.66, McGraw-Hill, New York, 1993.
- Stein, M., Space-time covariance functions, *Journal of the American Statistical Association*, 100, 310–321, 2005.
- Steutel, F., and K. Van Harn, *Infinite Divisibility of Probability Distributions on the Real Line (Pure and Applied Mathematics)*, Marcel Dekker, 2003.
- Suhaila, J., and A. Jemain, Daily rainfall amount in peninsular malaysia using several types of exponential distributions, *Journal of Applied Sciences Research*, 3(10), 1027–1036, 2007.
- Svensson, C., Review of methods for deriving areal reduction factors, *Tech. rep.*, Department for Environment Food and Rural Affairs - UK, 2007.
- Svensson, C., J. Olsson, and R. Berndtsson, Multifractal properties of daily rainfall in two different climates, *Water Resources Research*, 32, 2463–2472, 1996.
- Taylor, G., The spectrum of turbulence, *Proceedings Royal Society London A*, 164(919), 1938.
- Tessier, Y., S. Lovejoy, and D. Schertzer, Universal multifractals: Theory and observations for rain and clouds, *Journal of Applied Meteorology*, 32(2), 223–250, 1993.
- Tessier, Y., P. Lovejoy, P. Hubert, D. Schertzer, and S. Pecknold, Multifractal analysis and modeling of rainfall and river flows and scaling causal transfer functions, *Journal of Geophysical Research*, 101(26371-26392), 1996.
- Todorovic, P., and D. Woolhiser, A stochastic model of n-day precipitation, *Journal of Applied Meteorology*, 14(1), 17–24, 1975.
- U.S. Weather Bureau, ., Rainfall intensity-frequency regime parts 1 and 2, Technical Paper no. 29., *Tech. rep.*, US Weather Bureau, 1958.
- Vainshtein, S., K. Sreenivasan, R. Pierrehumbert, V. Kashyap, and A. Juneja, Scaling exponents for turbulence and other random processes and their relationships with multifractal structure, *Physical Review E*, 50(3), 1823–1835, 1994.
- Veneziano, D., and P. Furcolo, Multifractality of rainfall and scaling of intensity-duration-frequency curves, *Water Resources Research*, 38(12), 1306, 2002.
- Veneziano, D., and A. Langousis, The areal reduction factor: A multifractal analysis, *Water Resources Research*, 41, W07,008, 2005.

- Veneziano, G., P. Furcolo, and V. Jacobellis, Multifractality of iterated pulse processes with pulse amplitudes generated by a random cascade, *Fractals*, 10(2), 209–222, 2002.
- Venugopal, V., E. Foufoula-Georgiou, and V. Sapozhnikov, Evidence of dynamic scaling in space-time rainfall, *Journal of Geophysical research*, 104(D24), 31,599–31,610, 1999.
- Venugopal, V., S. Basu, and E. Foufoula-Georgiou, A new metric for comparing precipitation patterns with an application to ensemble forecasts, *Journal of Geophysical Research*, 110(D8), D08,111.1–D08,111.11, 2005.
- Venugopal, V., S. Roux, E. Foufoula-Georgiou, and A. Arneodo, Revisiting multifractality of high resolution temporal rainfall using a wavelet-based formalism, *Water Resources Research*, 42(110), 2006a.
- Venugopal, V., S. Roux, E. Foufoula-Georgiou, and A. Arneodo, Scaling behavior of high resolution temporal rainfall: New insights from a wavelet-based cumulant analysis, *Physics Letters A*, 348(3-6), 335–345, 2006b.
- Vogel, R., and N. Fennessey, L-moment diagrams should replace product moment diagrams, *Water Resources Research*, 29(6), 1745–1752, 1993.
- Wackernagel, H., *Multivariate Geostatistics - 2nd ed.*, Springer, Berlin, 1998.
- Waymire, E., and V. Gupta, The mathematical structure of rainfall representations. a review of the stochastic rainfall models, *Water Resources Research*, 17(5), 1261–1272, 1981.
- Waymire, E., V. Gupta, and I. Rodriguez-Iturbe, A spectra theory of rainfall intensity at the meso-beta scale, *Water Resources Research*, 20, 1465–1483, 1984.
- Wheater, H., V. Isham, R. Cox, R. Chandler, A. Kakou, P. Northrop, L. Oh, C. Onof, and I. Rodriguez-Iturbe, Spatial-temporal rainfall fields: Modelling and statistical aspects, *Hydrological Earth System Science*, 4, 581–601, 2000.
- Willems, P., A spatial rainfall generator for small spatial scales, *Journal of Hydrology*, 252, 126–144, 2001.
- Wood, A., and G. Chan, A simple and efficient space domain implementation of the turning bands method, *Water Resources Research*, 31, 147–156, 1995.
- Yan, X., V. Ducrocq, P. Poli, M. Hakam, G. Jaubert, and A. Walpersdorf, Impact of gps zenith delay assimilation on convective-scale prediction of mediterranean heavy rainfall, *Journal of Geophysical Research*, 114, 2009.
- Yates, E., J. Creutin, S. Anquetin, and J. Rivoirard, A scale dependant quality index of areal rainfall prediction, *Journal of Hydrometeorology*, 8, 160–170, 2007.
- Yates, F., Tests of significance for 2x2 contingency tables, *Journal of the Royal Statistical Society - Series A*, 147(3), 426–463, 1984.

- Zawadzki, I., Statistical properties of precipitation patterns, *Journal of Applied Meteorology*, 12, 459–473, 1973.
- Zepeda-Arce, J., E. Foufoula-Georgiou, and K. Droegemeier, Space-time rainfall organization and its role in validating quantitative precipitation forecasts, *Journal of Geophysical Research*, 105, 10,129–10,146, 2000.
- Zhang, L., and V. P. Singh, Gumbel–hougaard copula for trivariate rainfall frequency analysis, *Journal of Hydrologic Engineering*, 12(4), 409–419, doi:10.1061/(ASCE)1084-0699(2007)12:4(409), 2007.
- Zhang, Q., C.-Y. X., Z. Zhang, Y. Chen, C. Liu, and H. Lin, Spatial and temporal variability of precipitation maxima during 1960–2005 in the yangtze river basin and possible association with large-scale circulation, *Journal of Hydrology*, 353, 215–227, 2008.
- Zolotarev, *One-dimensional Stable Distributions*, American Mathematical Society, 1986.

APPENDIX A

List of Variables and symbols

- α : Power-law decay exponent,
 α_s : Levy's stability index
 β_p : Generalized Pareto scale parameter
 β : spectral density slope in double logarithmic plot
 β_s : parameter of stable distribution
 γ_s : Singularity order:
 γ_e : Euler-Mascheroni constant ~ 0.577
 ϵ : ($\ln(1 - 1/T_R)$) in GEV-formulation
 ϵ : ($\ln(1 - 1/T_R)$) in GEV-formulation
 ϵ_d : energy dissipation rate in the Kolmogorov Theory
 $\zeta(p)$: generalized structure function (*Frisch and Parisi, 1985*) as a function of the moment order p
 κ : variance reduction factor relating point and areal rainfall
 λ : scale ratio, ratio between large and small resolution
 λ_i : kriging interpolation weights
 μ : position parameter of GEV distribution
 μ_s : parameter of stable distribution
 ν : fluid viscosity in Navier-Stokes equation
 ρ : correlation function
 σ : scale parameter of GEV distribution
 σ_e^2 : error variance
 τ : time scale in frozen-turbulence hypothesis
 $\tau(q)$: moment scaling function
 $\tau(r)$: L-moment ratio of order r
 ϕ : trigonometric function of Levy's stable characteristic function
 ψ : haar wavelet
 ω (geostatistics): realization number
 ω (scaling theory): frequency coordinate in Fourier Analysis
 Γ : Gamma function, generalization of factorial function for $x \in \mathfrak{R}$, $\Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t} dt$
 Θ : Gumbel parameters of spatial rainfall maxima (*Lebel and Laborde, 1988*)
 Λ : expected number of occurrences within an interval in Poisson distribution
 Φ : spectral density of a signal
 a, b (introduction): empirical parameters of the Z-R relationship
 a (geostatistics): exponential covariance range parameter
 $a(T_r)$ (scaling IDF): rainfall intensity for given T_R and for the reference duration (usually $D = 1h$)
 $c(h), c(\gamma)$: codimension function
 c_s : parameter of stable distribution

e : natural logarithm, $e \sim 2.71828$
 h : separation distance
 (h) : separation vector
 i : generic loop index
 j : generic loop index
 k : generic loop index
 m : cascade multiplicity
 p : moment order in Generalized Structure Function
 q : moment order in moment analysis
 r : long-range correlation fluctuation
 s : generic spatial coordinate
 t : time coordinate
 u : velocity component
 u_p : threshold of Generalized Pareto Distribution
 \mathbf{x} : location vector
 A : area (km^2)
 C_1 : mean co-dimension of the multi-fractal process
 C_β : fractal dimension of the beta model (intermittency)
 D : Euclidean dimension
 $D(h)$: singularity spectrum of a signal
 D_s : Sampling dimension of multifractal analysis
 D : accumulation duration (hours)
 H : Hurst exponent
 I : rainfall intensity ($mm\ h^{-1}$)
 K : Moment scaling function of Universal Multifractal Model
 L : 1D spatial scale
 N : total number of observations
 $P(x)$: Cumulative Density Function
 R : Rainfall depth (mm)
 T_R : return period $T_R = \frac{1}{1-f}$
 U : large-scale advection velocity in the “frozen field” theory
 X : generic random variate
 Y : generic random variate
 Z : generic random variate
 Z : radar reflectivity

Proportion of monthly rainfall

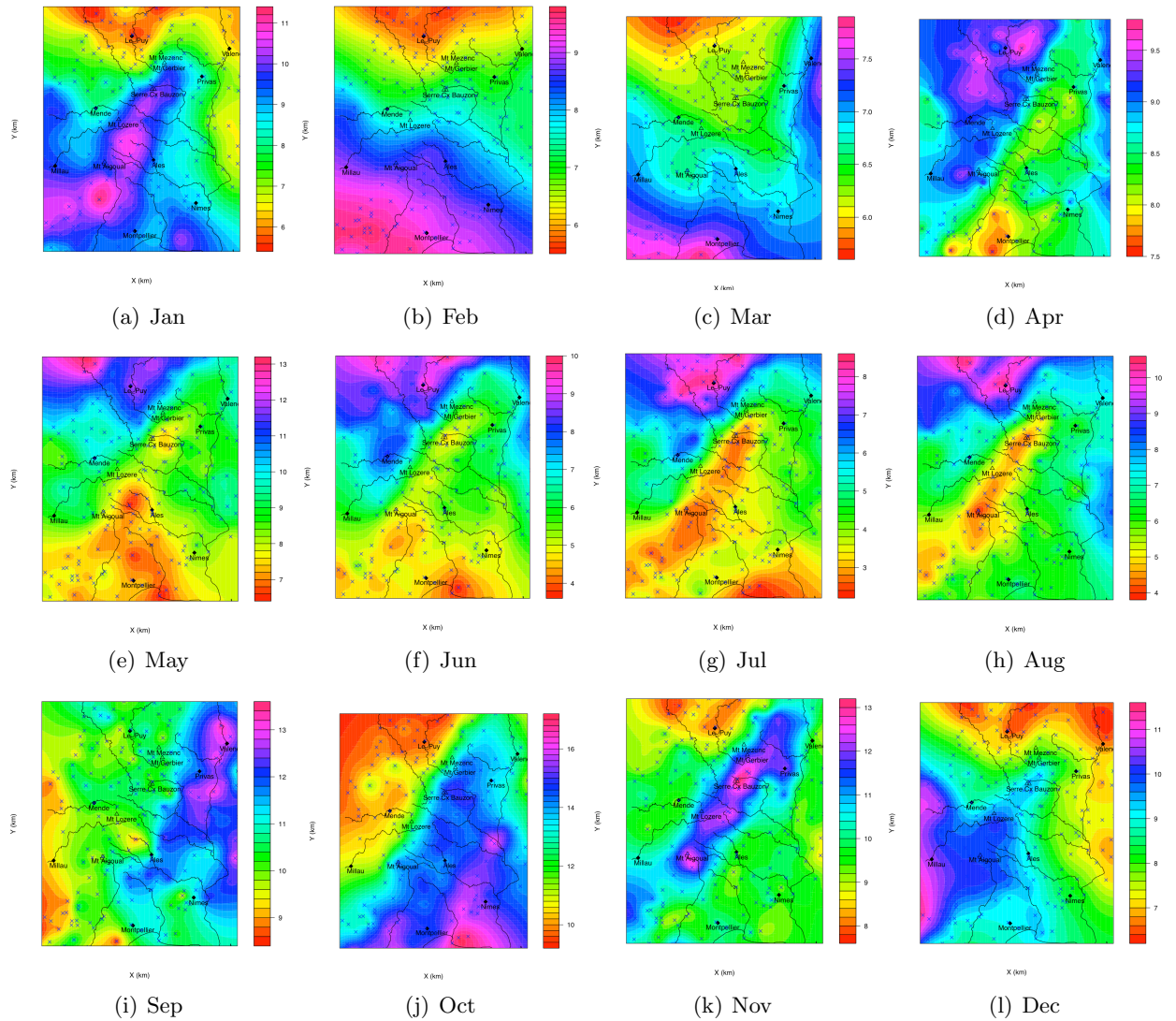


Figure 12.2: Proportion of monthly average rainfall on the total rainfall amount.

Monthly intermittency

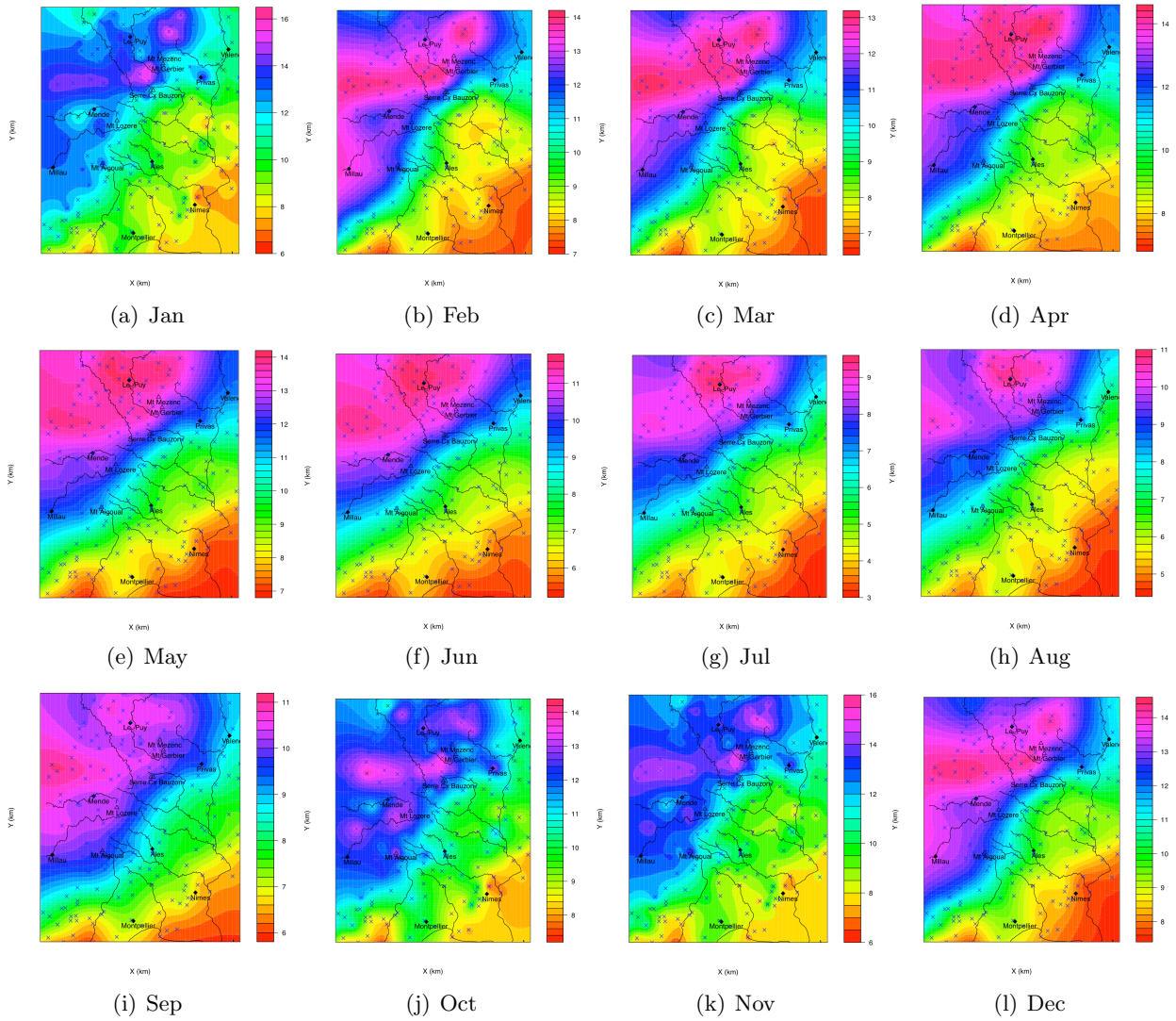


Figure 12.3: Proportion of monthly average rainfall on the total rainfall amount.

APPENDIX C

Extreme value analysis

```
### Packages
library(ismev)
library(VGAM)
### Random Generation of normal, lognormal, gumbel, GEV, GPD variates
rnorm(10000,0,1)
rlnorm(10000,0,1)
rgumbel(100,10,5)
rgev(100,10,5,0.1)
rgpd(100,10,5,0.1)
#### Gumbel (mu=10, sigma=5) quantile, probability and density function
dgumbel(1:100,10,5)
plot(dgumbel(1:100,10,5),type="l")
pgumbel(1:100,10,5)
plot(pgumbel(1:100,10,5),type="l")
qgumbel(seq(0.1,0.99,0.01),10,5)
plot(qgumbel(seq(0.1,0.99,0.01),10,5),type="l")
##### block maxima method
x=rgev(100,10,5,0)
x2=rgumbel(100,10,5)
y=gev.fit(x)
y2=gum.fit(x)
gev.diag(y)
gum.diag(y2)
##### POT method
##### autocorrelation real series
c<-rnorm(100) # insert a real series instead of rnorm(100)
acf(c) # choose based on this graph a decorrelation lag
#####
library(POT)
k=rlnorm(10000,0,1)
acf(k)
c2df=data.frame(obs=k,time=1:length(k))
events0 <- clust(c2df, u = 1, tim.cond = 4, clust.max = TRUE) #tim.cond is the decorr. lag
k2=array(0,length(k))
k2[events0[,1]]=events0[,2]
k3=as.numeric(k2)
mrlplot(k3)
gpd.fitrange(k3,umin=5,umax=20,nint=50)
fitting <- gpd.fit(k2,10,npj=365)
gpd.diag(fitting)
##### Return levels for Tr=100yrs
#GEV:
qgev(0.99,10,5,0)
#GPD
npj=length(k2[k2>10])/(length(k2)/365)
p100=rp2prob(retper = 100, npj = npj)
qgpd(p100[3],10,6,0)
#####
```

APPENDIX D

1D kriging example

```
library(gstat)
# Definition of point vectors (X,Y and the observed value Z)
x=seq(1,131,5)
y=x*0 #1D case
z=c( 10.311789, 9.930269, 15.175008, 14.258478, 6.690640, 7.455567,
    10.017836, 12.011963, 17.712395, 33.429268, 52.768895, 88.653022, 146.218722,
    173.169668, 147.796973, 135.701460, 140.147641, 142.966831
    ,117.150382, 70.210415, 24.394139, 3.971373, 4.621377, 5.568183,
    6.917091, 13.136782, 17.582994)
# creation of the interpolation grid
data=data.frame(x=x,y=y,z=z)
grid=expand.grid(x=seq(0.5,131.5,2),y=0)

# Sample variogram
v=variogram(z~x,~x+y,data,width=1)
# Variogram fitting
v2=fit.variogram(v,vgm(0,"Gau",30,1))

# Ordinary kriging
kri=krige(z~1,~x+y,model=v2,data=data,newdata=grid)
# Generation of 10 Conditional Simulations
kri3=krige(z~1,~x+y,model=v2,data=data,newdata=grid,nsim=10)
# Impose nugget=0 in variogram
v3=v2
v3$psill[1]=0
# Kriging with new variogram
kri2=krige(z~1,~x+y,model=v3,data=data,newdata=grid)
# plot of first kriging (red), second kriging (blue), experimental points (dots).
plot(kri$x,kri$var1.pred,type="l",col="red",lwd=2,ylim=c(0,170),xlab="Distance X (km)",
ylab=expression(paste("Rainfall intensity (mm h-1,)")))
points(kri2$x,kri2$var1.pred,type="l",col="blue",pch=19,lwd=2)
points(x,z,pch=19)

# plot of the conditional simulation results: experim. points (black), cond. simulations (black thin lines), kriging interp (red) and st.dev (green).
plot(x,z,pch=19,ylim=c(-20,170),xlab="Distance X (km)",ylab=expression(paste("Rainfall intensity (mm h-1,)")))
points(kri3$x,kri3$sim1,pch="19",cex=0.2,type="l",lwd=0.2)
points(kri3$x,kri3$sim2,pch="19",cex=0.2,type="l",lwd=0.2)
points(kri3$x,kri3$sim3,pch="19",cex=0.2,type="l",lwd=0.2)
points(kri3$x,kri3$sim4,pch="19",cex=0.2,type="l",lwd=0.2)
points(kri3$x,kri3$sim5,pch="19",cex=0.2,type="l",lwd=0.2)
points(kri3$x,kri3$sim6,pch="19",cex=0.2,type="l",lwd=0.2)
points(kri3$x,kri3$sim7,pch="19",cex=0.2,type="l",lwd=0.2)
points(kri3$x,kri3$sim8,pch="19",cex=0.2,type="l",lwd=0.2)
points(kri3$x,kri3$sim9,pch="19",cex=0.2,type="l",lwd=0.2)
points(kri3$x,kri3$sim10,pch="19",cex=0.2,type="l",lwd=0.2)
points(x,z,pch=19)
points(kri$x,kri$var1.pred,type="l",col="red",lwd=2)
points(kri$x,kri$var1.pred+sqrt(kri$var1.var),lwd=2,type="l",col="green")
points(kri$x,kri$var1.pred-sqrt(kri$var1.var),lwd=2,type="l",col="green")
```

APPENDIX E

Conditioned simulation

```
#### Create 100 fields 100x100 km, resolution 1 km, spherical covariance range 20 km.
library(RandomFields)
x<-y<-1:100
ms0 <- exp(GaussRF(x, y, grid=TRUE, n=100,model="spherical", param=c(0,1,0,20),method="circ"))
##### Variogram (verification)
d=EmpiricalVariogram(x,y, data=ms0, grid=TRUE, bin=seq(0,40,1))
plot(d$centers,d$emp.vario)
```

APPENDIX F

Derivation of the codimension function $c(\gamma)$

The Universal Multifractal Model can be expressed in terms of the singularity order γ , as a function either of the moment order q or of the singularity order γ . Step by step, we substitute Equation 5.18 for the case $\alpha \neq 1$ and $H = 0$ in Equation 5.19 and develop:

$$\gamma(q) = \frac{dK}{dq} = \frac{C_1}{\alpha - 1} (\alpha q^{\alpha-1} - 1) \quad (\text{F1})$$

Knowing from the functional equation of Legendre transform (also in *Veneziano and Furcolo (2002)*) that $c(q) = q\gamma(q) - K(q)$, we substitute the values of $\gamma(q)$ and $K(q)$ obtaining

$$c(q) = \frac{dK}{dq} = \frac{C_1}{\alpha - 1} (\alpha q^{\alpha-1} - 1) \quad (\text{F2})$$

The inversion of Equation D1 yields

$$q(\gamma) = \left(\frac{1}{\alpha} + \frac{\gamma(\alpha - 1)}{\alpha C_1} \right)^{\frac{1}{\alpha-1}} \quad (\text{F3})$$

and we can insert Equation D3 into D2, obtaining $c(\gamma)$:

$$c(\gamma) = C_1 \left(\frac{1}{\alpha} + \frac{\gamma}{C_1 \alpha'} \right)^{\alpha'} \quad (\text{F4})$$

where $\alpha' = \frac{\alpha}{\alpha-1}$.

APPENDIX G

Statistical moments computation of 1D process

```
library(e1071)
##### 1. Definition of the scales and moment values
magni=trunc(log(length(c2)/4,2)) # maximum analyzed scale
size=2(0:magni) # list of the aggregation sizes
scale=size #
lambda=max(scale)/scale # definition of the scale lambda
q=seq(0.5,2.5,0.5) # analyzed moments (> 0.5 because of the measure resolution, <2.5 to have finite moments)
eta=1 # In double trace moments eta can be a sequence from 1.2 to 2.5
mom2=array(0,c(length(q),length(size),length(eta))) # array of moments
momzero=array(0,c(length(q),length(size),length(eta))) # array of positive-rainfall moments
zero_prop=array(0,length(scale)) # proportion of zeros vector

##### 2. Moments computation
for (i6 in 1: length(eta))
{ print("Elaborating moments: eta value=")
etaval=eta[i6]
print(etaval)
c3=((c2)^etaval)/mean((c2)^etaval,na.rm=TRUE) # DTM theory: the field is raised to the eta power and averaged
for (i in 1: length(size))
{ print(size[i])
for (i3 in 1: size[i])
{
dime=trunc(length(c2)/size[i])+1
c2[(length(c2)+1):(dime*size[i])]=NA
w=colMeans(matrix(c2,size[i],length(c2)/size[i])) #fasten the aggregation process
for (i4 in 1: length(q))
{
mom2[i4,i,i6]=mom2[i4,i,i6]+moment(w,order=q[i4],na.rm=TRUE)/size[i]
momzero[i4,i,i6]=momzero[i4,i,i6]+moment(w[w>0],order=q[i4],na.rm=TRUE)/size[i]
}
}
if (eta[i6]==1)
{zero_prop[i]=zero_prop[i]+length(sort(w))/length(sort(w[w>0]))/size[i]}
}
}
}
Cb2=coef(lm(log10(zero_prop)~log10(lambda))) # zero computation
Cb=Cb2[2]
print("Fractal intermittency exponent:")
print(Cb)

##### 3. Print plot
plot(scale,mom2[length(q),,which(eta==1)],pch=19,ylim=c(min(mom2[, ,which(eta==1)],na.rm=TRUE),max(mom2[, ,which(eta==1)],
na.rm=TRUE)),cex=0,log="xy",xlab="Aggregation Scale",ylab="Raw Moments")
for (i in 1:(length(q)))
{points(scale,mom2[i, ,which(eta==1)],cex=0.7,pch=19+i)}
dev.off()
```

APPENDIX H

Trace Moment Method

For the two following example, the results of the Appendix G are used, in particular the matrix *mom2*, the moment vectors *q* and *η* and the scale vector *λ* .

```
mom=mom2
K=matrix(NA,length(q),length(eta))
for (i5 in 1: length(eta))
{
  for (j5 in 1: length(q))
  {
    a=coef(lm(log10(mom[j5,,i5])~log10(lambda)))
    #abline(lm(log10(mom[j5,,i5])~log10(lambda)))
    K[j5,i5]=a[2]#####q[j5]-1
    #points(lambda,mom[j5,,i5])
  }
}
# dev.off()
#xyz4=xyz[xyz$Kz==1,]
azz<-coef(nls(K[,which(eta==1)]~(C1*(q^alpha-q)/(alpha-1)),
start=list(C1=0.025,alpha=1.8),upper=list(C1=1,alpha=2),lower=list(C1=0.0001,alpha=0.0001),algorithm="port",trace=TRUE))
postscript(paste("/Users/davideceresetti/Documents/results/multifractal/TMb_wind_Kq_",contatore, ".eps",sep=""),width=6,height=6)
plot(q,K,xlab="Moment Order q",ylab="Moment scaling function K(q)",xlim=c(0,2.5))
print("TM COEFFICIENTS")
print(azz)
w=seq(0,2.5,0.1)
points(w,azz[1]*(w^azz[2]-w)/(azz[2]-1),type="l",lwd=1,lty=1)
dev.off()
alphaDTM=azz[2]
C1=azz[1]
```

Double Trace Moment Method

```
alpha_DTM=0
slope=0
#####NORMAL DTM
polyinflex=0
plot(eta,abs(K[length(q),]),log="xy",ylim=c(min(abs(K)),max(abs(K))),xlab=expression(paste(eta)),ylab=expression(paste("K(q, ",eta,")",sep="")))
slope=0
qq=0
for (i in 1: length(q[q>1.1]))
{print(length(q)-i+1)
points(eta,abs(K[length(q[q>1.1])-i+1,]))
qq[i]=q[length(q[q>1.1])-i+1]
alpha_DTM[i]=(max(diff(log10(abs(K[length(q)-i+1,]))))/diff(log10(eta))))
wcoeff=coef(lm(log10((K[length(q)-i+1,]))~I(log10(eta))+I(log10(eta)^2)+I(log10(eta)^3)))

points(eta,10^(wcoeff[1]+wcoeff[2]*log10(eta)+wcoeff[3]*log10(eta)^2-wcoeff[3]*log10(eta)^3),type="l")
poly=summary(wcoeff)#$coefficients[1:4]
polyinflex[i]=-2*wcoeff[3]/6/wcoeff[4]
lyinflex[i]^3,pch=19,col="red")

if ((polyinflex[i]>1.5)&(polyinflex[i]<1.5))
{
  slope[i]=wcoeff[2]+2*wcoeff[3]*polyinflex[i]+3*wcoeff[4]*polyinflex[i]^2
} else
{ slope[i]=wcoeff[2]+2*wcoeff[3]*0+3*wcoeff[4]*0^2}
}
alphaDTM=max(slope,na.rm=TRUE)
```

APPENDIX I

1D Multi-fractal cascade with atom at zero

We report in this section an example of 1D multi-fractal disaggregator of rainfall, according to (*Over and Gupta*, 1996; *Schmitt et al.*, 1998)

```
library(e1071)
library(fBasics)
library(ismev)
##### Initializing values #####
exponent=0.6 ;alpha=2 ;C1=0.1; g=runif(10)

ss=disaggregate(g,24,0.4,2,0.1,method="Micro")
disaggregate<-function(g,lambda,atom,alpha,C1,method="Micro")
{
##### GUIDE
# lambda=maximum disaggregation scale
# atom= fractal exponent of rain-no rain intermittence (plot proportion of dry events as function of scale)
# alpha=levy's alpha stable coefficient (alpha=2 normal cascade, alpha<2 levy's stable cascade)
# C1 mean codimension
# method "Micro"=microcanonical; "Canon"=Canonical
##### 1. Cascade steps definition ##### cascade at least branching number=3 because of intermittence
cont=0
divisor=0
num=lambda
while (num>1)
{ cont=cont+1
div=3
while (num-trunc(num/div)*div>0)
{ div=div+1
}
divisor[cont]=div
num=num/div
if (num==2)
{divisor[cont+1]=2
num=1}
}
disagg=divisor # defines the cascade
##### 2. Begin disaggregation
interm1=1 ;
for (i in 1 : length(g))
{
g3=g[i]
if (g3>0)
{
##### 2a. disaggregation steps
for (i2 in 1: length(disagg))
{
g2=array(0,prod(disagg[1:i2]))
for (i4 in 1:length(g3))
{
##### 2b. Intermittence definition
if (alpha==1) {alpha=0.99}
if (i2==1) ##### compute (partial) intermittency in disaggregating the positive observations
{
interm=interm1*disagg[i2]^(-atom)
interm2[i2]=(interm/interm1) #intermitt effettiva sui valori>0
} else {
interm=interm1*prod(disagg[1:i2])^(-atom)
interm2[i2]=(interm1*prod(disagg[1:i2])^(-atom))/(interm1*prod(disagg[1:(i2-1)])^(-atom) #intermitt effettiva sui valori>0
}
##### end 2b
}
}
##### 2c. levy-stable parameters
logvar=C1*2*log(disagg[i2])
if (C1>0) {sigma=C1^(1/alpha)*log(disagg[i2])^((1-alpha)/alpha)*(cos(3.141593*alpha/2)/(1-alpha))^(1/alpha) ;beta=-1 ; mu=C1/(1-alpha) } else {alpha=0.1}
##### 2d. random generator intermittence
ciao=runif(disagg[i2])
ciao[ciao<(1-interm2[i2])]=0 # aggiungo alla probabilita' anche il fatto che spesso ho tutti valori nulli e riestrarre mi distorce l'estimatore
ciao[ciao>(1-interm2[i2])]=1
while (length(ciao[ciao>0])==0)
{ ciao=runif(disagg[i2])
ciao[ciao<(1-interm2[i2])]=0
ciao[ciao>(1-interm2[i2])]=1 }

gg=length(ciao[ciao>0])
if (gg==0)
{gg=1
```

```

ciao[1]=1}
k=c(exp(rstable(gg,alpha=alpha,beta=beta,gamma=sigma,delta=mu,pm=2))) # in stable distr
##### 2e. if k explode is possible regenerate
while (length(which(is.na(k)))>0)
{print("K diverges: rigeration")
if (length(which(is.na(k)))>0)
{ k=c(exp(rstable(gg,alpha=alpha,beta=beta,gamma=sigma,delta=mu,pm=2))) #stable distr (not used)
}
}
while (sum(abs(k))==Inf)
{ k=c(exp(rstable(gg,alpha=alpha,beta=beta,gamma=sigma,delta=mu,pm=2))) #stable distr(not used)
}
ciao[ciao>0]=k
##### 2f. normalization of positive rainfall to balance the dry periods
if (sum(ciao)>0)
{ k=ciao*disagg[i2]/gg
}else{k=c(1,1,1,1)} # in case of error, split exactly the rainfall
##### 2g. Microcanonical cascade option
if (method=="Micro"){
k=k/mean(k) } # microcanonical cascade
for (i3 in 1: disagg[i2])
{g2[((i4-1)*disagg[i2])+i3]=g3[i4]*k[i3]
}
}
g3=g2 }
} else {g3=array(0,prod(disagg)) }
##### end disaggregation
##### 3. store new disaggregation in the vector g_fine at fine resolution
if (i==1)
{g_fine=g3} else {g_fine=c(g_fine,g3)}
}

return(g_fine) # output variable
}

```